

บทที่ 1

บทนำ

1.1 หลักการและเหตุผล

ภาษามือ คือ ภาษาสำหรับคนหูหนวก โดยใช้มือ สีสหน้าและกิริยาท่าทางในการประกอบในการสื่อความหมาย และถ่ายทอดอารมณ์แทนการพูด ภาษามือของแต่ละชาติมีความหมายแตกต่างกัน เช่นเดียวกับภาษาพูด ซึ่งแตกต่างกันตามขนบธรรมเนียม ประเพณี วัฒนธรรมและลักษณะภูมิศาสตร์ เช่น ภาษามือจีน ภาษามืออเมริกัน และภาษามือไทย เป็นต้น ภาษามือเป็นภาษาที่นักการศึกษาทางด้านการศึกษาคคนหูหนวกตกลงและยอมรับกันแล้วว่าเป็นภาษาหนึ่งสำหรับการติดต่อสื่อความหมายระหว่างคนหูหนวกกับคนหูหนวกด้วยกัน และระหว่างคนปกติกับคนหูหนวก (bkkthon, 2563: ออนไลน์)

เทคโนโลยีในปัจจุบันมีหลากหลายเทคโนโลยีและมีหลากหลายศาสตร์ที่จะนำมาช่วยแก้ปัญหาให้กับมนุษย์และลดแรงงานของมนุษย์ลง เช่น เทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) ที่เกิดจากการเรียนรู้ของเครื่อง (Machine Learning) การเรียนรู้เชิงลึก (Deep Learning) และ โครงข่ายประสาทเทียม (Neural Networks) โดยได้มีนักวิจัยและพัฒนากระบวนการรู้จำภาษามือด้วยเทคนิคต่าง ๆ เช่น งานวิจัยของ A. Chaikaew, K Somkuan and T. Yuyen (2564) วัตถุประสงค์ของงานวิจัยนี้คือเพื่อพัฒนาแอปพลิเคชันสำหรับการรู้จำภาษามือที่เป็นภาษาไทยแบบเรียลไทม์โดยการใช้ MediaPipe Framework มาช่วยในการสกัดแลนด์มาร์กจากวิดีโอท่าทางภาษามือและใช้แลนด์มาร์กเพื่อสร้างโมเดลสำหรับการรู้จำท่าทางภาษามือด้วย Recurrent Neural Network (RNN) ผลที่ได้จากการวิจัยคือ โมเดลที่สร้างโดย LSTM, BiLSTM และ GRU มีความถูกต้องมากกว่า 90% วิธีนี้สามารถสร้างความแม่นยำได้ใกล้เคียงกับวิธีการแบบดั้งเดิมและงานวิจัยของ Gerages H. Samaan, Abanoub R. Widie, Abanoub K. Attia, Abanoub M. Asaad, Andrew E. Kamel, Salwa O. Slim, Mohamed S. Abdallah and Young-Im Cho (2022) ในงานวิจัยนี้ได้ใช้ MediaPipe ในการเชื่อมเข้ากับ RNN โมเดล เพื่อแก้ปัญหาการรู้จำภาษามืออังกฤษแบบไดนามิก MediaPipe ถูกใช้เพื่อสร้าง Landmarks บนร่างกายแล้วสกัด Keypoints ของมือ ตัวและหน้า ส่วน RNN โมเดล เช่น GRU, LSTM และ BiLSTM ถูกใช้เพื่อการรู้จำภาษามืออังกฤษเนื่องจากไม่มีชุดข้อมูลภาษามือ จึงได้สร้าง DSL 10 Dataset ซึ่งมีคำศัพท์ 10 คำที่ซ้ำกัน 75 ครั้งโดยที่ปรึกษา 5 คนซึ่งให้คำแนะนำขั้นตอนในการสร้างคำศัพท์ดังกล่าว มีการทดลองสองครั้งในชุดข้อมูล DSL 10 Dataset โดยใช้แบบจำลอง RNN เพื่อเปรียบเทียบความแม่นยำของการรู้จำภาษามือแบบไดนามิกที่มีและไม่มี Keypoint ผลการทดลองคือโมเดลมีความแม่นยำมากกว่า 90%

จากที่กล่าวมาข้างต้นผู้วิจัยจึงมีความสนใจที่จะพัฒนาระบบการรู้จำภาษาไทยและท่าทางด้วยเทคนิคโครงข่ายประสาทเทียมแบบวนกลับ โดยสร้างเป็นคำที่ใช้ในชีวิตประจำวัน เพื่อใช้ในการแปลภาษาไทยของผู้พิการทำให้สามารถเข้าใจความหมายที่ต้องการจะสื่อได้

1.2 วัตถุประสงค์

1.2.1 เพื่อพัฒนาระบบการรู้จำท่าทางภาษาไทยด้วยโครงข่ายประสาทเทียมแบบวนกลับ

1.2.2 เพื่อประเมินประสิทธิภาพระบบการรู้จำท่าทางภาษาไทยด้วยโครงข่ายประสาทเทียมแบบวนกลับ

1.3 ขอบเขตและข้อตกลงเบื้องต้นของการวิจัย

1.3.1 การรวบรวมข้อมูล

ข้อมูลที่ใช้ในการศึกษาครั้งนี้ ผู้วิจัยได้นำคำศัพท์ท่าทางต่าง ๆ ของภาษาไทยมาจากเว็บไซต์ highlightkapook.com ที่มีเนื้อหาเสนอตัวอย่างภาษาไทยที่ใช้ในชีวิตประจำวันของผู้พิการจำนวน 10 คำ โดยที่ผู้วิจัยได้คัดเลือกมาเป็นจำนวน 8 คำ ได้แก่ สบายดี ชอบ รัก ป่วย หิว ขอบคุณ ขอโทษ สวัสดี และอีก 2 คำจะเป็นที่ใช้สำหรับเรียกแทนบุคคล เช่น ฉัน คุณ เพื่อที่จะทดลองสร้างการรู้จำภาษาไทยที่เป็นคำและสามารถเป็นประโยคได้ เช่น ฉันรักคุณ ฉันสบายดีขอบคุณ คุณขอบฉัน ฉันขอโทษ ฉันหิว เป็นต้น



คำภาษาไทย : ขอบคุณ
คำภาษาอังกฤษ : Thank you
ความหมาย : กล่าวแสดงความรู้สึกถึงบุญคุณหรือกล่าวเมื่อได้รับความช่วยเหลือ
การใช้ภาษามือ : แขนงทั้ง 2 ข้าง โดยให้ทุกนิ้วชิดกัน จากนั้นยกมือขึ้นมาชนกันในแนวตั้ง แล้วจึงดึงมือทั้งสองข้างออกจากกัน

ภาพที่ 1.1 ตัวอย่างคำภาษามือไทย

ที่มา : [highlightkapook](http://highlightkapook.com) (2565: ออนไลน์)

ทางผู้วิจัยได้มีการออกหนังสือขอความอนุเคราะห์จากศูนย์บริการสนับสนุนการนักศึกษาพิการระดับอุดมศึกษา (DSS) เพื่อเก็บข้อมูลสำหรับการเทรนโมเดลสำหรับการรู้จำภาษาไทยจากทั้งผู้เชี่ยวชาญภาษามือและผู้พิการที่ใช้ภาษามือเป็นหลักในการสื่อสาร

1.3.2 การประเมินประสิทธิภาพ

1.3.2.1 Confusion Matrix

Confusion Matrix ถือเป็นเครื่องมือสำคัญในการประเมินผลลัพธ์ของการทำนาย หรือ Prediction ที่ทำนายจาก Model ที่สร้างขึ้น ใน Machine learning โดยมีไอดีจากการวัดว่า สิ่งที่คิด (Model ทำนาย) กับ สิ่งที่เกิดขึ้นจริง มีสัดส่วนเป็นอย่างไร

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

ภาพที่ 1.2 ตาราง Confusion Matrix

True Positive (TP) = โมเดลทำนายว่าเป็น Positive และคำตอบก็เป็น Positive จริง

True Negative (TN) = โมเดลทำนายว่าเป็น Negative และคำตอบก็เป็น Negative จริง

False Positive (FP) = โมเดลทำนายว่าเป็น Positive และคำตอบก็เป็น Negative

False Negative (FN) = โมเดลทำนายว่าเป็น Negative และคำตอบก็เป็น Positive

โดย TP, TN, FP, FN ในตารางจะแทนด้วยค่าความถี่ สามารถใช้ Confusion Matrix มาคำนวณ การประเมินประสิทธิภาพของการทำนายด้วย Model ของ ในรูปแบบค่าต่างๆได้หลายค่า (Pagon Gatchalee. 2565: Online)

1.3.2.2 Accuracy

Accuracy (ความถูกต้อง) = $(TPs + TNs) / (TPs + TNs + FPs + FNs)$

หรือกล่าวได้ว่า Accuracy = ผลรวมของตัวเลขบนเส้นทแยงมุมในตาราง Confusion Matrix /

จำนวน Observations ทั้งหมด โดย ความเป็นจริงแล้ว Confusion matrix ไม่จำเป็นต้องเป็นแบบ

2x2 หรือมีผลลัพธ์แค่ 2 แบบ เสมอไป โดยอาจเป็น 3x3, 4x4, nxn ก็ได้ โดยวิธีการหา Accuracy ก็

ใช้แบบเดิม คือ ผลรวมของตัวเลขบนเส้นทแยงมุมในตาราง Confusion Matrix / จำนวน

Observations ทั้งหมด (Pagon Gatchalee. 2565: ออนไลน์)

1.4 ข้อตกลงเบื้องต้น

1.4.1 เทคนิคหรือเทคโนโลยีที่ใช้

- 1.4.1.1 การเรียนรู้เชิงลึก (Deep Learning)
- 1.4.1.2 หน่วยความจำระยะสั้นยาว (Long Short-Term Memory: LSTM)
- 1.4.1.3 หน่วยเกทแบบวนกลับ (Gated Recurrent Unit: GRU)
- 1.4.1.5 หน่วยความจำระยะสั้นยาวแบบสองทิศทาง (Bidirectional Long Short-Term Memory: BiLSTM)

1.4.2 เครื่องมือวิจัย

- 1.4.2.1 Tensorflow
- 1.4.2.2 OpenCV
- 1.4.2.3 Midaipipe
- 1.4.2.4 Keras

1.4.3 เครื่องมือที่ใช้ในการพัฒนา

- 1.4.3.1 ภาษาคอมพิวเตอร์
 - ภาษา Python
- 1.4.3.2 ซอฟต์แวร์
 - โปรแกรม Anaconda
- 1.4.3.3 ฮาร์ดแวร์
 - เครื่องคอมพิวเตอร์ Notebook ที่ใช้ทำโครงงาน หน่วยประมวลผล AMD

Ryzen 5 4600H with Radeon RX Graphics หน่วยความจำหลัก (SSD): 512 GB หน่วยความจำชั่วคราว (RAM): 20 GB ระบบปฏิบัติการ (OS): Windows 11 64-bit

1.4.4 วิธีการดำเนินงาน

- 1.4.1 กำหนดหัวข้อและนำเสนอหัวข้อ
- 1.4.2 ค้นหาปัญหา โอกาสและเป้าหมาย
- 1.4.3 ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง
- 1.4.4 เสนอเค้าโครงงาน
- 1.4.5 ทำความเข้าใจข้อมูลและเตรียมข้อมูล
- 1.4.6 ดำเนินการพัฒนาโมเดล
- 1.4.7 ประเมินประสิทธิภาพการพัฒนาโมเดล
- 1.4.8 จัดทำเอกสารประกอบโครงงาน
- 1.4.9 นำเสนอโครงงานจบ
- 1.4.10 รายงานด้วยเล่มสมบูรณ์

1.4.5 แผนการดำเนินการ

ตารางที่ 1.1 ระยะเวลาการดำเนินงาน

กิจกรรม	ระยะเวลาในการดำเนินงาน (พ.ศ. 2565 - 2566)								
	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.
1. กำหนดหัวข้อและนำเสนอหัวข้อ									
2. ค้นหาปัญหา โอกาสและเป้าหมาย									
3. ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง									
4. เสนอเค้าโครงงาน									
5. ทำความเข้าใจข้อมูลและเตรียมข้อมูล									
6. ดำเนินการพัฒนาโมเดล									
7. ประเมินประสิทธิภาพการพัฒนาโมเดล									
8. จัดทำเอกสารประกอบโครงงาน									
9. นำเสนอโครงงานจบ									
10. รายงานด้วยเล่มสมบูรณ์									

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ได้ระบบการรู้จำท่าทางภาษามือไทยด้วยโครงข่ายประสาทเทียมแบบวนกลับ

1.5.2 สามารถต่อยอดเป็นแอปพลิเคชันแปลภาษามือไทยของผู้พิการได้ในอนาคต