การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกเพื่อการสื่อสารระหว่างนิสิตครูกับนักเรียนที่มีความต้องการพิเศษ: การประยุกต์โมเดลซีเอ็นเอ็นร่วมกับอาร์เอ็นเอ็น

Development of Thai sign language recognition model using deep learning for communication between student teachers and student with special needs: application of CNN and RNN models

ธีรวิชญ์ คะรานรัมย์[1*] และ สิวะโชติ ศรีสุทธิยากร[2]
Teerawit Karanram[1*] and Siwachoat Srisuttiyakorn[2]
(Received: October 17, 2024; Revised: November 29, 2024; Accepted: December 2, 2024)

บทคัดย่อ

การวิจัยนี้มีจุดมุ่งหมาย 1) เพื่อวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน และ 2) เพื่อพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก ตัวอย่างวิจัย ได้แก่ ครูการศึกษาพิเศษในโรงเรียนสอนคนหนูหนวก และอาสาสมัครที่มีความสามารถในการใช้ภาษามือ เครื่องมือวิจัย ได้แก่ แบบสอบถามความถี่การใช้งานประโยคภาษามือ การวิเคราะห์ข้อมูลแบ่งออกเป็น 2 ตอน ตอนที่ 1 การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียนจากแบบสอบถามโดยใช้ค่าเฉลี่ยและร้อยละ ตอนที่ 2 การพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก ผู้วิจัยเลือกใช้โมเดลการเรียนรู้เชิงลึก convolutional neural network (CNN) ร่วมกับ recurrent neural network (RNN) จำนวน 4 โมเดล ในการพัฒนาโมเดลรู้จำภาษามือไทยจำนวน 20 ประโยค โดยใช้ข้อมูลจากวิดีโอที่ถ่ายทำกับอาสาสมัคร จากนั้นเปรียบเทียบประสิทธิภาพของโมเดลโดยพิจารณาจากความถูกต้อง ขนาดของโมเดล และเวลาที่ใช้ในการฝึกสอนโมเดล เพื่อนำโมเดลที่มีประสิทธิภาพสูงสุดไปใช้ประเมินประสิทธิภาพของโมเดลในการจำแนกประโยคภาษามือในภาพรวมและในแต่ละประโยคโดยใช้ค่าความถูกต้อง ค่าความแม่นยำ ค่าการเรียกคืน และคะแนน F1

ผลการวิจัยพบว่า 1) ภาษามือที่ใช้ในชั้นเรียนมักเป็นประโยคที่สั้นและกะทัดรัดมากกว่าภาษาพูดและภาษามือที่ใช้มากที่สุดจัดอยู่ในหมวดการเสริมแรง 2) การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้โมเดล EfficientNetB0 ร่วมกับ GRU สามารถจำแนกประโยคภาษามือได้มีประสิทธิภาพสูงสุดทั้งในด้านความถูกต้อง

[1] นิสิตปริญญาโท สาขาวิชาวิธีวิทยาการพัฒนานวัตกรรมทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
[2] ผู้ช่วยศาสตราจารย์ ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
[1] Master's degree student, Methodology for Innovation Development in Education Program, Faculty of Education, Chulalongkorn University
[2] Assistant Professor, Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University
* Corresponding Author E-mail: teerawit.karanram@gmail.com

ขนาดของโมเดล และเวลาที่ใช้ในการฝึกสอนโมเดล และ 3) การประเมินประสิทธิภาพของโมเดลรู้จำภาษามือไทยโดยใช้ข้อมูลชุดทดสอบ โมเดลมีความถูกต้องร้อยละ 73 โดยประโยคที่โมเดลสามารถจำแนกได้ดี ได้แก่ ภูมิใจ พร้อมยัง ดูภาพ และเข้าใจ และประโยคที่โมเดลจำแนกได้ไม่ดี ได้แก่ งานสวย ทำได้ไหม และถูกหรือผิด

คำสำคัญ การรู้จำภาษามือ การเรียนรู้เชิงลึก ภาษามือไทย

Abstract

This research aims to 1) analyze the Thai sign language used by teachers to communicate with students with special needs in the classroom, and 2) develop and evaluate the performance of a Thai sign language recognition model using deep learning. The research subjects include special education teachers in schools for the deaf and volunteers proficient in sign language. The research tools include a questionnaire on the frequency of sign language sentences. Data analysis is divided into two parts: Part 1 involves analyzing the Thai sign language used by teachers to communicate with students with special needs in the classroom using mean and percentage from the questionnaire data. Part 2 involves developing and evaluating the performance of a Thai sign language recognition model using deep learning. The researcher selected four deep learning models combining CNN and RNN to develop a model that recognizes 20 Thai sign language sentences using video data recorded with volunteers. The models' performance was compared in terms of accuracy, model size, and training time, to identify the most effective model for evaluating the overall and individual sentence recognition using accuracy, precision, recall, and F1-score.

The research findings revealed that 1) the sign language used in the classroom tends to be shorter and more concise than spoken language, with reinforcement phrases being the most commonly used; 2) the development of a Thai sign language recognition model using the EfficientNetB0 combined with GRU achieved the highest performance in terms of accuracy, model size, and training time; and 3) the evaluation of the Thai sign language recognition model's performance using a test dataset showed that the model achieved an accuracy of 73%. Sentences that the model recognized well included "proud," "are you ready?" "look at the picture," and "understand," while sentences it struggled with included "beautiful work," "can you do it?" and "right or wrong."

Keywords: sign language recognition, deep learning, Thai sign language

บทนำ

ภาษามือเป็นภาษาที่ใช้ในการสื่อสารระหว่างคนหูหนวกด้วยกันเองและคนที่มีความสามารถในการแปลภาษามือ ซึ่งอาศัยองค์ประกอบของภาษามือ ได้แก่ ท่ามือ ตำแหน่งของมือ การเคลื่อนไหวของมือ การหันฝ่ามือ และการแสดงออกทางสีหน้า เพื่อใช้ในการสื่อสารแทนเสียงพูด อย่างไรก็ตามภาษามือของแต่ละชุมชนหรือท้องถิ่นจะมีลักษณะแตกต่างกันไปตามวัฒนธรรม ขนบธรรมเนียมประเพณี และลักษณะทางภูมิศาสตร์ เช่นเดียวกับภาษาพูดที่มีคำศัพท์ความหมายเหมือนกันแตกต่างกันในแต่ละท้องถิ่น (Boonya, 2008; Buangamdee, 2017; Petchporee, 2005)

ภาษามือนอกจากจะเป็นเครื่องมือการสื่อสารของคนหนูหนวกแล้ว ภาษามือยังเป็นเครื่องมือสำคัญในการเรียนรู้ของคนหูหนวก การสื่อสารระหว่างครูและนักเรียนหูหนวกเป็นปัจจัยสำคัญต่อผลลัพธ์การเรียนรู้ โดยเฉพาะการจัดการเรียนรู้แบบเรียนรวม (inclusive education) การขาดครูการศึกษาพิเศษหรือครูที่มีความรู้ด้านภาษามืออาจทำให้การสื่อสารระหว่างครูกับนักเรียนไม่มีประสิทธิภาพเท่าที่ควร และทำให้การจัดการเรียนการสอนขาดคุณภาพ (Kelly et al., 2020) ซึ่งครูการศึกษาพิเศษในประเทศไทยมีจำนวนที่จำกัด ถ้าหากมีเครื่องมือหรือตัวช่วยในการรู้จำภาษามือให้กับครูหรือนิสิตนักศึกษาครูที่ไม่มีพื้นฐานภาษามือก็จะช่วยลดช่องว่างในการสื่อสารระหว่างครูกับนักเรียนหูหนวก และช่วยให้นักเรียนมีผลลัพธ์การเรียนรู้ที่ดียิ่งขึ้น

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับภาษามือพบว่า มีการนำเทคโนโลยีเกี่ยวกับการรู้จำภาษามือ (sign language recognition) เข้ามาใช้เป็นเครื่องมือที่ช่วยในการให้ความหมายของท่าภาษามือเป็นข้อความหรือคำพูด โดยรูปแบบของภาษามือที่ถูกใช้ในการศึกษามีอยู่สองรูปแบบ รูปแบบแรกคือการสะกดนิ้วมือ (finger spelling) เป็นท่ามือที่แทนพยัญชนะ สระ และวรรณยุกต์ รูปแบบที่สองคือท่าภาษามือที่แสดงถึงคำศัพท์ วลี หรือประโยค (Buangamdee, 2017) ทั้งนี้ ในการวิจัยนี้จะเป็นการศึกษาและพัฒนาโมเดลในส่วนของการรู้จำประโยคภาษามือไทยพื้นฐานที่ใช้ในห้องเรียน เพื่อให้ได้โมเดลที่สามารถนำไปประยุกต์ใช้เป็นเครื่องมือให้ครูสามารถเรียนรู้เกี่ยวกับภาษามือและสื่อสารกับนักเรียนหูหนวกในชั้นเรียนของตนเองได้

การรู้จำภาษามือต้องอาศัยข้อมูลที่สะท้อนถึงท่าทางของผู้ทำภาษามือ จากการศึกษาพบว่ามีวิธีการที่ใช้ในการรับข้อมูลดังกล่าวอยู่ 2 วิธี (Sharma & Singh, 2023) วิธีแรกคือการใช้เซนเซอร์ (sensor-based approaches) ซึ่งเซนเซอร์นี้จะอยู่ในอุปกรณ์ที่สวมใส่ไปบนร่างกายของผู้ทำภาษามือ ซึ่งในแต่ละการวิจัยเกี่ยวกับการรู้จำภาษามือไทยแต่ละงานจะมีการใช้อุปกรณ์ต่างกัน เช่น Jaruwanawat (2007); Saengsri et al. (2012) ใช้ถุงมืออิเล็กทรอนิกส์ซึ่งเป็นอุปกรณ์ที่จะส่งสัญญาณไฟฟ้าเมื่อเกิดการเคลื่อนไหว ส่วน Amatanon et al. (2014); Vachirapipop et al. (2017) ใช้เครื่องวัดสัญญานไฟฟ้ากล้ามเนื้อ (electromyogram) ซึ่งจะรับสัญญาณโดยตรงจากกล้ามเนื้อเมื่อเกิดการเคลื่อนไหว ทั้งนี้การใช้อุปกรณ์สวมใส่ที่กล่าวมาข้างต้นมีข้อดีคือเซนเซอร์สามารถรับข้อมูลการเคลื่อนไหวได้อย่างแม่นยำ และไม่จำเป็นต้องอาศัยการเตรียมข้อมูลที่ซับซ้อนก่อนนำไปพัฒนาโมเดล แต่อย่างไรก็ตามอุปกรณ์ดังกล่าวมีต้นทุนที่สูง และการสวมใส่อุปกรณ์จะทำให้ความยืดหยุ่นในการทำภาษามือลดลงด้วย ซึ่งจากข้อจำกัดนี้ทำให้ในเวลาต่อมามักไม่ค่อยพบการศึกษาเกี่ยวกับการรู้จำภาษามือโดยการใช้อุปกรณ์ดังกล่าว จากการศึกษางานวิจัยเกี่ยวกับการรู้จำภาษามือไทยในเวลาต่อมา (Chaikaew, 2022; Pariwat & Seresangtakul, 2021; Plianwong & Kangkachit, 2021; Sriboonruang et al., 2022; Vijitkunsawat et al., 2023) พบว่าวิธีการรับข้อมูลจากการทำภาษามืออีกวิธีที่เป็นที่นิยม

มากกว่า คือ การใช้ภาพ (image-based approaches) ซึ่งมีทั้งในรูปแบบของภาพนิ่งและวิดีโอขึ้นอยู่กับภาษามือที่ถูกนำมาใช้ในการศึกษา โดยวิธีนี้มีข้อได้เปรียบคือมีต้นทุนที่ต่ำกว่า และผู้ทำภาษามือสามารถทำท่าทางต่าง ๆ ได้อย่างอิสระ แต่วิธีการนี้ต้องอาศัยเทคนิคในการเตรียมข้อมูลที่มีความซับซ้อน ทั้งนี้ วิธีการที่ผู้วิจัยเลือกมาใช้ในการวิจัยนี้คือการใช้ภาพซึ่งจะใช้ข้อมูลในรูปแบบวิดีโอ เนื่องจากครูและนิสิตนักศึกษาครูที่นำโมเดลไปใช้สามารถใช้กล้องในการรับข้อมูลเพื่อรู้จำภาษามือได้ ซึ่งมีความสะดวกในการใช้งานและประหยัดต้นทุนกว่าการใช้อุปกรณ์สวมใส่

การพัฒนาโมเดลรู้จำภาษามือเพื่อใช้ในการแปลภาษามือนั้นเป็นการประยุกต์ใช้โมเดลการเรียนรู้ของเครื่อง (machine learning) ในการจำแนกท่าทางของผู้ทำภาษามือ จากการศึกษางานวิจัยที่เกี่ยวข้องกับการรู้จำภาษามือไทยโดยการใช้ภาพ พบว่าโมเดลการเรียนรู้ของเครื่องที่ถูกนำมาใช้มีความหลากหลาย โดยแบ่งเป็น2 กลุ่ม กลุ่มแรกคือโมเดลการเรียนรู้ของเครื่องแบบดั้งเดิม เช่น hidden Markov model (Sakulsujirapa, 2011) decision tree (Veerasakulthong, 2007) KNN (Klomsae et al., 2017; Pariwat & Seresangtakul, 2021) อย่างไรก็ตามในกรณีข้อมูลที่เป็นรูปภาพหรือวิดีโอ การพัฒนาโมเดลจะมีเตรียมข้อมูลให้มีความเหมาะสมโดยการสร้างคุณลักษณะ (feature engineering) ซึ่งในการใช้โมเดลการเรียนรู้ของเครื่องแบบดั้งเดิมนั้นจะมีขั้นตอนการสร้างคุณลักษณะที่ยาก และมีข้อจำกัดในการเรียนรู้ที่มีความซับซ้อน จากปัญหาดังกล่าวจึงมีการนำโมเดลกลุ่มที่สองคือโมเดลการเรียนรู้เชิงลึก (deep learning) มาใช้แทนโมเดลการเรียนรู้ของเครื่องแบบดั้งเดิม เช่น artificial neural networks หรือ ANN (Amatanon et al., 2014; Chansri & Srinonchat, 2016) เป็นโมเดลที่ได้แรงบันดาลใจมากจากการทำงานของเซลล์ประสาทในสมองมนุษย์ในการแก้ปัญหาที่มีความซับซ้อน convolutional neural networks หรือ CNN (Gedkhaw, 2022; Pariwat & Seresangtakul, 2021; Sripairojthikoon & Harnsomburana, 2019) เป็นโมเดลที่สามารถดึงคุณลักษณะจากข้อมูลรูปภาพได้ recurrent neural networks หรือ RNN (Chaikaew, 2022; Vijitkunsawat et al., 2023; Sriboonruang et al., 2022) เป็นโมเดลที่สามารถเรียนรู้กับข้อมูลที่เป็นลำดับได้ และนอกจากนี้ยังมีโมเดลการเรียนรู้เชิงลึกอื่นอีกหลายโมเดล โดยการวิจัยนี้ผู้วิจัยเลือกใช้โมเดล CNN ร่วมกับ RNN เนื่องจากโมเดล CNN มีจุดเด่นในการดึงคุณลักษณะ (feature extraction) จากข้อมูลรูปภาพ สามารถนำมาใช้ดึงคุณลักษณะจากภาพในวิดีโอภาษามือได้ โดยโมเดลในกลุ่ม CNN นั้นมีอยู่หลายโมเดล การวิจัยนี้ผู้วิจัยเลือกใช้โมเดล CNN จำนวน 2 โมเดล ได้แก่ DenseNet121 และ EfficientNetB0 เนื่องจากเป็นโมเดลที่มีขนาดเล็ก เป็นโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) บนชุดข้อมูลขนาดใหญ่ ทำให้มีประสิทธิภาพในการทำนาย สำหรับโมเดล RNN มีจุดเด่นในการเรียนรู้ข้อมูลที่เป็นลำดับ (sequence learning) ซึ่งสามารถนำมาใช้กับข้อมูลวิดีโอภาษามือซึ่งเกิดจากรูปภาพที่นำมาต่อกันตามลำดับเวลา โดยโมเดลในกลุ่ม RNN นั้นมีอยู่หลายโมเดล การวิจัยนี้ผู้วิจัยเลือกใช้โมเดล RNN จำนวน 2 โมเดล ได้แก่ long-short term memory และ gated recurrent unit (GRU) เนื่องจากเป็นโมเดลที่นิยมใช้ในการเรียนรู้ข้อมูลที่เป็นลำดับเวลา ดังนั้นการใช้โมเดล CNN ร่วมกับ RNN จึงมีความเหมาะสมในการนำไปพัฒนาโมเดลการรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกจากข้อมูลที่เป็นวิดีโอ

จากที่กล่าวมาการพัฒนาโมเดลรู้จำภาษามือโดยใช้การเรียนรู้เชิงลึกจึงเป็นทางเลือกที่เหมาะสมสำหรับการแก้ปัญหาการสื่อสารระหว่างครูกับนักเรียนหูหนวก โดยโมเดลที่พัฒนาขึ้นนี้สามารถนำไปใช้เป็น

เครื่องมือในการแปลภาษามือไทยพื้นฐานที่ใช้ในห้องเรียน ช่วยให้ครูหรือนิสิตนักศึกษาครูที่ไม่มีความรู้พื้นฐานภาษามือ สามารถเข้าใจและสื่อสารกับนักเรียนหูหนวกได้อย่างมีประสิทธิภาพและมีความยืดหยุ่นมากขึ้น นอกจากนี้โมเดลที่พัฒนาขึ้นยังสามารถนำไปใช้เป็นโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) สำหรับงานวิจัยอื่น ๆ ซึ่งช่วยต่อยอดการพัฒนาโมเดลรู้จำภาษามือ และช่วยให้การฝึกสอนโมเดลในงานใหม่เป็นไปได้อย่างรวดเร็วและมีประสิทธิภาพมากขึ้นโดยไม่จำเป็นต้องเริ่มฝึกสอนโมเดลใหม่ตั้งแต่ต้น

ความมุ่งหมายของการวิจัย

1. เพื่อวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน
2. เพื่อพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

วิธีดำเนินการวิจัย

การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกในการวิจัยนี้ มีรายละเอียดวิธีดำเนินการวิจัย ได้แก่ ตัวอย่างวิจัย เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล โดยมีการวิจัย 2 ระยะ ได้แก่ การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน และการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

1. ตัวอย่างวิจัย

ตัวอย่างวิจัยที่ใช้การวิจัยนี้แบ่งออกเป็น 2 ระยะ ดังนี้

ระยะที่ 1 การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน ตัวอย่างวิจัย คือ ครูการศึกษาพิเศษ จำนวน 13 คน โดยการเลือกแบบเจาะจง มีเกณฑ์ในการคัดเลือก ได้แก่ 1) เป็นครูหรืออาจารย์ที่ทำการสอนประจำในโรงเรียนสอนคนหูหนวก 2) มีประสบการณ์อย่างน้อยหนึ่งปีในการสอนนักเรียนหูหนวกในระดับการศึกษาขั้นพื้นฐาน และ 3) มีความรู้ความเข้าใจในการใช้ภาษามือไทยเพื่อการสื่อสารกับนักเรียนหูหนวก

ระยะที่ 2 การพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก ตัวอย่างวิจัย คือ อาสาสมัครที่มีความสามารถในการใช้ภาษามือ จำนวน 4 คน โดยการเลือกแบบเจาะจง มีเกณฑ์ในการคัดเลือกได้แก่ 1) เป็นผู้ที่มีการใช้ภาษามือสื่อสารในโรงเรียนสอนคนหูหนวก 2) มีทักษะในการใช้ภาษามือไทยในการสื่อสารอยู่ในระดับดี และ 3) สามารถทำภาษามือตามประโยคที่ใช้ในการวิจัยได้ ทั้งนี้จำนวนตัวอย่างที่ใช้ในการวิจัยมีขนาดเพียงพอต่อการพัฒนาโมเดลรู้จำภาษามือโดยใช้การเรียนรู้เชิงลึก เนื่องจากผู้วิจัยจะใช้การเพิ่มจำนวนข้อมูล (data augmentation) และใช้โมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) มาต่อยอดในการพัฒนาโมเดลรู้จำภาษามือ ทำให้สามารถใช้ตัวอย่างที่มีขนาดเล็กได้

2. เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยนี้ คือ แบบสอบถามความถี่การใช้งานประโยคภาษามือ ซึ่งจะนำมาใช้ในส่วนของการวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน โดยแบบสอบถามนี้เป็นแบบสอบถามแบบมาตรประมาณค่า 4 ระดับ โดยที่ระดับ 4 หมายถึง ใช้ทุกคาบเรียน ระดับ 3 หมายถึง ใช้ในคาบเรียนส่วนใหญ่ ระดับ 2 หมายถึง ใช้ในบางคาบเรียน และระดับ 1 หมายถึง ไม่เคยใช้เลย

แบบสอบถามมีจำนวน 58 ข้อ ประกอบด้วยประโยคภาษามือที่แบ่งออกเป็น 5 หมวด ได้แก่ การนำเข้าสู่บทเรียน การตรวจสอบความเข้าใจ การออกคำสั่ง การเสริมแรง และการโต้ตอบของนักเรียน ทั้งนี้ ประโยคภาษามือที่ใช้ในแบบสอบถามนี้ได้มีการตรวจสอบความเหมาะสมและปรับแก้ตามคำแนะนำของผู้ทรงคุณวุฒิด้านการศึกษาพิเศษก่อนนำไปใช้เก็บรวบรวมข้อมูล

3. การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในการวิจัยนี้ผู้วิจัยแบ่งออกเป็น 2 ระยะ ได้แก่ การเก็บรวบรวมข้อมูลภาษามือที่ใช้ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน เพื่อนำมาวิเคราะห์หาประโยคภาษามือที่ใช้บ่อยในชั้นเรียนสำหรับนำไปใช้พัฒนาโมเดลรู้จำภาษามือ และการเก็บรวบรวมข้อมูลสำหรับการฝึกสอนและประเมินประสิทธิภาพของโมเดลรู้จำภาษามือ

ระยะที่ 1 การเก็บรวบรวมข้อมูลภาษามือที่ใช้ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน เป็นการเก็บรวบรวมข้อมูลเพื่อวิเคราะห์ประโยคภาษามือที่ครูมักใช้ในสื่อสารชั้นเรียนเพื่อนำประโยคดังกล่าวมาใช้ในการพัฒนาโมเดลรู้จำภาษามือ โดยผู้วิจัยได้ศึกษาประโยคที่ครูใช้ในชั้นเรียนจากนั้นนำไปปรึกษาผู้ทรงคุณวุฒิเพื่อตรวจสอบความเหมาะสมของประโยคที่ใช้ในชั้นเรียนของโรงเรียนสอนคนหูหนวกและปรับแก้ตามคำแนะนำของผู้ทรงคุณวุฒิ จากนั้นจึงจัดทำแบบสอบถามความถี่การใช้งานประโยคภาษามือและส่งแบบสอบถามไปยังครูการศึกษาพิเศษที่ทำการสอนในโรงเรียนสอนคนหูหนวก

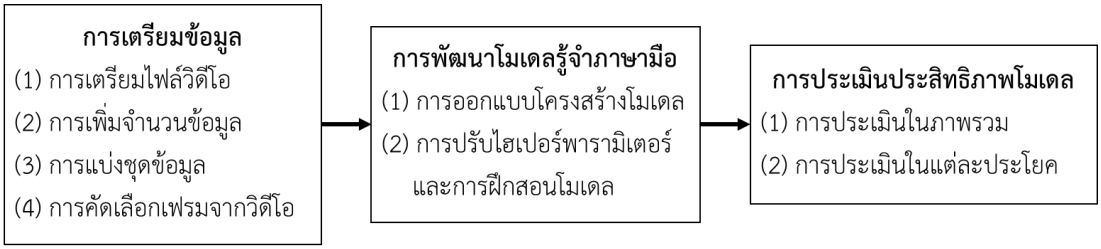
ระยะที่ 2 การเก็บรวบรวมข้อมูลสำหรับการฝึกสอนและประเมินประสิทธิภาพของโมเดลรู้จำภาษามือ เป็นการเก็บรวบรวมข้อมูลในรูปแบบวิดีโอการทำภาษามือไทย โดยนำประโยคภาษามือที่มีการใช้สื่อสารในชั้นเรียนจำนวน 20 ประโยค มาถ่ายทำกับอาสาสมัครจำนวน 4 คน และมีการถ่ายซ้ำจำนวน 4 ครั้ง ทำให้ได้วิดีโอสำหรับการพัฒนาโมเดลการรู้จำภาษามือไทยรวมทั้งสิ้น 320 วิดีโอ โดยหนึ่งวิดีโอจะแทนหนึ่งประโยค และแต่ละประโยคจะมีจำนวนวิดีโอเท่ากัน ทั้งนี้การเก็บรวบรวมข้อมูลในการวิจัยนี้ได้รับความอนุเคราะห์จากโรงเรียนเศรษฐเสถียร ในพระราชูปถัมภ์ โรงเรียนเฉพาะทางสำหรับนักเรียนหูหนวก ซึ่งมีการใช้ภาษามือในการจัดการเรียนการสอนเป็นหลัก และระหว่างการถ่ายวิดีโอมีผู้ทรงคุณวุฒิด้านภาษามือ 2 ท่าน คอยตรวจสอบความถูกต้องในการทำภาษามือของอาสาสมัคร

4. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยนี้แบ่งออกเป็น 2 ระยะ ได้แก่ การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียนเพื่อให้ได้ภาษามือที่จะใช้ในการพัฒนาโมเดล และการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกเพื่อหาโมเดลที่มีประสิทธิภาพสูงที่สุดและวิเคราะห์ความสามารถในการรู้จำภาษามือในภาพรวมและในแต่ละประโยค

ระยะที่ 1 การวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน เป็นการวิเคราะห์ข้อมูลจากแบบสอบถามการใช้งานประโยคภาษามือ เพื่อคัดเลือกประโยคภาษามือที่มีการใช้สื่อสารบ่อยในชั้นเรียนจำนวน 20 ประโยค โดยใช้ค่าเฉลี่ยของคะแนนความถี่ในการใช้งานประโยคภาษามือเพื่อให้ได้ประโยคภาษามือที่มีคะแนนความถี่ในการใช้งานมากที่สุด 20 อันดับแรก สำหรับนำมาใช้ในการพัฒนาโมเดลรู้จำภาษามือ และใช้ค่าร้อยละเพื่อเปรียบเทียบจำนวนประโยคภาษามือที่ใช้ในแต่ละหมวด

ระยะที่ 2 การพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลโดยใช้ Google Colab ในการประมวลผลบน TPU ด้วยภาษา Python ผ่าน Jupyter notebook ในการเขียนคำสั่ง และสร้างโมเดลการเรียนรู้เชิงลึกผ่าน Tensorflow และ Keras API โดยการพัฒนาโมเดลประกอบด้วย 3 ขั้นตอน ดังภาพ 1 ได้แก่ การเตรียมข้อมูลเพื่อให้ได้ข้อมูลในรูปแบบที่สามารถนำไปใช้ฝึกสอนโมเดลการเรียนรู้เชิงลึกได้ การพัฒนาโมเดลเพื่อออกแบบโครงสร้างโมเดลและฝึกสอนโมเดล และการประเมินประสิทธิภาพของโมเดลเพื่อวิเคราะห์ความสามารถในการรู้จำภาษามือของโมเดลในภาพรวมและในแต่ละประโยค



ภาพ 1 ขั้นตอนการวิเคราะห์ข้อมูลในส่วนของการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทย

1. การเตรียมข้อมูล สำหรับการพัฒนาโมเดลรู้จำภาษามือในการวิจัยนี้แบ่งเป็น 4 ขั้นตอน ประกอบด้วย การเตรียมไฟล์วิดีโอเพื่อให้ได้วิดีโอเฉพาะส่วนที่มีสารสนเทศสำหรับการพัฒนาโมเดล การเพิ่มจำนวนข้อมูลเพื่อให้โมเดลสามารถเรียนรู้ได้มีประสิทธิภาพมากขึ้น การแบ่งชุดข้อมูลสำหรับฝึกสอน ตรวจสอบ และทดสอบประสิทธิภาพของโมเดล และการคัดเลือกเฟรมจากวิดีโอเพื่อให้ได้เฟรมรูปภาพในจำนวนที่เหมาะสมกับการพัฒนาโมเดล

(1) การเตรียมไฟล์วิดีโอ เนื่องจากแต่ละวิดีโอที่ผู้วิจัยบันทึกไว้มีส่วนที่ไม่ได้มีการสารสนเทศเกี่ยวกับภาษามือ ได้แก่ ช่วงที่อาสาสมัครอยู่นิ่ง และภาพพื้นหลัง ดังนั้นผู้วิจัยจึงได้ใช้โปรแกรม CapCut ตัดวิดีโอให้เหลือเฉพาะช่วงที่มีการเคลื่อนไหวของมือ และลบพื้นหลังของวิดีโอออกเป็นสีดำให้เหลือเฉพาะผู้ทำภาษามือ เพื่อให้ได้เฉพาะวิดีโอในส่วนที่มีสารสนเทศที่จำเป็นในการนำไปสร้างโมเดลรู้จำภาษามือ ทั้งนี้ ชื่อไฟล์วิดีโอจะถูกตั้งให้ตามความหมายของประโยคภาษามือในแต่ละวิดีโอ

(2) การเพิ่มจำนวนข้อมูล (data augmentation) เนื่องจากวิดีโอที่ผู้วิจัยรวบรวมมาได้ในแต่ละประโยคมีจำนวนที่น้อย แต่ในการพัฒนาโมเดลการเรียนรู้เชิงลึกหากมีข้อมูลให้โมเดลได้เรียนรู้จำนวนมากจะช่วยให้โมเดลได้เรียนรู้จากข้อมูลที่มีความหลากหลายและทำให้โมเดลมีประสิทธิภาพมากขึ้น ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงได้ทำการเพิ่มจำนวนข้อมูลของวิดีโอภาษามือ โดยนำวิดีโอที่มีอยู่เดิมมาทำการดัดแปลงให้มีความหลากหลายมากขึ้นผ่านไลบรารี OpenCV ด้วยวิธีการเพิ่มจำนวนข้อมูล 7 วิธี ได้แก่ การกลับด้าน (flip) การหมุน (rotate) การเพิ่มจุดรบกวน (noise) การเบลอ (blur) การปรับเป็นสีขาวดำ (greyscale) การเพิ่มความเร็ว (speed up) และการลดความเร็ว (slow down) ทำให้ได้วิดีโอใหม่อีก 2,240 วิดีโอ ซึ่งรวมกับวิดีโอเดิมที่มีอยู่ 320 วิดีโอ ทำให้ได้วิดีโอสำหรับนำไปพัฒนาโมเดลจำนวนทั้งสิ้น 2,560 วิดีโอ

(3) การแบ่งชุดข้อมูล การพัฒนาโมเดลการเรียนรู้เชิงลึกในการวิจัยครั้งนี้มีได้นำข้อมูลที่เก็บรวมรวมได้และนำไปเพิ่มจำนวนวิดีโอจนได้เป็น 960 วิดีโอ มาแบ่งออกเป็น 3 ชุดข้อมูล เพื่อใช้สำหรับฝึกสอน

ตรวจสอบ และทดสอบประสิทธิภาพของโมเดล ดังนี้ 1) ชุดข้อมูลฝึกสอน (training dataset) ใช้ข้อมูลร้อยละ 75 ของข้อมูลทั้งหมด เป็นชุดข้อมูลที่นำมาใช้ฝึกสอนโมเดลการเรียนรู้เชิงลึก โดยให้โมเดลแต่ละโมเดลเรียนรู้จากข้อมูลชุดนี้ 2) ชุดข้อมูลตรวจสอบ (validation dataset) ใช้ข้อมูลร้อยละ 12.5 ของข้อมูลทั้งหมด เป็นชุดข้อมูลที่นำมาใช้ตรวจประสิทธิภาพโมเดลการเรียนรู้เชิงลึกขณะที่กำลังฝึกสอนโมเดล และใช้เปรียบเทียบประสิทธิภาพของแต่ละโมเดลที่ใช้ในการวิจัยครั้งนี้ และ 3) ชุดข้อมูลทดสอบ (test dataset) ใช้ข้อมูลร้อยละ 12.5 ของข้อมูลทั้งหมด เป็นชุดข้อมูลที่นำมาใช้ทดสอบประสิทธิภาพของโมเดลการเรียนรู้เชิงลึกที่มีความแม่นยำมากที่สุด และใช้ประเมินประสิทธิภาพการจำแนกของแต่ละประโยคที่ใช้ในโมเดลการเรียนรู้เชิงลึก ทั้งนี้ชุดข้อมูลฝึกสอนจะใช้ข้อมูลวิดีโอจากอาสาสมัคร 3 คน เพื่อให้มีข้อมูลจำนวนมากพอสำหรับการฝึกสอนโมเดล ส่วนชุดข้อมูลตรวจสอบและชุดข้อมูลทดสอบจะใช้ข้อมูลวิดีโอจากอาสาสมัคร 1 คน แล้วแบ่งข้อมูลวิดีโอออกเป็นสองส่วนสำหรับใช้ในสองชุดข้อมูลดังกล่าว

(4) การคัดเลือกเฟรมจากวิดีโอ วิดีโอภาษามือไทยที่ได้ทำการแบ่งชุดข้อมูลเรียบร้อยแล้วจะถูกนำมาคัดเลือกเฟรมและปรับขนาดให้มีความเหมาะสมเพื่อนำไปใช้สำหรับพัฒนาโมเดลการเรียนรู้เชิงลึก เนื่องจากหากใช้ทุกเฟรมจากวิดีโอหรือภาพในแต่ละเฟรมเป็นภาพขนาดใหญ่ จะทำให้โมเดลมีขนาดใหญ่และใช้ระยะเวลานานในการพัฒนาโมเดล ดังนั้นผู้วิจัยจึงได้กำหนดจำนวนเฟรมที่คัดเลือกจากแต่ละวิดีโอไว้ที่จำนวน 20 เฟรม และกำหนดขนาดภาพในแต่ละเฟรมให้อยู่ที่ 224×224 พิกเซล เนื่องจากเป็นขนาดที่ถูกกำหนดให้ใช้ในการพัฒนาโมเดลการเรียนรู้เชิงลึกด้วยโมเดลที่ผู้วิจัยเลือกใช้ในงานวิจัยนี้

2. การพัฒนาโมเดล แบ่งเป็น 2 ส่วน คือ การออกแบบโครงสร้างของโมเดลสำหรับการรู้จำภาษามือไทย และการปรับไฮเปอร์พารามิเตอร์และการฝึกสอนโมเดลเพื่อหาโมเดลที่มีประสิทธิภาพดีที่สุด

(1) การออกแบบโครงสร้างของโมเดล

การวิจัยนี้เป็นการพัฒนาโมเดลรู้จำภาษามือจากวิดีโอภาษามือแต่ละประโยค โดยหนึ่งวิดีโอแทนหนึ่งประโยคภาษามือ โมเดลที่ใช้ในการรู้จำภาษามือจึงเป็นโมเดลจำแนกแบบหลายกลุ่ม (multi-class classification model) และเนื่องจากวิดีโอเป็นข้อมูลที่เกิดจากรูปภาพที่นำต่อกันตามลำดับเวลา ซึ่งโมเดลการเรียนรู้เชิงลึกที่มีประสิทธิภาพในการทำงานกับข้อมูลรูปภาพคือโมเดลในกลุ่ม convolutional neural network (CNN) และโมเดลการเรียนรู้เชิงลึกโมเดลการเรียนรู้เชิงลึกที่มีประสิทธิภาพในการทำงานกับข้อมูลที่เป็นลำดับคือโมเดลในกลุ่ม recurrent neural network (RNN) ดังนั้นในการพัฒนาโมเดลการรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก ผู้วิจัยได้พัฒนาโมเดลการรู้จำภาษามือไทยโดยประยุกต์ใช้ CNN ร่วมกับ RNN

โมเดล CNN และ RNN ที่ใช้ในการวิจัยนี้ โมเดลกลุ่ม CNN ผู้วิจัยเลือกใช้โมเดล DenseNet121 และ EfficientNetB0 เนื่องจากเป็นโมเดลขนาดเล็ก และเป็นโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) กับชุดข้อมูลรูปภาพขนาดใหญ่ ImageNet สามารถนำมาต่อยอดในการดึงคุณลักษณะจากเฟรมภาพในวิดีโอภาษามือได้ ในส่วนของโมเดลกลุ่ม RNN ผู้วิจัยเลือกใช้โมเดล long-short term memory (LSTM) และโมเดล gated recurrent unit (GRU) เนื่องจากเป็นโมเดลที่นิยมใช้ในการจำแนกข้อมูลที่มีลักษณะเป็นลำดับ ดังนั้นจากโมเดลที่กล่าวมาข้างต้น การพัฒนาโมเดลการรู้จำภาษามือไทยในการวิจัยนี้มีโมเดลที่ถูกใช้ทั้งหมด 4 โมเดล ได้แก่ 1) DenseNet121 ร่วมกับ LSTM 2) DenseNet121 ร่วมกับ GRU 3) EfficientNetB0 ร่วมกับ LSTM และ 4) EfficientNetB0 ร่วมกับ GRU

(2) การปรับไฮเปอร์พารามิเตอร์ของโมเดลและการฝึกสอนโมเดล

ในส่วนของการปรับไฮเปอร์พารามิเตอร์ (hyperparameter) ซึ่งเป็นพารามิเตอร์ที่กำหนดไว้ก่อนการฝึกสอน ผู้วิจัยเลือกใช้วิธีการปรับการพารามิเตอร์ (optimizer) แบบ adam และใช้ฟังก์ชันค่าสูญเสีย (loss function) สำหรับวัดความแตกต่างของค่าจริงกับค่าที่โมเดลทำนายเป็น sparse categorical cross entropy เนื่องจากผู้วิจัยได้กำหนดผลลัพธ์การทำนายภาษามือ 20 ประโยคหรือคลาสสำหรับการทำนายในรูปแบบตัวเลข 0 ถึง 19 ในการฝึกสอนโมเดล จากนั้นกำหนดดัชนีที่ใช้ประเมินประสิทธิภาพของโมเดล (metrics) ระหว่างการฝึกสอนโมเดลแต่ละรอบเป็นค่าความถูกต้อง (accuracy) เนื่องจากโมเดลในการวิจัยนี้เป็นโมเดลจำแนกแบบหลายกลุ่ม ต่อมากำหนดจำนวนตัวอย่างข้อมูลในการปรับค่าพารามิเตอร์หนึ่งครั้ง (batch size) เท่ากับ 32 จำนวนรอบในการพัฒนาโมเดล (epoch) เท่ากับ 20 รอบ และกำหนดให้มีการหยุดฝึกสอนโมเดลก่อนครบจำนวนรอบ (early stopping) เมื่อค่าความสูญเสียบนข้อมูลชุดตรวจสอบ (validation loss) ไม่ลดลงติดต่อกัน 10 รอบ เพื่อป้องกันการเรียนรู้บนชุดข้อมูลฝึกสอนมากเกินไป (overfitting)

ผู้วิจัยดำเนินการฝึกสอนโมเดลโดยนำเข้าข้อมูลเฟรมภาพที่ได้จากการเตรียมข้อมูล ซึ่งเป็นข้อมูลที่ครอบคลุมในทุกท่ามือและการเคลื่อนไหวของภาษามือในแต่ละประโยค ในระหว่างการการฝึกสอนโมเดลผู้วิจัยจะดำเนินการติดตามค่าความสูญเสีย (loss) ซึ่งเป็นค่าความแตกต่างระหว่างค่าจริงกับค่าที่โมเดลทำนาย และค่าความถูกต้อง (accuracy) ซึ่งเป็นสัดส่วนข้อมูลตัวอย่างที่โมเดลทำนายได้ถูกต้อง โดยจะติดตามค่าดังกล่าวในชุดข้อมูลฝึกสอนและชุดข้อมูลตรวจสอบ เพื่อตรวจสอบความถูกต้องและความผิดพลาดของแต่ละโมเดล จากนั้นเมื่อฝึกสอนเสร็จสิ้นครบทุกโมเดลแล้ว ผู้วิจัยจะนำแต่ละโมเดลมาเปรียบเทียบประสิทธิภาพโดยเปรียบเทียบจากค่าความสูญเสียบนข้อมูลชุดตรวจสอบ (validation loss) ค่าความถูกต้องบนชุดข้อมูลตรวจสอบ (validation accuracy) ขนาดโมเดล และเวลาเฉลี่ยที่ใช้ในการฝึกสอนโมเดลต่อรอบ ทั้งนี้ โมเดลที่มีประสิทธิภาพดีที่สุดจะถูกนำไปใช้ประเมินประสิทธิภาพโมเดลกับข้อมูลชุดทดสอบ

(3) การประเมินประสิทธิภาพของโมเดลรู้จำภาษามือไทย

การประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทย ผู้วิจัยนำโมเดลที่มีประสิทธิภาพสูงที่สุดจากขั้นตอนการพัฒนาโมเดลมาประเมินประสิทธิภาพกับชุดข้อมูลทดสอบ ทั้งนี้ เนื่องจากโมเดลการแปลภาษามือไทยโดยใช้การเรียนรู้เชิงลึกเป็นโมเดลจำแนกแบบหลายกลุ่ม (multi-class classification model) ดังนั้นการประเมินประสิทธิภาพของโมเดลจึงใช้คอนฟิวชันเมทริกซ์ (confusion matrix) แบบหลายกลุ่ม (multi-class) ซึ่งเป็นเมทริกซ์ที่แสดงจำนวนตัวอย่างข้อมูลในประโยคที่โมเดลทำนายเทียบกับประโยคจริงว่าโมเดลทำนายได้ถูกต้องมากน้อยเพียงใด มีการทำนายผิดพลาดอย่างไรบ้าง และนำจำนวนตัวอย่างข้อมูลที่ได้จากคอนฟิวชันเมทริกซ์มาประเมินประสิทธิภาพของโมเดลรู้จำภาษามือในภาพรวมและประสิทธิภาพการรู้จำในแต่ละประโยคภาษามือโดยใช้ค่าต่อไปนี้

1) ค่าความถูกต้อง (accuracy) เป็นค่าที่แสดงถึงความถูกต้องในการทำนายของโมเดลในภาพรวมทุกประโยค คำนวณจากสมการ $\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$

2) ค่าความแม่นยำ (precision) เป็นค่าที่แสดงถึงความแม่นยำในการทำนายในแต่ละประโยค คำนวณจากสมการ $\text{precision} = \frac{TP}{TP + FP}$

3) ค่าการเรียนคืน (recall) เป็นค่าที่แสดงถึงการไม่พลาดการทำนายในแต่ละประโยค คำนวณจากสมการ $\text{recall} = \frac{TP}{TP + FN}$

4) คะแนน F1 (F1-score) เป็นค่าเฉลี่ยแบบฮาร์มอนิกของค่า precision และ recall คำนวณจากสมการ $\text{F1-score} = 2 \times \frac{\text{Precison} \times \text{Recall}}{\text{Precison} + \text{Recall}}$

โดยถ้ากำหนดให้ positive เป็นคลาสหรือประโยคที่สนใจ และคลาสอื่นหรือประโยคอื่นเป็น negative จะได้ว่า true positive (TP) คือ จำนวนตัวอย่างที่โมเดลทำนายว่าเป็นคลาส positive และอยู่ในคลาส positive จริง true negative (TN) คือ จำนวนตัวอย่างที่โมเดลทำนายว่าเป็นคลาส negative และอยู่ในคลาส negative จริง false positive (FP) คือ จำนวนตัวอย่างที่โมเดลทำนายว่าเป็นคลาส positive แต่ความจริงไม่ได้อยู่ในคลาส positive จริง และ false negative (FN) คือ จำนวนตัวอย่างที่โมเดลทำนายว่าเป็นคลาส negative แต่ความจริงไม่ได้อยู่ในคลาส negative

ผลการวิจัย

ผลการวิจัยแบ่งออกได้เป็น 2 ตอน ได้แก่ ผลการวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน และผลการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

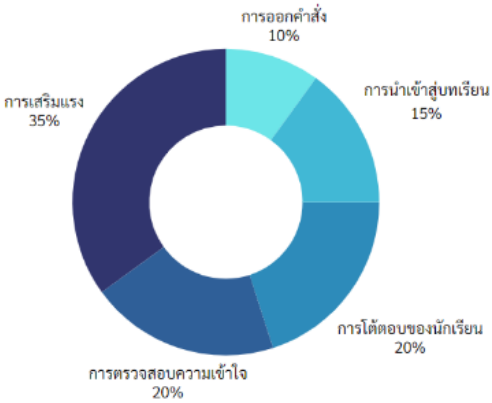
ตอนที่ 1 ผลการวิเคราะห์ภาษามือไทยที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษในชั้นเรียน

จากการวิเคราะห์ค่าเฉลี่ยของคะแนนความถี่ประโยคภาษามือที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษจากแบบสอบถามความถี่การใช้งานประโยคภาษามือในชั้นเรียน โดยคะแนนความถี่จะอยู่ในช่วง 1 ถึง 4 คะแนน คะแนนที่มีค่ามากจะหมายถึงการใช้ประโยคภาษามือในชั้นเรียนบ่อยครั้ง ผลการวิเคราะห์พบว่าประโยคภาษามือที่มีการใช้งานบ่อยที่สุด 20 อันดับแรก ได้แก่ สวัสดี พร้อมยัง ถูกหรือผิด เก่งมาก เรียบร้อย ตั้งใจเรียน เข้าใจ ถูกไหม ดูภาพ ทำได้ไหม ใครตอบได้ งานสวย ภูมิใจ รู้ไหม ทำใบงาน เก็บอุปกรณ์ เยี่ยม น่ารัก ผิดไหม และทำอย่างไร โดยแต่ละประโยคมีค่าเฉลี่ยคะแนนความถี่จำแนกตามหมวด ดังตาราง 1

ตาราง 1 ค่าเฉลี่ยคะแนนความถี่ของประโยคภาษามือ 20 ประโยคที่ใช้งานบ่อยที่สุดในแต่ละหมวด

หมวด	ประโยค	ค่าเฉลี่ยคะแนนความถี่	หมวด	ประโยค	ค่าเฉลี่ยคะแนนความถี่
1. การนำเข้าสู่บทเรียน	สวัสดี	4.00	4. การเสริมแรง	เก่งมาก	3.92
	พร้อมยัง	3.92		เรียบร้อย	3.92
	ดูภาพ	3.85		ตั้งใจเรียน	3.92
2. การตรวจสอบความเข้าใจ	ถูกหรือผิด	3.92		งานสวย	3.85
	ทำได้ไหม	3.85		ภูมิใจ	3.85
	ใครตอบได้	3.85		เยี่ยม	3.77
	รู้ไหม	3.77		น่ารัก	3.77
3. การออกคำสั่ง	ทำใบงาน	3.77	5. การโต้ตอบของนักเรียน	เข้าใจ	3.92
	เก็บอุปกรณ์	3.77		ถูกไหม	3.92
				ผิดไหม	3.77
				ทำอย่างไร	3.69

จากตาราง 1 พบว่าประโยคภาษามือทั้ง 20 ประโยคมีการใช้งานบ่อยในชั้นเรียนเนื่องจากมีค่าคะแนนความถี่ใกล้เคียงกับ 4 ซึ่งเป็นค่าสูงสุด และเมื่อเปรียบเทียบจำนวนประโยคภาษามือในที่ถูกคัดเลือกแต่ละหมวดมาพบว่าหมวดการเสริมแรงมีจำนวนประโยคมากที่สุด (ร้อยละ 35) รองลงมาคือหมวดการตรวจสอบความเข้าใจ (ร้อยละ 20) หมวดการโต้ตอบของนักเรียน (ร้อยละ 20) หมวดการนำเข้าสู่บทเรียน (ร้อยละ 15) และหมวดการออกคำสั่งมีจำนวนประโยคน้อยที่สุด (ร้อยละ 10) ดังภาพ 2 โดยประโยคภาษามือที่ได้มาในขั้นตอนนี้จะถูกนำไปใช้ในการพัฒนาโมเดลการรู้จำภาษามือโดยใช้การเรียนรู้เชิงลึกในขั้นตอนต่อไป



ภาพ 2 แผนภูมิแสดงจำนวนประโยคภาษามือที่ใช้บ่อยจำแนกตามหมวด

ตอนที่ 2 ผลการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

ผลการพัฒนาและประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก แบ่งออกเป็น 3 ส่วน ได้แก่ ข้อมูลที่ใช้ในการพัฒนาโมเดล ผลการพัฒนาโมเดลรู้จำภาษามือไทย และผลการประเมินประสิทธิภาพของโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

1. ข้อมูลที่ใช้ในการพัฒนาโมเดล

ข้อมูลที่ใช้ในการพัฒนาโมเดลรู้จำภาษามือไทยเป็นวิดีโอท่าภาษามือที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษจำนวน 20 ประโยค โดยมีลักษณะท่ามือและการเคลื่อนไหว รวมไปถึงความยาวของวิดีโอที่ใช้ในการพัฒนาโมเดลแสดงได้ดังตาราง 2

ตาราง 2 ลักษณะท่ามือและความยาววิดีโอในแต่ละประโยคภาษามือที่ใช้ในการวิจัย

ประโยค	ท่ามือและการเคลื่อนไหว	ความยาววิดีโอ (วินาที)
พร้อมยัง	(1) คว่ำฝ่ามือลงทั้งสองข้างไว้บริเวณกลางลำตัวและมือห่างกัน เคลื่อนฝ่ามือซ้ายไปขวา (2) มือข้างหนึ่งชูนิ้วโป้งและนิ้วก้อย จากนั้นขยับมือกลับไปกลับมาช้า ๆ	2.4
งานสวย	(1) กำมือสองข้างข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง เคลื่อนที่เข้ามาหากันบริเวณหน้าอก (2) แบมือแล้วเคลื่อนที่วนรอบบริเวณใบหน้า (3) ชูนิ้วโป้งไว้ที่บริเวณด้านข้างศีรษะ	1.9
ทำได้ไหม	(1) กำมือสองข้างข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง เคลื่อนที่เข้ามาหากันบริเวณกลางลำตัว และทำซ้ำอีกครั้ง (2) ชูนิ้วโป้งบริเวณด้านหน้า (3) ใช้นิ้วชี้เคลื่อนที่เป็นรูปเครื่องหมายคำถาม	1.7
รู้ไหม	(1) ชี้นิ้วชี้และนิ้วกลางไปยังบริเวณข้างศีรษะ (2) ใช้นิ้วชี้เคลื่อนที่เป็นรูปเครื่องหมายคำถาม	1.3
ทำใบงาน	(1) กำมือสองข้างข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง เคลื่อนที่เข้ามาหากันบริเวณกลางลำตัว และทำซ้ำอีกครั้ง (2) ตั้งฝ่ามือสองข้างไว้ที่ข้างศีรษะ หันออกมาด้านหน้า ขยับมือซ้ายขวาเล็กน้อย	2.1

ตาราง 2 (ต่อ)

ประโยค	ท่ามือและการเคลื่อนไหว	ความยาววิดีโอ (วินาที)
	พร้อมเคลื่อนมือลงมาด้านล่าง (3) กำมือหนึ่งข้างไว้ด้านโดยไขว้กับมืออีกข้าง จากนั้นขยับมือด้านบนไปข้างหน้าสองจังหวะ	
เยี่ยม	(1) ชี้นิ้วชี้ของมือข้างหนึ่งไปด้านหน้า (2) ชูนิ้วโป้งโดยใช้มือข้างเดียวกัน	0.9
สวัสดี	(1) แบมือและหันฝ่ามือไปแตะที่ศีรษะแล้วเคลื่อนฝ่ามือออกมาด้านหน้า	1.0
ทำอย่างไร	(1) กำมือสองข้างข้างหนึ่งอยู่ด้านบน อีกข้างอยู่ด้านล่าง เคลื่อนที่เข้ามาหากันบริเวณกลางลำตัว (2) แบมือไว้บริเวณกลางลำตัว แล้วม้วนฝ่ามือออกไปด้านนอก	1.6
ถูกไหม	(1) มือข้างหนึ่งชูนิ้วก้อยแล้วเคลื่อนที่ขึ้นบริเวณมืออีกข้างหนึ่งที่กำลังแบอยู่ (2) ใช้นิ้วชี้เคลื่อนที่เป็นรูปเครื่องหมายคำถาม	1.7
ผิดไหม	(1) มือข้างหนึ่งชูนิ้วก้อยแล้วเคลื่อนที่ปัดลงบริเวณมืออีกข้างหนึ่งที่กำลังแบอยู่ (2) ใช้นิ้วชี้เคลื่อนที่เป็นรูปเครื่องหมายคำถาม	1.5
ดูภาพ	(1) ตั้งฝ่ามือมือสองข้าง ข้างหนึ่งอยู่บริเวณศีรษะ ข้างหนึ่งอยู่บริเวณไหล่ (2) เคลื่อนมือด้านบนมาหามือที่อยู่ด้านล่าง โดยนิ้วหัวแม่มือและนิ้วทั้งสี่ขยับเข้าหากัน (3) มือข้างหนึ่งใช้นิ้วชี้และนิ้วกลางชี้ไปยังมือที่แบอยู่ แล้วเคลื่อนมือลง	2.3
เรียบร้อย	(1) มือข้างหนึ่งหงายวางไว้ด้านล่าง มืออีกข้างหนึ่งคว่ำไว้ด้านบนบริเวณกลางลำตัว แล้วเคลื่อนมือออกจากกันไปด้านข้าง	1.6
น่ารัก	(1) ขยับนิ้วโป้งและนิ้วทั้งสี่เข้าและออกจากกันบริเวณข้างใบหน้า	1.1
ภูมิใจ	(1) ตั้งฝ่ามือข้างหนึ่งไว้กลางลำตัว แล้วเคลื่อนฝ่ามือขึ้นไปบริเวณหน้าอก	1.8
ถูกหรือผิด	(1) มือข้างหนึ่งชูนิ้วก้อยแล้วเคลื่อนที่ขึ้นบริเวณมืออีกข้างหนึ่งที่กำลังแบอยู่ (2) มือข้างที่ชูนิ้วก้อยเคลื่อนที่ลงบริเวณมืออีกข้างหนึ่งที่กำลังแบอยู่ (3) ใช้นิ้วชี้เคลื่อนที่ซ้ายขวากลับไปกลับมา	2.6
เก็บอุปกรณ์	(1) กางนิ้วหัวแม่มือ นิ้วชี้ และนิ้วกลางของทั้งสองมือให้อยู่ในแนวตั้งฉากกัน มือข้างหนึ่งอยู่ด้านบนอีกข้างอยู่ด้านล่าง แล้วเคลื่อนที่เข้าหากัน (2) หงายฝ่ามือโดยงอนิ้วเข้าหากัน ขยับมือขึ้นสลับลงเล็กน้อยพร้อมกับเคลื่อนมือไปด้านข้าง	2.0
ตั้งใจเรียน	(1) ตั้งฝ่ามือสองข้างไว้ที่บริเวณข้างศีรษะแล้วเคลื่อนที่ลงมา (2) แบมือข้างหนึ่งไว้บริเวณหน้าอก มืออีกข้างนิ้วทั้งห้าแตะกันและเคลื่อนที่ขึ้นไปบริเวณศีรษะ	1.7
เข้าใจ	(1) กำมือไว้บริเวณข้างศีรษะ (2) ชูนิ้วชี้บริเวณข้างศีรษะจากนั้นขยับนิ้วขึ้นลง	1.4
เก่งมาก	(1) แบมือข้างหนึ่งไว้ด้านล่าง มืออีกข้างอยู่ด้านบนชูนิ้วชี้กับนิ้วกลางแล้วเคลื่อนที่ลงมาหามือด้านล่าง (2) ตั้งฝ่ามือสองข้างแล้วเคลื่อนฝ่ามือออกจากกัน	1.7
ใครตอบได้	(1) นิ้วชี้ข้างหนึ่งเคลื่อนที่วนรอบบริเวณใบหน้า (2) ชี้นิ้วทั้งสองมือไปบริเวณปากและปัดออกมาด้านหน้า (3) กำมือข้างหนึ่งและชูนิ้วโป้ง	2.2

จากตารางพบว่าภาษามือในแต่ละประโยคจะมีท่ามือและการเคลื่อนไหวแตกต่างกันทำให้ความยาววิดีโอที่ได้แตกต่างกันออกไปตามความซับซ้อนของท่าภาษามือ เมื่อพิจารณาท่ามือและการเคลื่อนไหวของแต่ละประโยคภาษามือ พบว่า ประโยคภาษามือที่ใช้ในการวิจัยนี้มีจำนวนท่ามือ 1-3 ท่า และยังพบว่าบางประโยคมีท่ามือและการเคลื่อนไหวแตกต่างกันอย่างชัดเจน ในขณะที่บางประโยคมีท่ามือที่มีลักษณะคล้ายกันซึ่งอาจส่งผลต่อประสิทธิภาพของโมเดลรู้จำภาษามือไทย

2. ผลการพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

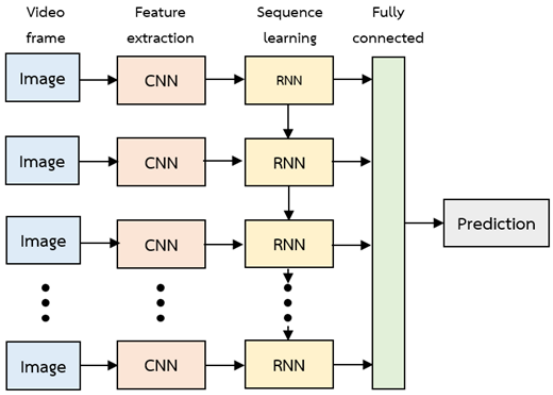
การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกในการวิจัยนี้เป็นการการประยุกต์ใช้โมเดล CNN ร่วมกับ RNN โดยการออกแบบโครงสร้างของโมเดลแสดงได้ดังภาพ 3 ซึ่งผู้วิจัยได้กำหนดชั้น (layer) ของโมเดลไว้ 4 ส่วน แต่ละส่วนมีรายละเอียด ดังนี้

(1) rescaling layer เป็นชั้นที่นำข้อมูลนำเข้า (input) จากภาพในแต่ละเฟรมของวิดีโอซึ่งมีค่าตั้งแต่ 0 ถึง 255 มาจัดกระทำให้มีค่าอยู่ระหว่าง 0 ถึง 1 เพื่อให้โมเดลสามารถเรียนรู้ได้มีประสิทธิภาพมากยิ่งขึ้น

(2) layer ของโมเดลในกลุ่ม CNN สำหรับดึงคุณลักษณะ (feature extraction) ได้แก่ time distributed layer ของโมเดล EfficientNetB0 และ DenseNet121 ซึ่งโดยทั่วไปโมเดลทั้งสองนี้เป็นโมเดลสำหรับข้อมูลที่เป็นรูปภาพ หากนำมาครอบด้วย time distributed layer จะทำให้โมเดลสามารถนำไปใช้กับข้อมูลที่เป็นวิดีโอได้ โดย time distributed layer จะแบ่งเฟรมภาพของวิดีโอแต่ละเฟรมเข้าไปในโมเดล EfficientNetB0 และ DenseNet121 และผู้วิจัยนำชั้นดังกล่าวมาเชื่อมต่อกับ time distributed global average pooling 2D เพื่อลดขนาดของข้อมูลสำหรับนำไปใช้ในส่วนต่อไป ทั้งนี้ในส่วนฐานของโมเดล CNN ทั้งสองนี้จะใช้ค่าพารามิเตอร์จากโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) บนชุดข้อมูลขนาดใหญ่ ImageNet โดยจะตัดในส่วนการจำแนก (classifier) ของโมเดลเดิมออก แล้วแทนที่ด้วยชั้นของโมเดลส่วนต่อไปที่ใช้ในการวิจัยนี้เพื่อให้สามารถจำแนกประโยคภาษามือได้อย่างมีประสิทธิภาพ

(3) layer ของโมเดลในกลุ่ม RNN สำหรับเรียนรู้ข้อมูลที่เป็นลำดับ (sequence learning) ได้แก่ โมเดล LSTM และ GRU โดยผู้วิจัยกำหนดจำนวนหน่วยประมวลผล (neuron) ไว้เท่ากับ 64

(4) dense layer หรือ fully connected layer เป็นชั้นสุดท้ายสำหรับการจำแนกวิดีโอประโยคภาษามือ 20 ประโยค หรือ 20 คลาส ผู้วิจัยจึงกำหนดจำนวนหน่วยประมวลผล (neuron) ไว้เท่ากับ 20



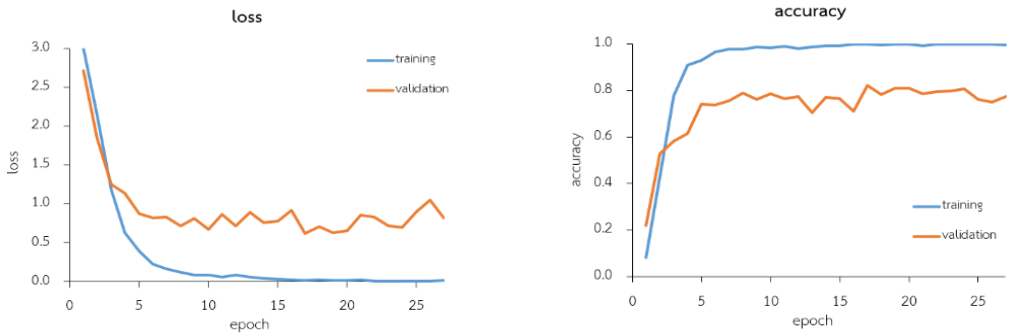
ภาพ 3 โครงสร้างของโมเดลการเรียนรู้เชิงลึกที่ใช้ในการวิจัย

โมเดลการเรียนรู้เชิงลึกสำหรับรู้จำภาษามือไทยที่มีโครงสร้างดังที่กล่าวมาข้างต้นเมื่อนำมาฝึกสอนและนำประสิทธิภาพของโมเดลมาเปรียบเทียบกันดังตาราง 3 พบว่า การใช้โมเดล EfficientNetB0 ร่วมกับ GRU ทำให้โมเดลการรู้จำภาษามือมีประสิทธิภาพมากที่สุดทั้งในด้านความถูกต้อง ขนาดของโมเดล และเวลาที่ใช้ในการฝึกสอนโมเดล รองลงมาคือ โมเดล EfficientNetB0 ร่วมกับ LSTM โมเดล DenseNet121 ร่วมกับ GRU และโมเดล DenseNet121 ร่วมกับ LSTM ตามลำดับ ซึ่งแสดงให้เห็นว่าโมเดลที่มีประสิทธิภาพในการทำนาย นั้นไม่จำเป็นต้องมีขนาดใหญ่และใช้เวลาฝึกสอนมาก นอกจากนี้เมื่อพิจารณาการฝึกสอนโมเดลจากกราฟในภาพ 4 ซึ่งแสดงค่าความสูญเสีย (loss) และค่าความถูกต้อง (accuracy) ในการฝึกสอนโมเดล EfficientNetB0 ร่วมกับ GRU ซึ่งพบว่าค่าความสูญเสียบนชุดข้อมูลฝึกสอน (training loss) และค่าความ

ถูกต้องบนชุดข้อมูลฝึกสอน (training accuracy) มีแนวโน้มลู่เข้าสู่ค่าคงที่ แสดงให้เห็นว่าโมเดลสามารถเรียนรู้จากชุดข้อมูลฝึกสอนได้ ในขณะที่ค่าความสูญเสียบนชุดข้อมูลตรวจสอบ (validation loss) และค่าความถูกต้องบนชุดข้อมูลตรวจสอบ (validation accuracy) มีการเพิ่มขึ้นสลับลดลงและมีแนวโน้มคงที่ แสดงให้เห็นว่าโมเดลยังไม่เกิดการเรียนรู้มากเกินไป (overfitting)

ตาราง 3 การเปรียบเทียบโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

model	validation loss	validation accuracy	size (MB)	average training time per epoch (s)
EfficientNetB0+GRU	0.8238	0.7750	16.44	541
EfficientNetB0+LSTM	1.2220	0.6969	16.77	549
DenseNet121+GRU	1.5096	0.5969	27.65	944
DenseNet121+LSTM	2.1585	0.4719	27.91	952

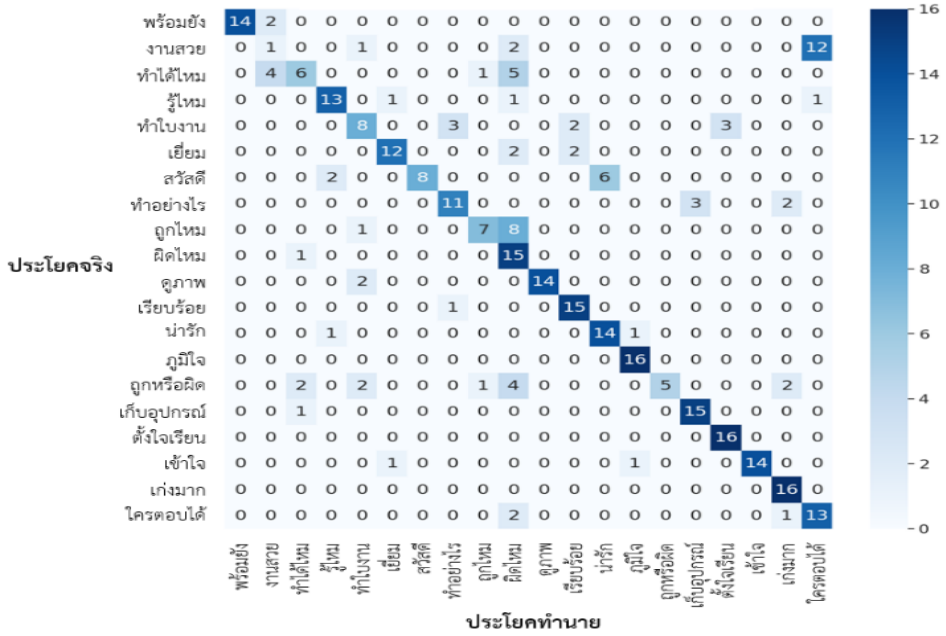


ภาพ 4 กราฟแสดงค่าความสูญเสียและค่าความถูกต้องของการฝึกสอนโมเดล EfficientNetB0 ร่วมกับ GRU

3. ผลการประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก

โมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกที่ผู้วิจัยพัฒนาขึ้น เป็นโมเดลจำแนกแบบหลายกลุ่มโดยใช้โมเดล CNN ร่วมกับ RNN จำนวนทั้งสิ้น 4 โมเดล และสามารถจำแนกประโยคภาษามือไทยได้จำนวน 20 ประโยค จากการพัฒนาโมเดลรู้จำภาษามือไทยพบว่าโมเดลที่มีประสิทธิภาพสูงที่สุดคือโมเดล EfficientNetB0 ร่วมกับ GRU ผู้วิจัยจึงนำโมเดลดังกล่าวมาประเมินประสิทธิภาพโดยใช้ชุดข้อมูลทดสอบ โดยวิเคราะห์ข้อมูลจากคอนฟิวชันเมทริกซ์ (confusion matrix) ซึ่งแสดงจำนวนตัวอย่างข้อมูลในประโยคที่โมเดลทำนายเทียบกับประโยคจริงว่าโมเดลทำนายได้ถูกต้องมากน้อยเพียงใด และมีการทำนายพลาดอย่างไรบ้าง

เมื่อให้โมเดลรู้จำภาษามือไทยจำแนกประโยคภาษามือในชุดข้อมูลทดสอบ โดยในแต่ละประโยคในชุดข้อมูลทดสอบจะมีข้อมูลจากวิดีโอจำนวนเท่ากันคือ 16 วิดิโอ ผลการจำแนกเมื่อเปรียบเทียบระหว่างประโยคจริงกับประโยคที่โมเดลทำนาย แสดงได้ด้วยคอนฟิวชันเมทริกซ์ขนาด 20×20 ดังภาพ 5 พบว่าประโยคที่โมเดลทำนายได้ถูกต้องทั้งหมดมีจำนวน 4 ประโยค ได้แก่ พร้อมยัง ภูมิใจ ตั้งใจเรียน และเก่งมาก ส่วนประโยคที่โมเดลทำนายได้ถูกต้องน้อยที่สุด คือ งานสวย ซึ่งมักมีการทำนายพลาดเป็น ใครตอบได้ นอกจากนี้ยังมีประโยคอื่น ๆ ที่โมเดลมักทำนายพลาดเกินครึ่งหนึ่งของข้อมูลตัวอย่างในประโยคนั้น ได้แก่ ทำได้ไหม ถูกไหม และถูกหรือผิด



ภาพ 5 คอนฟิวชันเมทริกซ์ของโมเดลรู้จำภาษามือไทย

เมื่อนำค่าที่ได้จากคอนฟิวชันเมทริกซ์มาคำนวณค่าความถูกต้อง (accuracy) เพื่อประเมินประสิทธิภาพการจำแนกของโมเดลรู้จำภาษามือในภาพรวม คำนวณค่าความแม่นยำ (precision) ค่าการเรียกคืน (recall) และคะแนน F1 (F1-score) เพื่อประเมินประสิทธิภาพการจำแนกของโมเดลรู้จำภาษามือในแต่ละประโยค พบว่า โมเดลมีค่าความถูกต้องเท่ากับ 0.73 ค่าความแม่นยำเฉลี่ยเท่ากับ 0.76 ค่าการเรียกคืนเฉลี่ยเท่ากับ 0.73 และคะแนน F1 เฉลี่ยเท่ากับ 0.72 ดังตาราง 4 ซึ่งแสดงให้เห็นว่าโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกที่พัฒนาขึ้นสามารถจำแนกประโยคภาษามือได้โดยมีประสิทธิภาพอยู่ในระดับปานกลาง

เมื่อประเมินประสิทธิภาพของโมเดลรู้จำภาษามือในแต่ละประโยค โดยพิจารณาประโยคภาษามือที่มีค่าความแม่นยำ ค่าการเรียกคืน และคะแนน F1 สูงสุดและต่ำสุด 3 อันดับแรก ดังตาราง 4 และภาพ 6 พบว่า

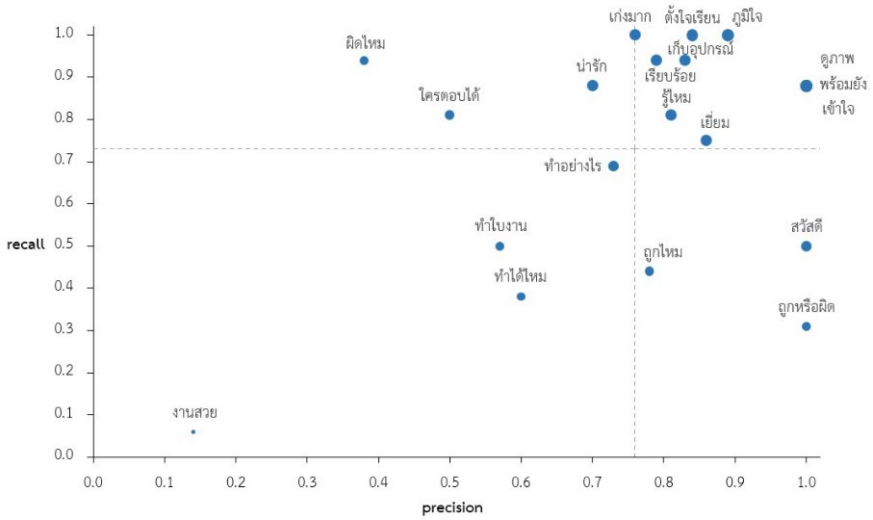
(1) ประโยคที่มีค่าความแม่นยำ (precision) มากที่สุด คือ พร้อมยัง ดูภาพ เข้าใจ สวัสดี และถูกหรือผิด มีค่าความแม่นยำเท่ากันคือ 1.00 เป็นประโยคที่โมเดลสามารถทำนายได้แม่นยำ ส่วนประโยคที่มีค่าความแม่นยำต่ำที่สุด คือ งานสวย ผิดไหม และทำใบงาน เป็นประโยคที่โมเดลสามารถทำนายได้ไม่แม่นยำ

(2) ประโยคที่มีค่าการเรียกคืน (recall) มากที่สุด คือ ภูมิใจ ตั้งใจเรียน และเก่งมาก มีค่าการเรียกคืนเท่ากันคือ 1.00 เป็นประโยคที่โมเดลไม่ค่อยพลาดการทำนาย ส่วนประโยคที่มีค่าการเรียกคืนต่ำที่สุด คือ “งานสวย” “ถูกหรือผิด” และ “ทำได้ไหม” เป็นประโยคที่โมเดลพลาดการทำนายบ่อยครั้ง

(3) ประโยคที่มีคะแนน F1 (F1-score) มากที่สุด คือ ภูมิใจ พร้อมยัง ดูภาพ และเข้าใจ เป็นประโยคที่โมเดลสามารถจำแนกได้ดี ส่วนประโยคที่มีค่าคะแนน F1 ต่ำที่สุดคือ คือ งานสวย ทำได้ไหม และถูกหรือผิด เป็นประโยคที่โมเดลยังไม่สามารถจำแนกได้ดี

ตาราง 4 ประสิทธิภาพของโมเดลรู้จำภาษามือไทย

ประโยค	precision	recall	F1 – score
ภูมิใจ	0.89	1.00	0.94
พร้อมยัง	1.00	0.88	0.93
ดูภาพ	1.00	0.88	0.93
เข้าใจ	1.00	0.88	0.93
ตั้งใจเรียน	0.84	1.00	0.91
เก็บอุปกรณ์	0.83	0.94	0.88
เรียบร้อย	0.79	0.94	0.86
เก่งมาก	0.76	1.00	0.86
รู้ไหม	0.81	0.81	0.81
เยี่ยม	0.86	0.75	0.80
น่ารัก	0.70	0.88	0.78
ทำอย่างไร	0.73	0.69	0.71
สวัสดี	1.00	0.50	0.67
ใครตอบได้	0.50	0.81	0.62
ถูกไหม	0.78	0.44	0.56
ผิดไหม	0.38	0.94	0.55
ทำใบงาน	0.57	0.50	0.53
ถูกหรือผิด	1.00	0.31	0.48
ทำได้ไหม	0.60	0.38	0.46
งานสวย	0.14	0.06	0.09
เฉลี่ย	0.76	0.73	0.72
accuracy = 0.73			



ภาพ 6 แผนภาพการกระจายค่า precision และ recall ของแต่ละประโยคภาษามือ โดยที่ขนาดจุดแทนค่า F1-score

อภิปรายผล

1. การศึกษาภาษามือที่ครูใช้สื่อสารกับนักเรียนที่มีความต้องการพิเศษ พบว่า ภาษามือที่ใช้มักเป็นประโยคที่สั้นและกะทัดรัดมากกว่าภาษาพูด เนื่องจากภาษามือไม่ได้มีคำศัพท์หลากหลายเท่ากับภาษาพูด และประโยคที่สั้นและกะทัดรัดนี้จะช่วยให้ครูและนักเรียนที่มีความต้องการพิเศษสามารถสื่อสารและทำความเข้าใจกันได้อย่างสะดวกและรวดเร็ว นอกจากนี้จากการสอบถามครูที่ใช้ภาษามือในการจัดการเรียนการสอนพบว่าภาษามือส่วนใหญ่ที่ครูใช้จัดอยู่ในหมวดการเสริมแรง ซึ่งเป็นการกล่าวชื่นชมและให้กำลังใจแก่นักเรียน ทั้งนี้เป็นเพราะว่าการเสริมแรงเป็นเทคนิคในการสร้างกำลังใจให้กับนักเรียนที่สามารถใช้ได้กับการสอนในทุกรายวิชา โดยประโยคที่ใช้ในการเสริมแรงแต่ละรายวิชามักจะเป็นประโยคในลักษณะเดียวกันทำให้ประโยคภาษามือในหมวดนี้มีการใช้งานมากที่สุด ในขณะที่ประโยคภาษามือในหมวดอื่นนั้นอาจมีการใช้งานที่แตกต่างกันออกไปตามแต่ละรายวิชา ดังนั้นในการเรียนรู้และฝึกฝนเกี่ยวกับการใช้ภาษามือเพื่อการสื่อสารของครูหรือนิสิตนักศึกษาครูที่ไม่ได้มีพื้นฐานเกี่ยวกับภาษามือสามารถเริ่มศึกษาได้จากภาษามือในหมวดการเสริมแรงที่สามารถใช้ได้กับทุกรายวิชา จากนั้นจึงศึกษาภาษามือในหมวดอื่นเพิ่มเติมที่แตกต่างกันไปในแต่ละรายวิชา

2. การพัฒนาโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึก พบว่า โมเดลรู้จำภาษามือไทยที่พัฒนาขึ้นโดยใช้โมเดล EfficientNetB0 ร่วมกับ GRU สามารถจำแนกประโยคภาษามือได้มีประสิทธิภาพสูงสุดในด้านความถูกต้อง เนื่องจากเป็นการใช้กลุ่มโมเดล CNN ที่สามารถดึงคุณลักษณะจากรูปภาพในแต่ละเฟรมของวิดีโอ และโมเดลกลุ่ม RNN ที่สามารถเรียนรู้ข้อมูลที่เป็นลำดับจากแต่ละเฟรมของวิดีโอ สอดคล้องกับงานวิจัยของ Puranik et al. (2022) ที่มีการใช้โมเดล VGG16 ร่วมกับ LSTM ในการรู้จำภาษามืออเมริกัน และงานวิจัย Huang & Chouvatut (2024) ที่มีการใช้โมเดล ResNet ร่วมกับ LSTM ในการรู้จำภาษามืออาร์เจนตินา ทำให้สามารถรู้จำภาษามือได้อย่างมีประสิทธิภาพ นอกจากนี้การใช้โมเดล EfficientNetB0 เป็นโมเดลที่มีผ่านการฝึกมาแล้ว (pretrained model) กับชุดข้อมูลขนาดใหญ่ และโมเดล EfficientNetB0 กับ GRU เป็นโมเดลที่มีขนาดเล็ก ทำให้ใช้เวลาในการฝึกสอนโมเดลน้อยกว่าโมเดลอื่น ๆ อย่างไรก็ตามโมเดลในการวิจัยนี้จะสามารถรู้จำภาษามือได้ดีในระดับหนึ่ง แต่ถ้าหากมีขนาดตัวอย่างที่ใช้ในการพัฒนาโมเดลจำนวนมากขึ้นก็จะช่วยให้โมเดลมีประสิทธิภาพสูงขึ้นได้อีก

3. การประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทย พบว่า โมเดลรู้จำภาษามือไทยที่มีประสิทธิภาพสูงสุดในการวิจัยนี้สามารถจำแนกประโยคภาษามือจำนวน 20 ประโยคได้ถูกต้องร้อยละ 73 จากประโยคภาษามือทั้งหมด เมื่อพิจารณาประสิทธิภาพในการจำแนกในแต่ละประโยคภาษามือโดยใช้ค่าคะแนน F1 พบว่าประโยคที่โมเดลสามารถจำแนกได้ดี ได้แก่ ภูมิใจ พร้อมยัง ดูภาพ และเข้าใจ อาจเป็นเพราะประโยคดังกล่าวเป็นประโยคที่มีท่ามือเพียง 1-2 ท่า ไม่มีการเคลื่อนไหวของมือที่ซับซ้อน เป็นเพียงการเคลื่อนไหวแค่ 1-2 จังหวะ และท่ามือกับการเคลื่อนไหวของมือมีความแตกต่างจากประโยคอื่น ๆ ในขณะที่ประโยคที่โมเดลจำแนกได้ไม่ดี ได้แก่ งานสวย ทำได้ไหม ถูกหรือผิด อาจเป็นเพราะประโยคดังกล่าวมีการเปลี่ยนท่ามือถึง 3 ท่า มีการเคลื่อนไหวของมือ 2-3 จังหวะ และท่ามือกับการเคลื่อนไหวของมือมีความคล้ายกับประโยคอื่น ๆ เช่นเดียวกับงานวิจัยของ Balaha et al. (2022) ที่พบว่าโมเดลรู้จำภาษามืออาหรับจำแนกภาษามือที่มีการเคลื่อนไหวคล้ายกันได้ไม่ดี สำหรับในการวิจัยนี้ประโยคที่โมเดลจำแนกได้ไม่ดี คือ “งานสวย” มักถูกทำนายผิดเป็น “ใคร

ตอบได้” เนื่องจากในท่าภาษามือของประโยคทั้งสองมีการเคลื่อนที่ของมือที่คล้ายกันคือมีการเคลื่อนที่ของมือวนรอบบริเวณใบหน้า นอกจากนี้คำว่า “งาน” กับ “ทำ” มีท่าทางและการเคลื่อนไหวที่คล้ายกันคือมีการกำมือและเคลื่อนมือเข้าหากัน และประโยคที่มีท่าทางคล้ายกันอีกคือ “ถูกไหม” “ผิดไหม” และ “ถูกหรือผิด” เนื่องจากประโยคทั้งสามมีคำที่เหมือนกันทำให้ท่าทางคล้ายกัน ซึ่งท่าทางที่คล้ายกันนี้อาจเป็นสาเหตุที่ทำให้โมเดลทำนายพลาดไป ทั้งนี้ สาเหตุที่โมเดลจำแนกประโยคที่มีท่ามือซับซ้อนและมีท่าคล้ายประโยคอื่น ๆ ได้ไม่ดีนั้น อาจเกิดขึ้นได้จากการกำหนดจำนวนเฟรมที่ดึงจากแต่ละวิดีโอมีจำนวนน้อย ทำให้การดึงรายละเอียดการเคลื่อนไหวของท่ามืออาจได้มาไม่ครบถ้วน ดังนั้นในการพัฒนาโมเดลครั้งต่อไปหากมีการกำหนดจำนวนเฟรมในแต่ละวิดีโอที่เหมาะสมในการพัฒนาโมเดล ก็จะช่วยให้โมเดลสามารถเรียนรู้รายละเอียดต่าง ๆ ของการเคลื่อนไหวแต่ละท่ามือได้มากขึ้น และทำให้โมเดลมีประสิทธิภาพสูงขึ้นได้

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1) ภาษามือที่ครูใช้ในการสื่อสารกับนักเรียนที่มีความต้องการพิเศษในการวิจัยนี้ ซึ่งประกอบด้วยหมวดการนำเข้าสู่บทเรียน การตรวจสอบความเข้าใจ การออกคำสั่ง การเสริมแรง และการโต้ตอบของนักเรียน สามารถนำไปใช้ในเป็นภาษามือพื้นฐานสำหรับครูและนิสิตนักศึกษาครูในการเรียนรู้และฝึกฝนการใช้ภาษามือในการสื่อสารกับนักเรียนที่มีความต้องการพิเศษ และภาษามือเหล่านี้สามารถนำไปพัฒนาโมเดลรู้จำภาษามือในชั้นเรียนที่มีประสิทธิภาพมากขึ้นได้

2) โมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกที่มีการใช้โมเดล EfficientNetB0 ร่วมกับ GRU ซึ่งในการวิจัยนี้พบว่ามีประสิทธิภาพสูงสุด สามารถนำไปพัฒนาเป็นโมเดลในแอปพลิเคชันสำหรับให้ครูและนิสิตครูสามารถเรียนรู้ประโยคภาษามือไทยพื้นฐานในชั้นเรียนและฝึกสอนการทำภาษามือ เพื่อให้สามารถสื่อสารกับนักเรียนที่มีความต้องการพิเศษได้อย่างมีประสิทธิภาพ

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

1) การศึกษาภาษามือที่ครูใช้ในการสื่อสารกับนักเรียนที่มีความต้องการพิเศษ สามารถศึกษาภาษามือที่ใช้ในแต่ละรายวิชา เนื่องจากแต่ละรายวิชามีคำศัพท์และประโยคที่เฉพาะตามเนื้อหาหรือวิธีการจัดการเรียนการสอนในรายวิชานั้น ๆ ซึ่งจะทำให้ภาษามือที่ใช้ในการพัฒนาโมเดลมีความหลากหลายมากยิ่งขึ้น ซึ่งสามารถนำโมเดลรู้จำภาษามือจากการวิจัยนี้ไปเป็นโมเดลที่ผ่านการฝึกมาแล้ว (pretrained model) สำหรับพัฒนาโมเดลรู้จำภาษามือครั้งต่อไปได้โดยไม่จำเป็นต้องเริ่มใหม่ตั้งแต่ต้น

2) การประเมินประสิทธิภาพโมเดลรู้จำภาษามือไทยโดยใช้การเรียนรู้เชิงลึกจากไฟล์วิดีโอ พบว่าโมเดลมีข้อจำกัดในการจำแนกบางประโยคภาษามือได้อย่างแม่นยำ ปัญหานี้เกิดจากความซับซ้อนของการเคลื่อนไหวในวิดีโอ เช่น การเปลี่ยนตำแหน่งของมือ ความสัมพันธ์เชิงเวลาของแต่ละเฟรมวิดีโอ และการเคลื่อนไหวที่คล้ายกันของประโยคที่ภาษามือแตกต่างกัน ส่งผลให้โมเดลมีแนวโน้มที่จะสับสนระหว่างประโยคที่มีลักษณะการเคลื่อนไหวคล้ายกัน ดังนั้นเพื่อปรับปรุงประสิทธิภาพของโมเดลสำหรับการใช้งานจริงในงานคอมพิวเตอร์วิทัศน์ (computer vision) โดยข้อมูลนำเข้าเป็นไฟล์วิดีโอ สามารถดำเนินการได้ดังนี้

(1) ขยายและปรับปรุงชุดข้อมูล เช่น เพิ่มตัวอย่างวิดีโอในกลุ่มประโยคที่มีการซับซ้อนร่วมกับการเพิ่มจำนวนข้อมูล (data augmentation) หรือเพิ่มจำนวนเฟรมของวิดีโอ และ (2) ปรับปรุงสถาปัตยกรรมของโมเดล โดยอาจเลือกใช้โมเดลในกลุ่ม attention หรือ transformer เช่น vision transformer (ViT) และ temporal attention mechanism ที่สามารถจับความสัมพันธ์ระหว่างเฟรมได้ดี ซึ่งอาจช่วยให้โมเดลมีประสิทธิภาพดีขึ้นได้

References

Amatanon, V., Chanhang, S., Naiyanetr, P., & Thongpang, S. (2014). Sign language-Thai alphabet conversion based on Electromyogram (EMG). In *2014 7th Biomedical Engineering International Conference (BMEiCON)* (pp. 1–4). IEEE. https://doi.org/10.1109/BMEiCON.2014.7017398

Balaha, M. M., El-Kady, S., Balaha, H. M., Salama, M., Emad, E., Hassan, M., & Saafan, M. M. (2022). A vision-based deep learning approach for independent-users Arabic sign language interpretation. *Multimedia Tools and Applications*, *82*(5), 6807-6826. https://doi.org/10.1007/s11042-022-13423-9

Chaikaew, A. (2022). An Applied Holistic Landmark with Deep Learning for Thai Sign Language Recognition. In *2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)* (pp. 1046-1049). IEEE. https://doi.org/10.1109/ITC-CSCC55581.2022.9895052

Chansri, C., & Srinonchat, J. (2016). Hand gesture recognition for Thai sign language in complex background using fusion of depth and color video. *Procedia Computer Science*, *86*, 257-260. https://doi.org/10.1016/j.procs.2016.05.113

Gedkhaw, E. (2022). The Performance of Thai Sign Language Recognition with 2D Convolutional Neural Network Based on NVIDIA Jetson Nano Developer Kit. *TEM Journal*, *11*(1), 411-419. https://doi.org/10.18421/TEM111-52

Huang, J., & Chouvatut V. (2024). Video-Based Sign Language Recognition via ResNet and LSTM Network. *Journal of Imaging*, *10*(6), 149. https://doi.org/10.3390/jimaging10060149

Kelly, J. F., McKinney E. L., & Swift, O. (2020). Strengthening teacher education to support deaf learners. *International Journal of Inclusive Education*, *26*(13), 1289–1307. https://doi.org/10.1080/13603116.2020.1806366

Klomsae, A., Auephanwiriyakul, S., & Theera-Umpon, N. (2017). A Novel String Grammar Unsupervised Possibilistic C-Medians Algorithm for Sign Language Translation Systems. *Symmetry*, *9*(12), 321. https://doi.org/10.3390/sym9120321

Pariwat, T., & Seresangtakul, P. (2021). Multi-stroke thai finger-spelling sign language recognition system with deep learning. *Symmetry*, *13*(2), 262. https://doi.org/10.3390/sym13020262

Puranik, V., Gawande, V., Gujarathi, J., Patani, A., & Rane, T. (2022). Video-based Sign Language Recognition using Recurrent Neural Networks. In *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)* (pp. 1-6). IEEE. https://doi.org/10.1109/ASIANCON55314.2022.9909061

Saengsri, S., Niennattrakul, V., & Ratanamahatana, C. A. (2012). TFRS: Thai finger-spelling sign language recognition system. In *2012 second international conference on digital information and communication technology and it's applications (DICTAP)* (pp. 457-462). IEEE. https://doi.org/10.1109/DICTAP.2012.6215407

Sharma, S., & Singh, S. (2023). ISL recognition system using integrated mobile-net and transfer learning method. *Expert Systems with Applications*, *221*, 119772. https://doi.org/10.1016/j.eswa.2023.119772

Sripairojthikoon, N., & Harnsomburana, J. (2019). Thai Sign Language Recognition Using 3D Convolutional Neural Networks. In *Proceedings of the 2019 7th International Conference on Computer and Communications Management* (pp. 186-189). Association for Computing Machinery. https://doi.org/10.1145/3348445.3348452

Vachirapipop, M., Soymat, S., Tiraronnakul, W., & Hnoohom, N. (2017, November). Sign Translation with Myo Armbands. In *2017 21st International Computer Science and Engineering Conference (ICSEC)* (pp. 1-5). IEEE. https://doi.org/10.1109/ICSEC.2017.8443836

Vijitkunsawat, W., Racharak, T., Nguyen, C., & Le Minh, N. (2023). Video-Based Sign Language Digit Recognition for the Thai Language: A New Dataset and Method Comparisons. In *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods* (pp. 775-782). SciTePress. https://doi.org/10.5220/0011643700003411

Translate Thai References

Boonya, R. (2008). Sign Language: Language of the Deaf. *Journal of Ratchasuda Institute, 4*(1), 77-94. (in Thai)

Buangamdee, C. (2017) *A Translation of Thai Language to Thai Sign Language with Rule-based Technique* [Master's thesis]. Naresuan University. (in Thai)

Jaruwanawat, A. (2007). *Continuous Thai Sign Language Recognition Using Hiden Markov Models* [Master's thesis]. King Mongkut Institute of Technology Ladkrabang. (in Thai)

Petchporee, S. (2005). *A Study on Sign Language Vocabulary Memory of Hearing Impaired Children in Prathomsuksa 1 by Minor Game* [Master's thesis]. Srinakarinwirot University. (in Thai)

Plianwong, N., & Kangkachit, T. (2021). Thai Sign Language Translation System Through Few Shot Learning. *DPU Graduate Studies Journal*, *10*(2), 119-132. (in Thai)

Sakulsujirapa, B. (2011). *Development of a Thai Sign Language Recognition System using Hiden Markov Models* [Master's thesis]. Prince of Songkla University. (in Thai)

Sriboonruang, P., Tanachai, P., Khaggathog, K. Vijitkunsawat, W., & Anunvrapong, P. (2022). Software Sign Language Translator to Text and Speech by Using the Landmarks Technique of MediaPipe. *The Journal of Industrial Technology : Suan Sunandha Rajabhat University*, *10*(2), 66-76. (in Thai)

Veerasakulthong, W. (2007). *Thai Hand Sign Recognition* [Master's thesis]. National Institute of Development Administration. (in Thai)