

บทที่ 1

บทนำ

1.1 หลักการและเหตุผล

ในปัจจุบันการใช้ภาษามือเพื่อการสื่อสารได้ถูกนำมาใช้โดยกลุ่มคนที่มีปัญหาทางด้านร่างกาย เช่น ผู้ที่บกพร่องทางการได้ยินหรือผู้ที่บกพร่องทางการสื่อสารได้ ซึ่งการใช้ภาษามือนั้นมีส่วนสำคัญอยู่หลายส่วนเช่น การใช้ท่าทาง การใช้มือ หรือบางประเภทต้องใช้สีหน้าในการบอกอารมณ์และความรู้สึกด้วย จึงเป็นการยากที่ผู้ที่ไม่ได้ฝึกภาษามือจะทำการเข้าใจ

ด้วยเหตุผลดังกล่าว จึงทำให้การที่ผู้ที่ไม่ได้ฝึกภาษามือไม่เข้าใจว่าการใช้ท่าทางเหล่านี้แปลว่าอะไร แต่ในปัจจุบัน มีเทคโนโลยีที่ตอบสนองความต้องการของมนุษย์มากขึ้น ทำให้เราสามารถใช้ชีวิตได้สะดวกสบายมากยิ่งขึ้น เทคโนโลยีทางด้านปัญญาประดิษฐ์หรือที่เรียกว่า Artificial Intelligence (AI) ได้ถูกนำมาประยุกต์ใช้ในสังคมได้อย่างแพร่หลาย เช่น ใช้ในการตรวจจับท่าทางภาษามือเพื่อใช้ในการสื่อสาร และเพื่อที่จะเป็นประโยชน์ต่อกลุ่มผู้คนที่เหลือในการสื่อสารกับผู้คนที่ปกติทั่วไปให้ได้ง่ายยิ่งขึ้น

ดังนั้น ผู้จัดทำจึงขอเสนอโปรแกรมในการแปลภาษามือ ซึ่งเป็นเดสก์ท็อปแอปพลิเคชัน ที่ผู้ใช้สามารถทำความเข้าใจภาษามือได้โดยง่าย โดยให้ผู้ใช้งานแสดงท่าทางภาษามือผ่านหน้ากล้อง จากนั้นเว็บไซต์จะทำการเปรียบเทียบกับโมเดลการตรวจจับท่าทางภาษามือ ด้วยอัลกอริทึม LSTM (Long Short-Term Memory) ว่าท่าทางเหล่านั้นมีความหมายว่าอะไร ด้วยเว็บไซต์นี้ไม่ใช่แค่จะทำให้ผู้ที่ไม่ได้ฝึกฝนภาษามือสามารถเข้าใจภาษามือได้ง่ายขึ้นเท่านั้น แต่ยังทำให้กลุ่มคนที่มีปัญหาทางด้านร่างกาย เช่น ผู้ที่บกพร่องทางการได้ยินหรือผู้ที่บกพร่องทางการสื่อสารนั้น สามารถสื่อสารกับผู้ที่ไม่ได้ฝึกฝนภาษามือได้ง่ายขึ้นอีกด้วย

1.2 วัตถุประสงค์ของโครงการ

โครงการนี้จัดทำเพื่อพัฒนาโมเดลตรวจจับภาษามือ ด้วยเทคโนโลยี Deep Learning โดยจัดทำโครงการนี้ขึ้นในรูปแบบของเดสก์ท็อปแอปพลิเคชัน ที่สามารถใช้ได้บนคอมพิวเตอร์และสามารถใช้งานได้ทุกเวลา

1.3 ขอบเขตของโครงการ

- 1.3.1 สามารถเริ่มต้นใช้งานโปรแกรมได้เลย โดยไม่ต้องสมัครสมาชิก
- 1.3.2 การจะนำเข้าข้อมูลท่าทางภาษามือเข้าสู่โปรแกรมได้ ผู้ใช้จำเป็นที่จะต้องมียกกล้องเว็บแคม
- 1.3.3 สามารถตรวจจับและประมวลผลท่าทางภาษามือได้ในทันที
- 1.3.4 ผลลัพธ์ที่ได้จะอยู่ในรูปแบบข้อความ มีเฉพาะข้อความที่ตรงกับรูปแบบท่าทางที่มีในผลเฉลยเท่านั้น
- 1.3.5 ชุดข้อมูลที่ใช้ในการเรียนรู้

1.3.5.1 ชุดข้อมูล ประกอบด้วยท่าทาง 20 ท่าทาง ซึ่งเป็นท่าทางที่คาดว่าจะถูกใช้ใน ชีวิตประจำวันบ่อยครั้ง ได้แก่

- 1) สวัสดี
- 2) ขอบคุณ
- 3) ขอโทษ
- 4) สบายดี
- 5) โชคดี
- 6) ช่วยด้วย
- 7) หิว
- 8) ชอบ
- 9) ไม่ชอบ
- 10) รัก
- 11) โกรธ
- 12) เศร้า
- 13) กิน
- 14) ลืม
- 15) ป่วย
- 16) ง่าย
- 17) ยาก
- 18) ผู้ชาย
- 19) เข้าใจ
- 20) คุณ

โดยแบ่งออกเป็น

- Training Data อย่างน้อย 80 วิดีโอต่อท่าทาง ซึ่งเก็บมาจากผู้ใช้ที่แตกต่างกัน
- Testing Data อย่างน้อย 20 วิดีโอต่อท่าทาง ซึ่งเก็บมาจากผู้ใช้ที่แตกต่างกัน

ตารางที่ 1.1 ชุดข้อมูลเรียนรู้ [1]

ความหมาย	ท่าทาง
สวัสดี	
ขอบคุณ	

ตารางที่ 1.1 ชุดข้อมูลเรียนรู้ [1] (ต่อ)

ความหมาย	ท่าทาง	
ขอโทษ		
สบายดี		
โชคดี		
ช่วยด้วย		
หิว		
ชอบ		
ไม่ชอบ		

ตารางที่ 1.1 ชุดข้อมูลเรียนรู้ [1] (ต่อ)

ความหมาย	ท่าทาง	
รัก		
โกรธ		
เศร้า		
กิน		
ลืม		
ง่าย		
ยาก		

ตารางที่ 1.1 ชุดข้อมูลเรียนรู้ [1] (ต่อ)

ความหมาย	ท่าทาง
ผู้ชาย	 
เข้าใจ	 
คุณ	 

1.3.5.1 ชุดข้อมูลจากฐานข้อมูลภาษามือไทย จากเว็บไซต์ th-sl.com [1] และจัดทำขึ้นเองโดยเก็บมาจากผู้ใช้ที่แตกต่างกัน และแต่ละวิดีโอกำหนดให้มี Frame per second (FPS) เป็น 30 FPS และมีความยาวไม่เกิน 4.00 วินาที

1.3.5.2 ตัวคนหรือผู้ใช้ในวิดีโอ ต้องไม่ใส่ถุงมือ

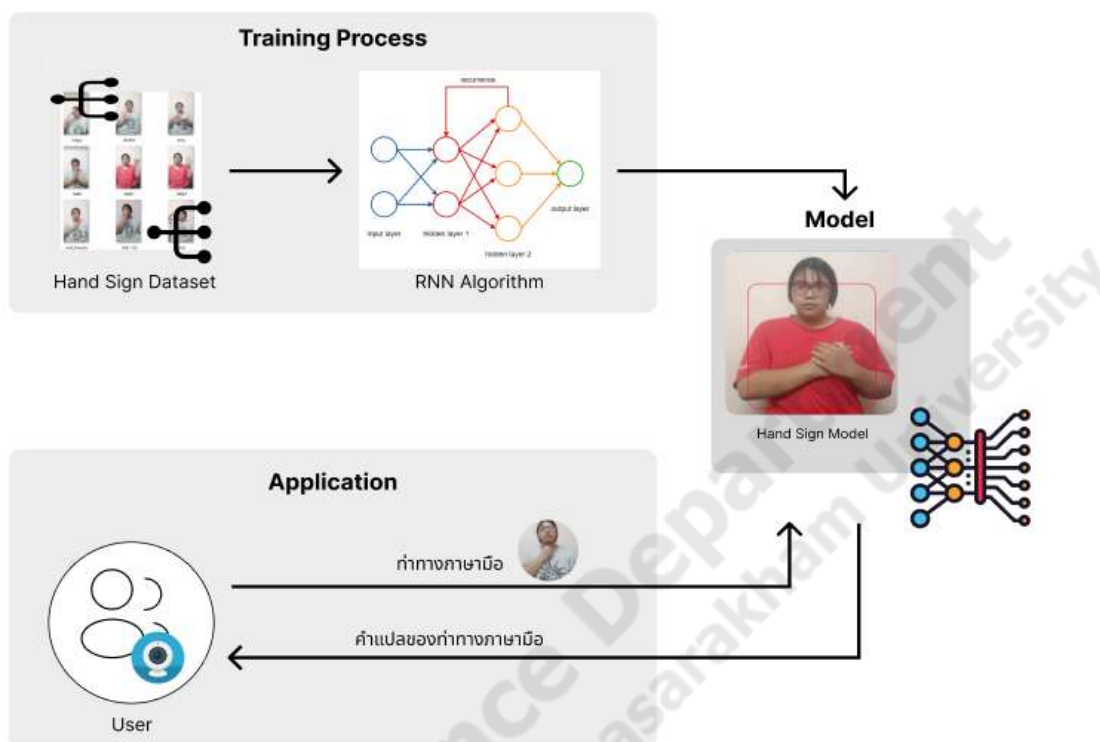
1.3.5.3 ระยะห่างระหว่างกล้องเว็บแคมเรากับคนหรือผู้ใช้ ต้องห่างกันประมาณ 60 เซนติเมตรถึง 150 เซนติเมตร เพื่อให้สามารถเห็นหน้าและมือของผู้ใช้ได้อย่างชัดเจน

1.3.6 ใช้อัลกอริทึม Long Short-Term Memory (LSTM) เพื่อใช้ในการสร้างโมเดลสำหรับทำนาย

1.3.7 การวัดประสิทธิภาพของโมเดล

1.3.7.1 วัดประสิทธิภาพโมเดลโดยใช้ค่า Accuracy โดยกำหนดให้มีค่า 70% ขึ้นไป

1.4 ภาพรวมของระบบ



ภาพประกอบที่ 1.1 ภาพรวมของระบบ

1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1.5.1 สามารถตรวจจับท่าทางภาษามือและแสดงผลพร้อมคำแปลได้อย่างถูกต้องและรวดเร็ว
- 1.5.2 สามารถสื่อสารกับผู้ที่บกพร่องทางการได้ยินหรือผู้ที่บกพร่องทางการสื่อสารได้ง่ายมากยิ่งขึ้น

1.6 อุปกรณ์และเครื่องมือที่ใช้ในการดำเนินงาน

1.6.1 ฮาร์ดแวร์

1) เครื่องคอมพิวเตอร์

Processor:	Intel Pentium G3250 (3.20 GHz)
Graphic System:	NVIDIA GeForce GTX 1050
Main Memory:	8 GB
Hard Disk Drive:	223 GB SSD PCIe M.2
Operating System:	Windows 10(64 bit)

2) คอมพิวเตอร์โน้ตบุ๊ก

Hardware:	Notebook (Acer Nitro)
Processor:	Intel Core i5-9300H (2.40 GHz)

Graphic System:	NVIDIA GeForce GTX 1050 Ti
Main Memory:	8 GB
Hard Disk Drive:	500 GB HDD TOSHIBA DT01ACA050
Operating System:	Windows 10(64 bit)

1.6.2 ซอฟต์แวร์

- Python
- Google Colaboratory
- Visual Studio Code

1.7 แผนการดำเนินงาน

โครงการปริญญาโทฉบับนี้ ดำเนินงาน ณ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคามระหว่างเดือน มิถุนายน 2565 ถึง มีนาคม 2566 ดังที่แสดงในตาราง 1.2

ตารางที่ 1.2 แผนการดำเนินงาน

กิจกรรม	เดือน												
	มิ.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	
1. ศึกษาและรวบรวมข้อมูล	■	■	■										
2. วิเคราะห์และกำหนดขอบเขต		■	■	■									
3. ออกแบบระบบ				■	■	■							
4. พัฒนาโปรแกรม				■	■	■	■	■					
5. ทดสอบระบบ				■	■	■	■	■	■				
6. ทำรายงานสรุป									■				
7. นำเสนอโครงการ										■			

1.8 ตัวอย่าง User Interface ของ Desktop Application

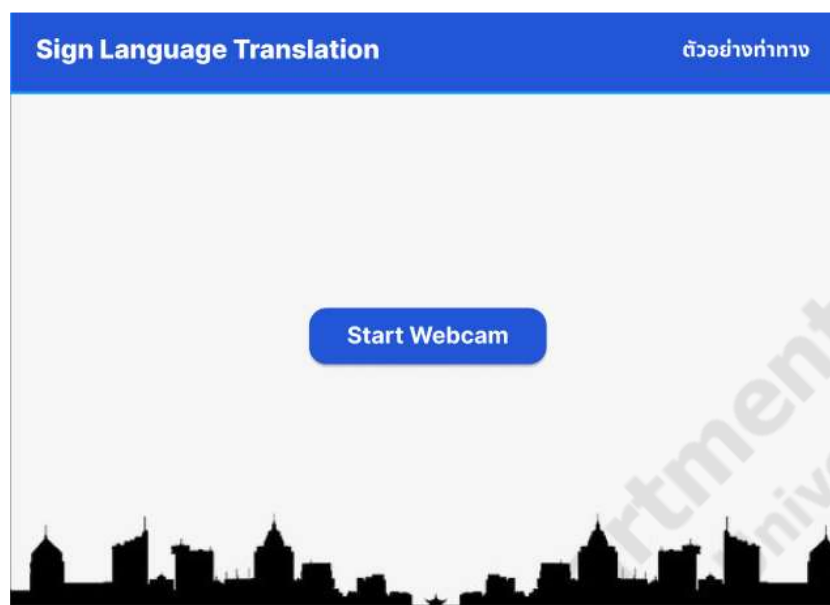
1.8.1 แอปพลิเคชันโดยรวม



ภาพประกอบที่ 1.2 ภาพรวมของแอปพลิเคชัน [1]

1.8.2 หน้าหลักของระบบ

มีปุ่ม Start Webcam เพื่อเปิดเข้าไปใน หน้าเปิดใช้งานกล้อง และมีปุ่มตัวอย่างท่าทาง เพื่อเข้าที่หน้า ตัวอย่างท่าทาง เพื่อที่จะสามารถดูได้ว่า มีท่าทางไหนที่มีผลเฉลยบ้าง



ภาพประกอบที่ 1.3 หน้าหลักของระบบ

1.8.3 หน้าตัวอย่างท่าทาง

ในหน้าตัวอย่างท่าทางจะมีท่าทางทั้งหมดพร้อมบอกความหมายของท่าทางแต่ละท่าทางที่จะสามารถใช้แล้วมีผลเฉลยออกมาได้ โดยเมื่อผู้ใช้กดไปที่วิดีโอ วิดีโอจะทำการขยายภาพขึ้นมาที่ตรงกลางหน้าจอเพื่อให้สามารถมองเห็นได้ง่ายยิ่งขึ้น และผู้ใช้สามารถค้นหาท่าทางได้ เพื่อทำการค้นหาว่าท่าทางที่ผู้ใช้ต้องการจะแสดงนั้นมีอยู่ในระบบหรือไม่



ภาพประกอบที่ 1.4 หน้าตัวอย่างท่าทาง ก [1]



ภาพประกอบที่ 1.5 หน้าตัวอย่างท่าทาง ข [1]

1.8.4 หน้าเปิดใช้งานกล้อง

ในหน้านี้เมื่อผู้ใช้กดที่ปุ่ม Switch จะทำการเปิดกล้อง Webcam ของตัวเครื่องที่ผู้ใช้ใช้อยู่ และหากต้องการที่จะปิดกล้อง ผู้ใช้สามารถกดที่ปุ่มเดิมเพื่อปิดกล้องได้ เพื่อที่จะทำให้มีผลเฉลยออกมา ผู้ใช้จะต้องแสดงท่าทางภาษามือต่างๆ โดยที่ผู้ใช้จะต้องทำการเปิดกล้อง และ ผู้ใช้จะต้องแสดงท่าทางอย่างถูกต้อง และ ท่าทางที่ผู้ใช้แสดงนั้นจะต้องมีอยู่ในหน้าตัวอย่างท่าทางด้วย



ภาพประกอบที่ 1.6 หน้าเปิดใช้งานกล้อง