

Effects of Using Applications to Thai Phonetic Recognition According to the Reading Test of Grade 1 Students

ผลการใช้แอปพลิเคชันการรู้จำแนกเสียงภาษาไทยตามแบบทดสอบการอ่าน ของนักเรียนระดับชั้นประถมศึกษาปีที่ 1

Somphon Passon^{1,*}, and Wudhijaya Philuek²

สมพร พัสสร^{1,*}, และ วุฒิชัย พิลึก²

Received: 4 July 2024;

Revised: 3 August 2024;

Accepted: 9 August 2024;

Published: 15 October 2024;

Abstract

This research aims to study the effects of using applications to Thai phonetic recognition according to the reading test of grade 1 students. The data used in the study were 800 voice readings with 40 students by using the sound feature extraction including Spectrogram and Mel Frequency Cepstral Coefficient techniques in conjunction with the use of Neural Networks to train data and test models for application development. The study found that pronunciation training and similarity comparison from sound signals after using an application that can analyze the pronunciation of 800 sounds, and Thai pronunciation of 160 words in the reading test has an accuracy of more than 50%, which is 93.75%, and less than 50%, which is 6.25%, after using the application to detect the pronunciation, the accuracy is at a very good level.

Keywords: Application Development, Thai Phonetic Recognition, Thai Pronunciation.

บทคัดย่อ

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อศึกษาผลการใช้แอปพลิเคชันด้วยการรู้จำเสียงเพื่อจำแนกเสียงภาษาไทยตามแบบทดสอบการอ่านของนักเรียนระดับชั้นประถมศึกษาปีที่ 1 ข้อมูลที่ใช้ในการศึกษา คือ การอ่านออกเสียง จำนวน 800 เสียง กับนักเรียน 40 คน จากการใช้เทคนิควิธีการสกัดคุณลักษณะของเสียง ได้แก่ เทคนิคสเปกโตรแกรม และเทคนิคสัมประสิทธิ์เซปสตรัมบนสเกลเมล (MFCC)

^{1,*} Student, Computer Education Division, Faculty of Education, Nakhon Sawan Rajabhat University, Nakhon Sawan 60130, Thailand; นักศึกษา สาขาวิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏนครสวรรค์ จังหวัดนครสวรรค์ 60130 ประเทศไทย; Email: somporn.b5901821@gmail.com

² Assistant Professor, Dr., Computer Education Division, Faculty of Education, Nakhon Sawan Rajabhat University, Nakhon Sawan 60130, Thailand; ผู้ช่วยศาสตราจารย์ ดร. สาขาวิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏนครสวรรค์ จังหวัดนครสวรรค์ 60130 ประเทศไทย; Email: wudhijaya@nsru.ac.th

*Corresponding authors: Somphon Passon (somporn.b5901821@gmail.com)



ร่วมกับการนำโครงข่ายประสาทเทียมมาใช้ในการฝึกข้อมูล, และทดสอบโมเดลในการพัฒนาแอปพลิเคชัน ผลศึกษาพบว่า การฝึกอ่านออกเสียงและเปรียบเทียบความคล้ายคลึงจากสัญญาณเสียงในหลังการใช้แอปพลิเคชันที่สามารถวิเคราะห์การอ่านออกเสียงได้ 800 เสียง และการอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่านจำนวน 160 คำ มีค่าความแม่นยำมากกว่า 50% คิดเป็น 93.75% และต่ำกว่า 50% คิดเป็น 6.25% โดยหลังใช้แอปพลิเคชันในการตรวจจับการอ่านออกเสียง มีความถูกต้องอยู่ในระดับดีมาก

คำสำคัญ: การพัฒนาแอปพลิเคชัน, การรู้จำแนกเสียงภาษาไทย, การอ่านออกเสียงภาษาไทย

1. บทนำ (Introduction)

การศึกษาเป็นปัจจัยพื้นฐานสำคัญสำหรับการพัฒนาประชากรของประเทศให้มีคุณภาพอันจะนำไปสู่การพัฒนาประเทศให้มีความเข้มแข็งในทุกด้าน กระทรวงศึกษาธิการจึงกำหนดให้มีการปฏิรูปการศึกษา โดยกำหนดจุดเน้นการพัฒนาผู้เรียนที่ครอบคลุมทั้งด้านความรู้ ทักษะ และคุณลักษณะที่พึงประสงค์ตามจุดเน้นของผู้เรียนในศตวรรษที่ 21 โดยเฉพาะอย่างยิ่งความสามารถด้านการอ่านของผู้เรียนถือเป็นความสามารถพื้นฐานที่สำคัญสำหรับผู้เรียนในการเรียนรู้ในกลุ่มสาระการเรียนรู้ต่าง ๆ ตามหลักสูตรและการดำเนินชีวิตประจำวันของผู้เรียน ดังนั้น กระทรวงศึกษาธิการและสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน จึงให้ความสำคัญกับความสามารถด้านการอ่าน โดยเฉพาะอย่างยิ่งระดับช่วงชั้นที่ 1 ของระดับประถมศึกษา (ป.1 – ป.3) โดยได้กำหนดเป็นนโยบายเร่งด่วนใน ปีการศึกษา 2558 เด็กที่จบชั้นประถมศึกษาปีที่ 1 ต้องอ่านออก เขียนได้ การฟังเป็นทักษะทางภาษาอย่างหนึ่งที่มนุษย์เป็นเครื่องมือในการเรียนรู้และการแสวงหาความรู้ ในการเรียนรู้ในชีวิตประจำวันการฟังเป็นส่วนประกอบของการพูด (Pinno et al., 2011)

การฟังมีความสำคัญอย่างยิ่ง เพราะหากฟังแล้วเข้าใจจะสามารถสื่อสารได้ราบรื่นและยังลดความผิดพลาดในการสื่อสาร ความผิดพลาดในการสื่อสารจากการฟัง อาจพบได้จาก การขาดความรู้และประสบการณ์ในเรื่องที่ฟัง ขาดสมาธิในการฟัง อวัยวะในการฟังบกพร่อง มีอคติต่อคู่สนทนา จึงมีความจำเป็นต้องใช้เทคโนโลยีการรู้จำเสียง (Speech Recognition) ที่สามารถแปลคำพูดของมนุษย์ได้ออกมาเป็นข้อความ

การรู้จำเสียงเป็นกระบวนการที่เสียงคำหรือวลีของมนุษย์ถูกแปลงเป็นสัญญาณไฟฟ้าจากนั้นจึงแปลงเป็นรูปแบบการเข้ารหัสซึ่งกำหนดความหมายไว้เพื่อให้คำพูดสามารถใช้เป็นตัวนำเข้าไปในโปรแกรมคอมพิวเตอร์ได้ (Rabiner & Juang, 1993) โดยระบบรู้จำเสียงพูดที่สามารถทำงานได้แบบไม่เจาะจงผู้พูดจะต้องคัดเลือกลักษณะเฉพาะของคลื่นเสียงโดยเลือกเฉพาะลักษณะคลื่นที่มาจากคำพูดและจัดลักษณะคลื่นที่เกิดจากตัวผู้พูดออกไป นั่นคือ เลือกลักษณะเฉพาะที่มีเสถียรภาพไม่เปลี่ยนแปลงตามตัวผู้พูดไปวิเคราะห์เปรียบเทียบเพื่อหาว่าลักษณะเฉพาะนี้คล้ายกับการออกเสียงคำใดจะทำให้สามารถทราบได้ว่าลักษณะเฉพาะของเสียงนี้ใกล้เคียงกับลักษณะเฉพาะของเสียงใคร การทำงานในลักษณะนี้ สามารถนำไปประยุกต์ใช้กับการระบุหรือยืนยันตัวตนด้วยเสียงพูดได้ (NovaBizz, n.d.)

ยุคที่เทคโนโลยีเข้ามามีบทบาทในชีวิตประจำวัน ทำให้การเรียนการสอนมีการนำเอาเทคโนโลยีเข้ามาเกี่ยวข้องในรูปแบบของสื่อต่าง ๆ ทั้งวีดิทัศน์ คอมพิวเตอร์ช่วยสอน รวมไปถึงแอปพลิเคชัน โดยการใช้สื่อที่เหมาะสมกับผู้เรียนจะช่วยให้ผู้เรียนสามารถเรียนรู้ได้อย่างมีประสิทธิภาพ ซึ่งทำให้เกิดการพัฒนาสื่อการสอนในรูปแบบแอปพลิเคชันบนสมาร์ตโฟนหรือแท็บเล็ต โดยเป็นอีกช่องทางที่ทำให้การเรียนรู้ไม่จำกัดอยู่แต่ภายในห้องเรียน อีกทั้งยังสะดวกสามารถทบทวนบทเรียนได้ทุกที่ตามต้องการจึงมีความจำเป็นที่จะต้องใช้อุปกรณ์การเรียนรู้ที่เหมาะสมแก่บุคคลเหล่านี้ (Klangbungrong, 2012) โดยเฉพาะอย่างยิ่งการนำเทคโนโลยีในรูปแบบห้องเรียนเสมือนจริง (Virtual Classes) มาใช้ในการเรียนการสอน ทำให้ตอบสนองต่อ

การเรียนรู้ของผู้เรียนมีการจับกลุ่มกันทางอินเทอร์เน็ตเพื่อแบ่งปันสิ่งที่สนใจร่วมกัน รวมทั้งแบ่งปันความรู้ความชำนาญและทักษะความสามารถต่าง ๆ ร่วมกัน

จากข้อความข้างต้นแสดงให้เห็นว่า การอ่านออกเสียงของชั้นประถมศึกษาปีที่ 1 มีความสำคัญเป็นอย่างมากในการวางรากฐานให้ผู้เรียนมีการพัฒนาการด้านการคิดวิเคราะห์ รวมทั้งมีทักษะการสื่อสารและใช้ภาษาที่สามในการต่อยอดการเรียนรู้ ทั้งนี้ผู้ตรวจด้านการอ่านออกเสียง ต้องมีความรู้ความสามารถในการฟังที่ดีและสามารถวิเคราะห์เสียงออกมาได้ดี เกิดข้อผิดพลาดน้อยที่สุด ผู้วิจัยได้เล็งเห็นข้อบกพร่อง ในการฟังของผู้ตรวจด้านการอ่านออกเสียง เช่น การขาดความรู้และประสบการณ์ในเรื่องที่ฟัง ขาดสมาธิในการฟัง อวัยวะในการฟังบกพร่อง มีอคติต่อคู่สนทนา ผู้วิจัยเห็นความสำคัญจึงคิดพัฒนาแอปพลิเคชันด้วยเทคโนโลยีรู้จำเสียงเข้ามามีส่วนช่วยในการวิเคราะห์คลื่นเสียงของผู้เรียน

2. วัตถุประสงค์งานวิจัย (Research Objectives)

เพื่อศึกษาผลการใช้แอปพลิเคชันเทคโนโลยีรู้จำเสียงเพื่อจำแนกเสียงภาษาไทยตามแบบทดสอบการอ่านของนักเรียนระดับชั้นประถมศึกษาปีที่ 1

3. กรอบแนวคิดงานวิจัย (Conceptual Framework)

งานวิจัยนี้กำหนดกรอบแนวคิดในการวิจัย ดัง Figure 1.

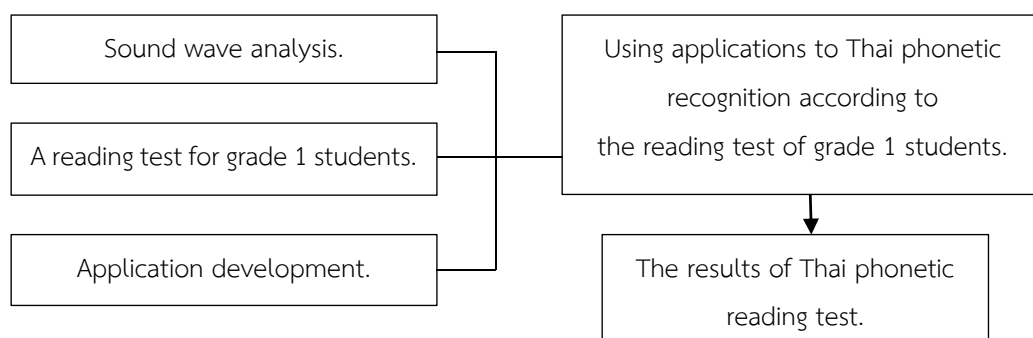


Figure 1. Conceptual Framework

4. การทบทวนวรรณกรรมและทฤษฎีที่เกี่ยวข้อง (Literature Review)

การพัฒนาแอปพลิเคชันเทคโนโลยีรู้จำเสียงเพื่อจำแนกประเภทผลการอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่านของนักเรียน ระดับชั้นประถมศึกษาปีที่ 1 ครั้นนี้ ผู้วิจัยได้ศึกษาเอกสาร แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้องในการนำมาเป็นข้อมูลพื้นฐานตลอดจนกรอบแนวคิดวิจัยข้างต้น อาทิ การรู้จำเสียงในการช่วยฝึกทักษะการอ่านออกเสียงการผันวรรณยุกต์ภาษาไทยสำหรับนักเรียนชั้นประถมศึกษา (Philuek & Puttasem, 2023) โดยเทคนิคที่ใช้ในการวิเคราะห์ข้อมูลมี 3 รูปแบบ ดังนี้ 1) Spectrogram 2) Wavelet Transform 3) Dynamic Time Warping โดยนำเทคนิคเปลี่ยนแปลงของคลื่นเสียง Spectrogram ร่วมกับ เทคนิคแปลงรูปแบบเสียง Wavelet Transform เพื่อแปลงสัญญาณที่มีอยู่เดิมให้เหมาะสมก่อนที่จะนำเทคนิค Dynamic Time Warping (DTW) ในการวัดความคล้ายกันของข้อมูลเสียง 2 ชุด จากการใช้แต่ละเทคนิคร่วมกันวิเคราะห์สามารถเปรียบเทียบและแสดงให้เห็นว่าจะได้ค่าความคล้ายสัญญาณเสียง

ผลการศึกษาเทคนิคการวิเคราะห์ข้อมูลของระบบรู้จำเสียงในการช่วยฝึกทักษะการอ่านออกเสียงการผันวรรณยุกต์ภาษาไทยสำหรับนักเรียนระดับชั้นประถมศึกษา พบว่า เทคนิคที่เกี่ยวข้องกับการวิเคราะห์เสียง 3 ประกอบไปด้วย เทคนิค

การเปลี่ยนแปลงของคลื่นเสียง Spectrogram เทคนิคแปลงรูปแบบเสียง Wavelet Transform และเทคนิคเปรียบเทียบความคล้ายของข้อมูล 2 ชุด Dynamic Time Warping (DTW) โดยใช้ผู้วิจัยได้นำเทคนิคต่าง ๆ ร่วมกันในการวิเคราะห์และเปรียบเทียบให้เห็นว่าทำให้มีค่าความคล้ายสัญญาณเสียงที่มากขึ้น กว่าการใช้เทคนิคเดียวในการวิเคราะห์ ผลการใช้ระบบรู้จำเสียงในการช่วยฝึกทักษะการอ่านออกเสียงการผันวรรณยุกต์ภาษาไทยสำหรับนักเรียนระดับชั้นประถมศึกษา พบว่า ผลการฝึกออกเสียงและเปรียบเทียบความคล้ายของสัญญาณเสียง หลังจากใช้เทคนิคต่าง ๆ ร่วมกันวิเคราะห์ จำนวน 5 ครั้ง มีผลค่าเฉลี่ยมากที่สุดคิดเป็นร้อยละ 93.52 ผลค่าเฉลี่ยน้อยที่สุดคิดเป็นร้อยละ 87.75 ผลเฉลี่ยรวมคิดเป็นร้อยละ 90.66

นอกจากนี้ยังมีผลงาน Malangpoo et al. (2022) ที่นำเสนอเรื่องผลการใช้ระบบรู้จำเสียงในการช่วยฝึกการออกเสียงภาษาจีนของนักเรียนระดับมัธยมศึกษา ดำเนินเรื่องวิจัยการสำรวจเก็บรวบรวมข้อมูลปัญหาในการเรียนภาษาจีน โดยใช้แบบสอบถามความคิดเห็นของครูที่สอนภาษาจีน นำผลการสำรวจเก็บรวบรวมข้อมูลปัญหาในการเรียนภาษาจีนโดยใช้แบบสอบถามเกี่ยวกับการออกเสียงภาษาจีน มาสรุปผลคำที่ออกเสียงผิดบ่อย และนำข้อมูลเสียงทั้งหมดมาเรียบเรียงวิเคราะห์คลื่นเสียงโดยใช้หลักการ การวิเคราะห์อนุกรมเวลา (Time Series Analysis) ลำดับของค่าที่เรียงตามเวลาโดยคลื่นเสียงได้รับการวิเคราะห์หลายเทคนิคโดยผู้วิจัยได้เลือกเทคนิคดังต่อไปนี้

เทคนิค Spectrogram เทคนิค Wavelet Transform เทคนิค Dynamic Time Warping (DTW) และเทคนิค Mel Frequency Cepstral Coefficient (MFCC) ในการจับคำบางคำเพื่อเทียบความคล้ายของสัญญาณเสียงเสียงพูด ผลวิเคราะห์การรู้จำเสียงในการช่วยฝึกการออกเสียงภาษาจีนของนักเรียนระดับมัธยมศึกษาประกอบไปด้วย เทคนิค Spectrogram เทคนิค Wavelet Transform เทคนิค Dynamic Time Warping (DTW) และเทคนิค Mel Frequency Cepstral Coefficient (MFCC) โดยใช้ 4 เทคนิคร่วมกันทำให้มีค่าความคล้ายของเสียงมากกว่าการใช้เทคนิคเดียวหรือใช้ 3 เทคนิคร่วมกัน ในการศึกษาผลการใช้ระบบรู้จำเสียงในการช่วยฝึกการออกเสียงภาษาจีนของนักเรียนระดับมัธยมศึกษา ผลของความคล้ายของสัญญาณเสียงโดยใช้เทคนิค Wavelet Transform มีความคล้ายของสัญญาณเสียงคิดเป็นร้อยละ 69.71 ผลของความคล้ายของสัญญาณเสียงโดยใช้เทคนิค Wavelet Transform ร่วมกับการดึงคุณลักษณะเด่นของสัญญาณเสียงพูดสัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel Frequency Cepstral Coefficient (MFCC)) มีความคล้ายของสัญญาณเสียงคิดเป็นร้อยละ 75.42 ผลของความคล้ายของสัญญาณเสียงโดยใช้เทคนิค Wavelet Transform ร่วมกับการดึงคุณลักษณะเด่นของสัญญาณเสียงพูดสัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel Frequency Cepstral Coefficient (MFCC)) และการเปรียบเทียบค่าความถี่ของสัญญาณเสียงหลังจากการปรับปรุงสัญญาณเสียงพูดแล้ว (Spectrogram) มีความคล้ายของสัญญาณเสียงคิดเป็นร้อยละ 80.36 ผลของความคล้ายของสัญญาณเสียงโดยใช้เทคนิค Wavelet Transform ร่วมกับการดึงคุณลักษณะเด่นของสัญญาณเสียงพูดสัมประสิทธิ์เซปสตรัมบนสเกลเมล (Mel Frequency Cepstral Coefficient (MFCC)) การเปรียบเทียบค่าความถี่ของสัญญาณเสียงหลังจากการปรับปรุงสัญญาณเสียงพูดแล้ว (Spectrogram) และการเปรียบเทียบความคล้ายกันระหว่างข้อมูล 2 ชุด (Dynamic Time Warping (DWT)) มีความคล้ายของสัญญาณเสียงคิดเป็นร้อยละ 82.39

ตลอดจน Angsuphan (2019) นำเสนอเรื่องการทดสอบความคงที่ของเสียงพูดสำหรับกลุ่มผู้ป่วยเด็กสมองพิการที่มีการพูดแบบดิสอาร์เทรีย โดยเล่าถึง การพัฒนาระบบรู้จำสำหรับเด็กสมองพิการที่บกพร่องทางเสียงนั้น เราจำเป็นต้องมีเครื่องมือในการประเมินผลเบื้องต้น สำหรับใช้เป็นแนวทางในการเลือกลักษณะของระบบที่เหมาะสมกับความบกพร่องแต่ละแบบ เช่น หน่วยการรู้จำที่เหมาะสมหรือสวณะลักษณะ (Acoustic Features) แบบต่าง ๆ เพื่อเพิ่มความถูกต้องในการรู้จำเสียงพูดโดยมีวิธีการดำเนินการ 1) นำสัญญาณเสียงพูดมาจากคลังข้อมูลมาสกัดสวณะลักษณะแบบ MFCC เพื่อใช้เป็นตัวแทนสัญญาณเสียงเหล่านั้น 2) นำ MFCC ที่ได้มาคำนวณหาค่าความคล้าย (SIM) และความต่าง (DIS) โดยใช้เทคนิคไดนามิกไทม์วอร์ปิง 3) หาค่าความคงที่ของเสียงพูด (Speech Consistency Score, SCS) จากค่าความคล้ายและความต่าง บางคนมีข้อจำกัดที่ทำให้ไม่สามารถใช้ระบบได้ ดังนั้นการประเมินผลเบื้องต้นก่อนพัฒนาระบบจึงมีความจำเป็นอย่างมาก อย่างไรก็ตาม

การประเมินมาตรฐานที่กล่าวมา ยังไม่สามารถนำมาใช้ทำนายประสิทธิภาพของระบบการรู้จำเมื่อใช้กับผู้พิการแต่ละคนได้ เก็บข้อมูลในรูปแบบของ คลังข้อมูลเสียงพูดประกอบด้วยเสียงพูด 68 คำ บันทึกจากกลุ่มเด็กปกติ 4 คนและกลุ่มเด็กสมองพิการ 8 คน จำนวน 5 รอบ ถูกนำมาใช้ในการคำนวณหาค่าความคงที่ของเสียงพูด (SCS) สำหรับผู้รับการทดสอบแต่ละคน เมื่อนำค่า SCS ไปเปรียบเทียบกับอัตราการเรียนรู้จำเสียงพูดที่ใช้แบบจำลองฮิดเดน มาร์คอฟ (HMMs) และข่ายงานประสาทเทียม (ANNS) พบว่าผู้รับการทดสอบที่มีค่า SCS สูง จะมีอัตราการเรียนรู้จำเสียงพูดสูง ส่วนผู้รับการทดสอบที่มีค่า SCS ต่ำก็จะมีอัตราการเรียนรู้จำเสียงพูด เช่นกัน จากนั้น ผู้วิจัยได้นำค่า SCS มาหาความสัมพันธ์เชิงเส้นกับอัตราการเรียนรู้จำเสียงพูดที่ใช้ HMM และ ANN ซึ่ง ได้ค่าสัมประสิทธิ์การกำหนด (R) เท่ากับ 0.89 และ 0.86 ตามลำดับ ซึ่งอยู่ในเกณฑ์ดีมากท้ายสุด ผู้วิจัยได้นำค่า SCS ไปทำนายอัตราการเรียนรู้จำเสียงพูด พบว่าผลการทำนายที่ได้มีคลาดเคลื่อน สัมบูรณ์เฉลี่ยจากอัตราการเรียนรู้จำจริงเท่ากับ 5% ซึ่งเป็นระดับที่น่าพอใจมาก นอกจากนี้ยังมีงานวิจัยของ Charoendee (2016) ที่นำเสนอเรื่องการคัดเลือกและสกัดคุณลักษณะเพื่อการจำแนกอารมณ์จากเสียงพูดภาษาไทย โดยมุ่งศึกษาการเรียนรู้จำอารมณ์จากเสียงพูด (Pitaksiriant, 2011) นำเสนอเรื่องการประเมินการพูดที่ฟังไม่รู้เรื่องอัตโนมัติในผู้ป่วยปากแหว่งเพดานโหว่ (Rockkittichareon, 2011) นำเสนอเรื่องพารามิเตอร์ทางเสียงสำหรับการจำแนกลักษณะการเปล่งเสียงในเสียงพูดต่อเนื่องภาษาไทย ซึ่งจากการทบทวนวรรณกรรมดังกล่าวทำให้ผู้วิจัยได้นำมาเป็นแนวทางในการศึกษาในครั้งนี้

5. วิธีดำเนินงานวิจัย (Research Methodology)

การวิจัยนี้ผู้วิจัยเลือกกลุ่มเป้าหมายเพื่อดำเนินการเก็บตัวอย่างเสียงแล้วนำมาวิเคราะห์เพื่อสร้างขึ้นเป็นโมเดลทดสอบโดยเก็บข้อมูลจากเด็กที่มีความสามารถในการอ่านออกเสียงหลังจากนั้นจึงดำเนินการฝึกข้อมูลเสียงจากโมเดลทดสอบและจึงพัฒนาแอปพลิเคชันแล้วจึงนำไปทดลองใช้เพื่อหาประสิทธิภาพและความแม่นยำโดยมีขั้นตอนดังนี้

1. ศึกษาค้นคว้าหาข้อมูลเกี่ยวกับคำภาษาไทยโดยนำมาจากแบบทดสอบการอ่านของสำนักทดสอบทางการศึกษาที่ใช้ในการทดสอบการอ่านของเด็กระดับประถมศึกษาปีที่ 1 จำนวน 20 คำ ดังนี้ 1) ดุแล 2) หน้าจอ 3) เน้นแนว 4) ลูกชิ้น 5) รังแก 6) กบข้าว 7) อาบน้ำ 8) น่ารัก 9) วิตามิน 10) นามสกุล 11) น้ำแข็ง 12) รถรับส่ง 13) เครื่องบิน 14) สวยงาม 15) ศาสนา 16) พิเศษ 17) ร้องเพลง 18) สร้างบ้าน 19) ขยัน 20) ตลอด
2. เตรียมแบบทดสอบการอ่านออกเสียงเพื่อนำไปเก็บเสียงกับกลุ่มเป้าหมายที่ใช้สำหรับเก็บข้อมูลเสียงเพื่อใช้พัฒนาโมเดล จำนวน 40 คน
3. จากนั้นสร้างแบบฟอร์มเก็บเสียงทั้งหมดจำนวน 20 คำ จากเด็กที่มีความสามารถในการอ่านออกจำนวน 40 คน รวมทั้งหมด 800 เสียง
4. การตัดเสียงและทำการแปลงเสียงให้เป็นนามสกุล .wav
5. ทำการปรับจนความถี่ให้เท่ากัน เพื่อง่ายต่อการเปรียบเทียบ
6. ทำการสร้างโมเดลใช้ในการเทรนวิเคราะห์เปรียบเทียบข้อมูลจากเสียงทั้งหมด 20 คำ จำนวน 40 คน โดยแต่ละคำจะใช้ข้อมูล 40 เสียงใช้ในการเทรนข้อมูล รวมทั้งหมด 800 เสียง
7. นำเสียงที่ทำการปรับจนแล้วทั้งหมด 20 คำ มาเปรียบเทียบกับเสียงต้นฉบับ นำผลแต่ละเทคนิคมาวิเคราะห์หาความแม่นยำ
8. ออกแบบแอปพลิเคชันโดยการออกแบบฐานข้อมูลเก็บข้อมูลเสียงนักเรียนและออกแบบหน้าเว็บแอปพลิเคชันในการใช้งานรูปแบบวิเคราะห์คลื่นเสียงและแสดงผลออกมาในรูปแบบที่สามารถใช้งานได้ง่าย

5.1 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย ประกอบด้วยแบบทดสอบการอ่านของนักเรียนระดับประถมศึกษาปีที่ 1 จัดทำในรูปแบบออนไลน์โดยนำแบบฟอร์มมาทำเป็น Google Forms อุปกรณ์ในการทำงาน ได้แก่ เครื่องคอมพิวเตอร์ เทคนิคที่ใช้ในการ

วิเคราะห์ข้อมูล ประกอบด้วย เทคนิค Spectrogram และ เทคนิค Mel Frequency Cepstral Coefficient (MFCC) โปรแกรม Ocenaudio โปรแกรม Visual Studio Code เครื่องมือที่ใช้ในการทำงาน ได้แก่ Database ใช้ phpMyAdmin และ FastAPI

5.2 การวิเคราะห์ข้อมูล

งานวิจัยวิจัยมีขั้นตอนการดำเนินการวิเคราะห์ข้อมูลผ่านกระบวนการพัฒนาโดยมีการทำงานแต่ละขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 ทำความเข้าใจปัญหา (Business Understanding) ผู้วิจัยได้ศึกษางานวิจัยที่เกี่ยวข้อง และได้ศึกษาข้อมูลเกี่ยวกับปัญหาในด้านการฟังของผู้ตรวจการอ่านออกเสียง (Read Test) การวิเคราะห์และเทคโนโลยีการรู้จำเสียงพูด การใช้เทคนิค Spectrogram และเทคนิค Mel Frequency Cepstral Coefficient (MFCC) มาวิเคราะห์รูปแบบเสียงโดยการthrenข้อมูล อันจะนำสู่ระบบและเครื่องมือช่วยในการคัดกรองการอ่านออกเสียงภาษาไทยตามแนวทางการทดสอบ (Read test) ในอนาคต

ขั้นตอนที่ 2 ทำความเข้าใจข้อมูล (Data Understanding) ผู้วิจัยได้ศึกษาข้อมูลที่มีอยู่พร้อมพิจารณาข้อมูลและความเป็นไปของการนำข้อมูลมาวิเคราะห์ มีวิธีการดังนี้ การเก็บรวบรวมข้อมูลข้างต้น ขั้นตอนนี้ผู้วิจัยได้ทำการคัดเลือกเด็กที่สามารถอ่านออกเสียงได้ นำมาอัดเสียง จากแบบทดสอบการอ่านออกเสียง (Reading Test) จากสำนักทดสอบทางการศึกษา จำนวน 20 คำ ของนักเรียนที่มีความสามารถอ่านออกเขียนได้ 40 คน รวมเสียงที่เก็บทั้งหมด 800 เสียง เพื่อใช้เป็นเสียงต้นแบบในการวิเคราะห์เสียง

ขั้นตอนที่ 3 การเตรียมข้อมูล (Data Predation) กระบวนการเตรียมข้อมูลโดยจะนำข้อมูลดิบที่มีแปลงให้เป็นข้อมูลที่จะต้องใช้ในขั้นตอนต่อไปตลอดจนการเลือกตัวแปรที่ต้องวิเคราะห์ให้เหมาะสม และแปลงรูปแบบของตัวแปรให้อยู่ในรูปแบบเดียวกันทั้งหมดเพื่อเตรียมข้อมูลให้มีความพร้อมสำหรับนำไปสร้างแบบจำลอง โดยข้อมูลแบ่งออกเป็นขั้นตอนย่อยอีก 3 ขั้นตอน ได้แก่

1) การเก็บข้อมูลเสียงผู้วิจัยได้ทำการแปลงข้อมูลเสียงที่ได้มา นำเสียงมาตัดให้เป็นคำ แล้วเปลี่ยนเสียงเป็นรูปแบบของ mono และแปลงไฟล์เป็นสกุลไฟล์ .wav เพื่อนำไปวิเคราะห์ในส่วนต่อไป โดยผู้วิจัยได้ใช้โปรแกรม Ocenaudio โดยมีวิธีการดังนี้ (1) เปิดใช้โปรแกรม Ocenaudio (2) นำเข้าข้อมูลเสียงเพื่อทำโดยการใช้โปรแกรม Ocenaudio ดัง Figure 2.

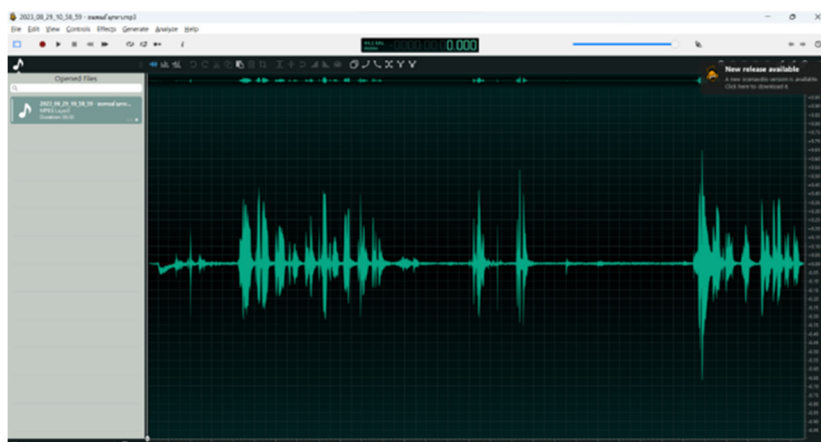


Figure 2. Importing audio data using the Ocenaudio software.

จากนั้น (3) ทำการแปลงไฟล์จาก stereo เป็น mono ทำให้ขนาดไฟล์เล็กง่ายต่อการคำนวณ (4) ทำการตัดเสียงออกเป็นแต่ละคำ (5) ทำการตัดเสียงออกเป็นแต่ละคำโดยบันทึกเสียงเป็นไฟล์นามสกุล .wav และทำการเก็บข้อมูลเสียงไว้ยังโฟลเดอร์

2) จากการตัดเสียงแบ่งข้อมูลออกเป็น 20 คำแล้วทำการปรับรูปแบบเสียง เวลาความถี่ คำที่ใช้ในการคำนวณให้ตรงกับข้อมูลให้ตรงกันในแต่ละคำ เพื่อจะนำมาวิเคราะห์ข้อมูลได้ง่ายโดยผู้ทำวิจัยได้มีวิธีการเขียนโค้ดเพื่อปรับค่าเวลา และความถี่ที่ใช้ในการคำนวณให้ตรงกัน ดัง Figure 3.

```
def Convert(inputs,outputs):
    target_wav_time = 5000
    original_segment = AudioSegment.from_wav(inputs)
    silence_duration = target_wav_time - len(sound)
    silenced_segment = AudioSegment.silent(duration=silence_duration)
    combined_segment = sound + silenced_segment
    combined_segment.export(outputs, format="wav")
```

Figure 3. Coding for equalizing time and frequency.

ขั้นตอนที่ 4 การวิเคราะห์ข้อมูล (Data Analytics) เป็นการนำข้อมูลเสี่ยงที่ผ่านการแปลงสกุลไฟล์ และตัดเป็นคำมาแล้ว นำมาใช้ในการวิเคราะห์เทียบความคล้ายของเสี่ยง โดยมีขั้นตอนดังนี้

1) ผู้วิจัยได้เลือกใช้เทคนิคการวิเคราะห์ 1) เทคนิค Spectrogram 2) เทคนิค Mel Frequency Cepstral Coefficient (MFCC) ในการวิเคราะห์ข้อมูล โดยใช้เครื่องมือ Jupyter Notebook ในการสร้างโมเดลการเรียนรู้ภาษาไพธอน

2) สร้างโมเดลการเรียนรู้โดยใช้ ด้วยไลบรารี Tensorflow และ Keras โดยในการสร้าง Deep Neural Network ใช้ข้อมูลเสียง 20 คำ โดยแต่ละคำจะมีเสียงในการเทรน 32 เสียง ซึ่งอีก 8 เสียง ใช้ในการทดสอบ รวมแล้วใช้เสียงทั้งหมด 800 เสียง ในแต่ละคำ ใช้เทคนิค Spectrogram ดัง Figure 4. (a) และการใช้เทคนิค Mel Frequency Cepstral Coefficient (MFCC) ดัง Figure 4. (b)

[illegible]

(a) Spectrogram

```

def record_to_mfcc(audio_file):
    x = AudioSegment.from_raw(audio_file)
    x = np.array(x.get_array_of_samples(), np.float32)
    return librosa.feature.mfcc(y=x)

X = []
path = ""
speech = ["cat", "dog", "elephant", "giraffe", "lion", "monkey", "tiger", "bird", "frog", "sheep", "goat", "horse", "cow", "pig", "kangaroo", "squirrel", "whale", "lioness", "kitten", "kitten", "kitten"]

for i in range(len(speech)):
    get_path = os.path.join(path + "speech" + "/" + speech[i] + ".mp3")
    for i in tqdm(range(len(speech))):
        file_1 = speech[i] + ".mp3"
        x = record_to_mfcc(path + "/" + file_1)
        print(x.shape)
        X.append(x)
    X.append(x.shape)

from sklearn.model_selection import train_test_split
X = np.array(X)
y = np.array(y)
X = X.reshape(X.shape[0], X.shape[1], X.shape[2], 1)
x_train, x_test, y_train, y_test = train_test_split(X, y, train_size = 0.8, random_state=42)
x = x_train

from tensorflow import keras
from sklearn.model_selection import train_test_split
import tensorflow as tf

model = keras.Sequential()
model.add(keras.layers.Conv2D(64, 5, input_shape=(25, 431, 1), activation='relu'))
model.add(keras.layers.Dropout(0.5))
model.add(keras.layers.MaxPooling2D())
model.add(keras.layers.Conv2D(64, 5, activation='relu'))
model.add(keras.layers.Dropout(0.5))
model.add(keras.layers.MaxPooling2D())
model.add(keras.layers.Flatten())
model.add(keras.layers.Dense(128, activation='relu'))
model.add(keras.layers.Dense(64, activation='relu'))
model.add(keras.layers.Dense(32, activation='softmax'))
model.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])

history=model.fit(x_train, validation_data=(x_test, y_test), epochs=100, verbose=1)

```

(b) Mel Frequency Cepstral Coefficient (MFCC)

Figure 4. Creating a learning model using Tensorflow and Keras libraries.

3) ทดสอบเสียงด้วยเสียงดังต้น 1 เสียง กับเสียงที่ทำการเทรนทั้งหมด 20 คำ จำนวนเสียงในแต่ละคำ 32 เสียง รวมทั้งหมด 640 เสียง จะได้ค่าความแม่นยำ และค่าความผิดพลาดออกมา ดัง Table 1.

Table 1. Accuracy and error values for selecting the feature extraction method for sound pattern analysis.

Words	Spectrogram		Mel Frequency Cepstral Coefficient	
	Accuracy values	Error values	Accuracy values	Error values
ดูแล (look after)	1.000	4.768	1.000	0.004
หน้าจอ (screen)	0.963	3.254	1.000	0.000
แนะแนว (guide)	1.000	2.365	1.000	0.002
ลูกชิ้น (meatball)	0.996	3.354	1.000	0.005
วิ่งแข่ง (race)	0.865	0.931	0.959	0.000
กับข้าว (side dish)	0.991	0.008	1.000	0.006
อาบน้ำ (shower)	0.831	4.365	0.964	0.008
น่ารัก (cute)	1.000	3.667	1.000	0.026
วิตามิน (vitamin)	1.000	0.985	1.000	0.000
นามสกุล (surname)	0.864	3.587	0.988	0.005
น้ำแข็ง (ice)	0.961	0.569	1.000	0.007
รถรับส่ง (shuttle)	0.782	3.551	0.960	0.000
เครื่องบิน (airplane)	0.812	2.251	1.000	0.000
สวยงาม (beautiful)	0.961	0.236	1.000	0.002
ศาสนา (religion)	1.000	0.112	0.955	0.000
พิเศษ (special)	0.954	0.237	1.000	0.025
ร้องเพลง (sing)	0.915	0.254	0.978	0.000
สร้างบ้าน (build a house)	1.000	3.254	1.000	0.000
ขยัน (hard-working)	1.000	2.446	1.000	0.028
ตลอด (always)	0.861	2.698	0.998	0.002

5.3 การพัฒนาแอปพลิเคชัน

งานวิจัยวิจัยออกแบบตามหลักการทำงานของโปรแกรมตามลำดับดังนี้ 1) ผู้ใช้งานใช้งานแอปพลิเคชัน 2) ส่งข้อมูลไปยังหลังบ้านโดยใช้ API พร้อมคำนวณเสียง 3) ส่งข้อมูลเพื่อเก็บข้อมูลไปยังดาต้าเบส 4) API คัดเลือกข้อมูลมาจาก Database 5) API ส่งข้อมูลไปแสดงผลต่อผู้ใช้งาน และการจำลองฐานข้อมูลเพื่อปรับแก้ไขในการพัฒนาแอปพลิเคชันสร้างตาราง 2 ตาราง ได้แก่ 1) reading_master และ 2) users และสร้างฐานข้อมูลโดยใช้ phpMyAdmin ที่เลือกมาเพราะเป็น Tool Database ที่สามารถใช้งานโดยไม่เสียค่าใช้จ่าย

จากนั้นใช้ FastAPI ในการสร้างเส้นส่งข้อมูลระหว่างผู้ใช้งานกับระบบหลังบ้าน เนื่องจาก FastAPI ใช้งานได้ง่ายและยังสามารถเข้ากับโมเดลของงานนี้ได้ดีโดยมีภาษาเขียนที่เหมือนกันคือภาษาไพธอน ดัง Figure 5. และสร้างหน้าอินเตอร์เฟซของการบันทึกเสียงเพื่อนำมาวิเคราะห์ โดยใช้โปรแกรม Vs Code ซึ่งมีหน้าดังต่อไปนี้ หน้าเข้าสู่ระบบเพื่อแสดงข้อมูลผู้เข้าใช้

งาน หน้าอินเทอร์เน็ตเข้าสู่ระบบ เพื่อนำเสียงมาวิเคราะห์จึงได้ทำการสร้างหน้า เลือกใช้งานและเก็บเสียง หน้าการเลือกผู้ใช้งานทดสอบเสียง (Figure 6.) และหน้าการใช้งานบันทึกเสียงเพื่อนำเสียงมาวิเคราะห์ของคำ (Figure 7.)

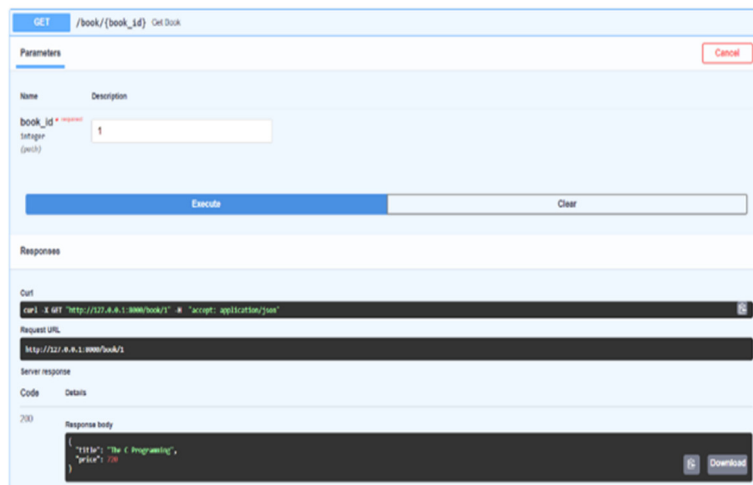


Figure 5. FastAPI testing

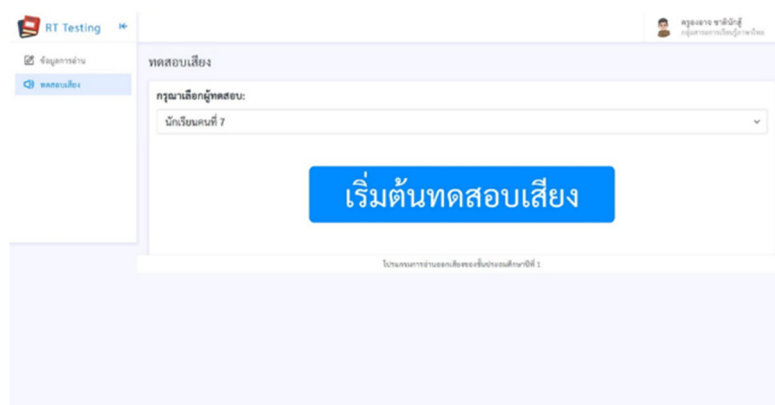


Figure 6. Page for selecting voice test user.

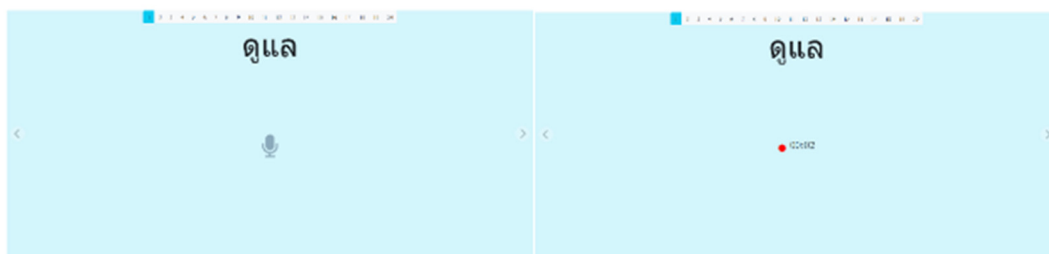


Figure 7. Page for using voice recording to analyze the sound of words.

6. ผลการวิจัย (Results)

6.1 ผลการใช้แอปพลิเคชันช่วยในการวิเคราะห์การอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่านของนักเรียนระดับชั้นประถมศึกษาปีที่ 1

ผลการใช้งานเมื่อเก็บเสียงทั้งหมดเสร็จ โปรแกรมจะแสดงผลการอ่าน ดัง Figure 8. และแสดงผลเป็นรูปแบบตารางเพื่อใช้ในการตรวจสอบอีกครั้งของแต่ละคน ดัง Figure 9.

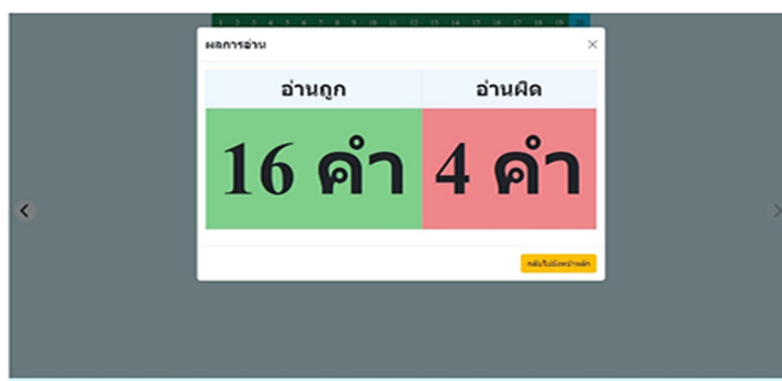


Figure 8. The reading test of grade 1 students.

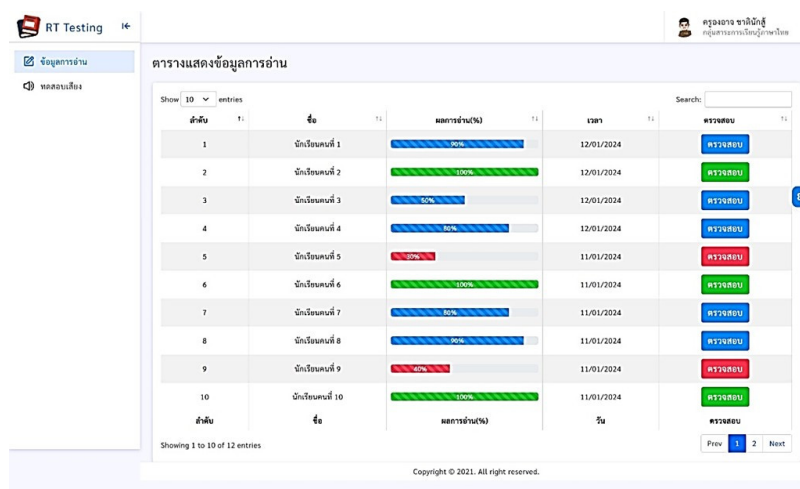


Figure 9. The display is in a table format for checking each student.

การวิจัยครั้งนี้เป็นการพัฒนาแอปพลิเคชันด้วยเทคโนโลยีรู้จำเสียงเพื่อจำแนกประเภทผลการอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่านของนักเรียนระดับประถมศึกษาปีที่ 1 ซึ่งผู้วิจัยได้ใช้เทคนิคต่าง ๆ ในการวิเคราะห์ชุดสัญญาณเสียงพร้อมรูปแบบการใช้งานแอปพลิเคชันซึ่งสามารถนำไปประยุกต์ใช้กับการเรียนการสอนภาษาไทยได้เพื่อช่วยในการแก้ปัญหาเรื่องการอ่านออกเสียงของผู้เรียนและยังเป็นแนวทางในการนำไปพัฒนาระบบหรือ เทคโนโลยีที่ช่วยในการเรียนการสอนภาษาต่อไปในอนาคต ซึ่งผลการใช้แอปพลิเคชันช่วยในการวิเคราะห์การอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่านของนักเรียนระดับชั้นประถมศึกษาปีที่ 1 แสดงใน Table 2.

Table 1. Accuracy values of using applications to Thai phonetic recognition according to the reading test of grade 1 students.

Words	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8
ดูแล (look after)	0.9823	0.9941	1.0000	0.9632	1.0000	0.8547	0.9986	1.0000
หน้าจอ (screen)	0.2365	0.8884	0.9658	0.6598	0.7231	1.0000	0.8941	0.8632
แนะแนว (guide)	0.5695	0.6897	1.0000	0.8635	1.0000	0.8549	1.0000	0.8996
ลูกชิ้น (meatball)	0.8954	0.2365	1.0000	0.3154	0.9697	0.9691	0.8613	1.0000
วิ่งแข่ง (race)	1.0000	0.7356	0.6985	1.0000	0.9587	1.0000	1.0000	0.6987
กับข้าว (side dish)	0.9632	1.0000	1.0000	0.9851	1.0000	0.3654	1.0000	0.9874
อาบน้ำ (shower)	0.6963	0.8774	1.0000	0.8753	0.7541	1.0000	0.6588	1.0000
น่ารัก (cute)	1.0000	0.8759	1.0000	0.9631	1.0000	0.8871	1.0000	0.8741
วิตามิน (vitamin)	0.2564	1.0000	1.0000	0.8565	1.0000	0.8521	0.9674	1.0000
นามสกุล (surname)	1.0000	0.6547	1.0000	0.8743	0.6987	1.0000	0.3654	1.0000
น้ำแข็ง (ice)	0.8999	1.0000	0.8111	0.7569	1.0000	0.9874	1.0000	0.6587
รถรับส่ง (shuttle)	1.0000	0.9151	0.7798	1.0000	0.6325	1.0000	0.9995	0.8965
เครื่องบิน (airplane)	0.2569	0.5698	1.0000	0.6581	0.9872	0.8321	0.5236	1.0000
สวยงาม (beautiful)	0.6981	1.0000	0.6223	0.6112	1.0000	0.9332	1.0000	0.9254
ศาสนา (religion)	1.0000	0.3256	0.9995	1.0000	0.9952	0.8569	1.0000	1.0000
พิเศษ (special)	0.8741	1.0000	0.5632	1.0000	1.0000	0.8974	1.0000	0.5244
ร้องเพลง (sing)	0.6954	1.0000	0.3652	0.8964	1.0000	0.8478	1.0000	0.2568
สร้างบ้าน (build a house)	0.8885	0.6987	1.0000	0.6922	0.8574	0.5548	0.8832	1.0000
ขยัน (hard-working)	1.0000	0.5893	0.6987	0.6972	1.0000	0.5236	1.0000	0.9687
ตลอด (always)	1.0000	0.6782	1.0000	0.7825	1.0000	0.3657	1.0000	1.0000

จาก Table 1. พบว่า ผลการใช้แอปพลิเคชันช่วยในการวิเคราะห์การอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่าน ของนักเรียนระดับชั้นประถมศึกษาปีที่ 1 ผู้อ่านจำนวน 8 คน อ่านจำนวนคำทั้งหมด 160 คำ มีความแม่นยำมากกว่า 50% อยู่ที่ 150 คำ คิดเป็น 93.75% และต่ำกว่า 50% อยู่ที่ 10 คำ คิดเป็น 6.25% หลังใช้แอปพลิเคชันช่วยตรวจจับการอ่านออกเสียง พบว่าส่วนมากมีค่าความถูกต้องอยู่ในระดับดีมาก

7. สรุปผลการวิจัย (Conclusion)

การศึกษาผลการใช้แอปพลิเคชันด้วยการรู้จำเสียงเพื่อจำแนกเสียงภาษาไทยตามแบบทดสอบการอ่านของนักเรียนระดับชั้นประถมศึกษาปีที่ 1 ผู้วิจัยได้ดำเนินการพัฒนาแอปพลิเคชันด้วยการใช้เทคนิคการรู้จำเสียง ได้แก่ เทคนิค Spectrogram 2) เทคนิค Mel Frequency Cepstral Coefficient (MFCC) นำมาเพื่อพัฒนาแอปพลิเคชันมาทดสอบกับข้อมูลของตัวอย่าง โดยดำเนินการเป็นระยะดังต่อไปนี้ ระยะที่ 1) การเลือกค่าและการเก็บข้อมูลคำจากกลุ่มเป้าหมายนักเรียนระดับประถมศึกษาแล้วดำเนินการจัดระเบียบไฟล์เพื่อนำไปใช้ในการวิเคราะห์คลื่นเสียงกับเสียงต้นฉบับ ระยะที่ 2) ขั้นตอนการพัฒนาแอปพลิเคชัน และระยะที่ 3) ขั้นตอนทดสอบโดยการทดลองใช้กับข้อมูลของกลุ่มเป้าหมาย โดยผลการวิจัย

พบว่า ผลการใช้แอปพลิเคชันช่วยในการวิเคราะห์การอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่าน ของนักเรียนระดับชั้นประถมศึกษาปีที่ 1 ผู้อ่านจำนวน 8 คน อ่านจำนวนคำทั้งหมด 160 คำ มีค่าความแม่นยำมากกว่า 50% อยู่ที่ 150 คำ คิดเป็น 93.75% และต่ำกว่า 50% อยู่ที่ 10 คำ คิดเป็น 6.25% หลังใช้แอปพลิเคชันช่วยตรวจจับการอ่านออกเสียง พบว่าส่วนมากมีค่าความถูกต้องอยู่ในระดับที่ดีมาก

8. อภิปรายผลการวิจัย (Discussion)

จากผลการวิจัยที่เกิดขึ้น พบว่า สามารถนำแอปพลิเคชันที่ใช้เทคนิคการรู้จำเสียงในการพัฒนาไปใช้ในการส่งเสริมและพัฒนาทักษะการออกเสียงภาษาไทยได้ ซึ่ง เทคโนโลยีการจดจำคำได้กลายเป็นเครื่องมือที่เปลี่ยนแปลงในการเรียนรู้ภาษา โดยเฉพาะอย่างยิ่งในการพัฒนาและส่งเสริมทักษะการออกเสียงโดยระบบจะให้ข้อเสนอแนะเฉพาะในทันทีซึ่งมีความสำคัญสำหรับผู้เรียนในการปรับปรุงความแม่นยำในการออกเสียง มีงานวิจัยและการศึกษาที่แสดงให้เห็นว่าการใช้การรู้จำเสียงอัตโนมัติ (Automatic Speech Recognition: ASR) สามารถปรับปรุงการออกเสียงของผู้เรียนได้อย่างมีนัยสำคัญผ่านระบบอินเทอร์เน็ตแบบโต้ตอบที่ให้มีการฝึกฝนโดยไม่ต้องอาศัยการดูแลของครูโดยตรง (Boyi & Guangliang, 2024); Sharma et al., 2024) ตัวอย่างเช่น การวิจัยระบุว่าระบบที่ใช้ ASR เช่น I Love Indonesia และ NovoLearning ได้นำไปใช้ในการฝึกการออกเสียงทั้งระดับคำและระดับประโยคในกลุ่มนักเรียน (Bashori et al., 2024) นอกจากนี้การใช้แอปพลิเคชันผ่านโทรศัพท์มือถือในระบบรู้จำคำพูดได้ปรับปรุงและเพิ่มประสิทธิภาพการเรียนรู้ภาษาทำให้ผู้เรียนได้รับข้อเสนอแนะที่นำเชื่อถือมากขึ้น (Jingning, 2024) โดยรวมแล้วการประยุกต์ใช้ ASR ในการศึกษาภาษาไม่เพียงแต่ช่วยให้ผลลัพธ์การเรียนรู้ที่ดีขึ้นแต่ยังเพิ่มแรงจูงใจและการมีส่วนร่วมของผู้เรียน (Samah, 2024)

เทคโนโลยีการรู้จำเสียงช่วยปรับปรุงคุณภาพการออกเสียงภาษาอังกฤษในหมู่ผู้เรียนอย่างมีนัยสำคัญด้วยการให้ข้อเสนอแนะทันทีที่ผู้เรียนสามารถระบุและแก้ไขข้อผิดพลาดในการออกเสียงได้อย่างมีประสิทธิภาพมากขึ้น นำไปสู่ทักษะการสื่อสารที่ดีขึ้น การใช้เทคโนโลยีการรู้จำเสียงในการเรียนรู้ภาษาจะช่วยสร้างประสบการณ์เชิงโต้ตอบและความสนใจสำหรับผู้เรียน เทคโนโลยีนี้ช่วยให้ผู้เรียนมีเส้นทางการเรียนรู้ส่วนบุคคลเป็นการตอบสนองความต้องการของแต่ละบุคคลและการเรียนรู้ซึ่งสามารถเพิ่มแรงจูงใจและโรงเรียนควรพิจารณาใช้เทคโนโลยีเทคโนโลยีการรู้จำเสียงเข้ากับหลักสูตรการเรียนรู้ภาษา ซึ่งอาจเกี่ยวข้องกับการใช้ซอฟต์แวร์หรือแอปพลิเคชันที่ใช้การจดจำคำพูดเพื่อให้ผู้เรียนมีโอกาสฝึกฝนและรับการตอบสนองข้อเสนอแนะแบบเรียลไทม์ ซึ่งอาจสรุปได้ว่า เทคโนโลยีเทคโนโลยีการรู้จำเสียงเป็นเครื่องมือที่มีประสิทธิภาพในการปรับปรุงการเรียนรู้การออกเสียงภาษาที่ช่วยเพิ่มประสบการณ์การเรียนรู้ให้ข้อเสนอแนะทันทีและเป็นสิ่งจำเป็นสำหรับการสื่อสารที่มีประสิทธิภาพซึ่งควรสนับสนุนการนำไปใช้ในวงกว้างในสภาพแวดล้อมการศึกษาและสนับสนุนการวิจัยในอนาคตเพื่อเพิ่มผลประโยชน์ที่อาจเกิดขึ้น (Boyi & Guangliang, 2024)

ทั้งนี้ ผลการวิจัยสามารถนำเทคโนโลยีการรู้จำเสียงไปใช้ในการเพิ่มประสบการณ์การเรียนรู้ให้กับผู้เรียนได้ โดยสอดคล้องกับงานวิจัยของ Sharma et al. (2024) ที่กล่าวว่า ผู้เรียนจะมีประสบการณ์การเรียนรู้แบบโต้ตอบ มีส่วนร่วมมากขึ้นโดยเทคโนโลยีนี้และทำให้ผู้เรียนสามารถฝึกการออกเสียงจากระยะไกลซึ่งมีประโยชน์โดยเฉพาะอย่างยิ่งในสถานการณ์ที่การโต้ตอบโดยตรงของครูมีจำกัด เครื่องมือนี้ใช้วิธี “ฟังและพูดซ้ำ” ซึ่งมีประสิทธิภาพสำหรับการเรียนรู้ภาษาวิธีนี้สนับสนุนให้ผู้เรียนเลียนแบบการออกเสียงคำและวลีช่วยให้การเก็บรักษาและเข้าใจภาษาได้ดีขึ้น ข้อดีที่สำคัญอย่างหนึ่งของการใช้เทคโนโลยีการรู้จำเสียงคือการให้ข้อเสนอแนะทันทีที่ผู้เรียนสามารถได้ยินเสียงของพวกเขาเมื่อเทียบกับแบบจำลองของครู (ระบบ) ทำให้พวกเขาสามารถปรับเปลี่ยนได้แบบเรียลไทม์และข้อเสนอแนะนี้มีความสำคัญต่อการเรียนรู้การออกเสียงเนื่องจากช่วยให้ผู้เรียนระบุและแก้ไขข้อผิดพลาดของพวกเขาทันที เทคโนโลยีการรู้จำเสียงมีลักษณะทั้งแบบออฟไลน์และออนไลน์ที่ทำให้การเรียนรู้ภาษาเข้าถึงได้มากขึ้นผู้เรียนสามารถฝึกได้ตามจังหวะและความสะดวกสบายของตนเอง ซึ่งเป็นสิ่งสำคัญอย่างยิ่งสำหรับผู้ที่ไม่สามารถเข้าถึงครูภาษาหรือห้องเรียนเป็นประจำ ผลการวิจัยชี้ให้เห็นว่าเทคโนโลยีดังกล่าว

สามารถนำไปใช้กับบริบทการเรียนรู้ภาษาต่าง ๆ ซึ่งอาจเปลี่ยนวิธีการสอนการออกเสียงและทักษะทางภาษาสิ่งนี้อาจนำไปสู่ประสบการณ์การเรียนรู้ส่วนบุคคลมากขึ้นซึ่งปรับให้เหมาะสมกับความต้องการของนักเรียนแต่ละคน

ผลการวิจัยยังสอดคล้องกับ Bashori et al. (2024) ที่ทำการศึกษายืนยันว่าระบบการเรียนรู้ภาษาที่ใช้การเรียนรู้ร่วมกับเทคโนโลยีการรู้จำเสียงมีประสิทธิภาพในการปรับปรุงการออกเสียงภาษา ข้อดีที่สำคัญอย่างหนึ่งคือความสามารถในการสะท้อนผลลัพท์ในทันทีและเป็นส่วนตัว คุณลักษณะนี้ช่วยให้ผู้เรียนเข้าใจข้อผิดพลาดในการออกเสียงแบบเรียลไทม์ ซึ่งสามารถปรับปรุงประสบการณ์การเรียนรู้ ซึ่งการศึกษาในอนาคตควรสำรวจผลกระทบระยะยาวของการใช้เทคโนโลยีการรู้จำเสียงต่อคุณภาพการออกเสียงของผู้เรียน โดยสามารถให้ข้อมูลเชิงลึกที่ลึกซึ้งยิ่งขึ้นเกี่ยวกับการใช้เทคโนโลยีว่าสามารถปรับการเรียนรู้การออกเสียงว่าเป็นอย่างไร นอกจากนี้งานวิจัยของ Jingning (2024) ยังได้เสนอแนะว่าการใช้อัลกอริทึมการจดจำคำเสียงช่วยเพิ่มประสิทธิภาพของการศึกษาภาษาอย่างมีนัยสำคัญ ซึ่งเป็นระบบช่วยการเรียนรู้อัจฉริยะที่พัฒนาขึ้นที่มีความแม่นยำและเชื่อถือมากขึ้นสำหรับผู้เรียนในการฝึกฝนทักษะการพูดภาษาด้วยการใช้ประโยชน์จากโหนดเช่นเซอร์หลายโหนดเพื่อรับสัญญาณการพูด ระบบช่วยนี้มั่นใจได้ว่าผู้เรียนจะได้รับข้อเสนอแนะในทันทีและแม่นยำซึ่งมีความสำคัญต่อการเรียนรู้ภาษา การวิจัยเน้นการประยุกต์ใช้อัลกอริทึมการเรียนรู้ของเครื่องที่สำหรับการประมวลผลล่วงหน้าและการแยกคุณสมบัติของสัญญาณพูดแนวทางเทคโนโลยีนี้เป็นกุญแจสำคัญในการบรรลุความแม่นยำในการรับรู้ที่สูงขึ้นเมื่อเทียบกับวิธีการแบบดั้งเดิม ซึ่งแสดงให้เห็นถึงศักยภาพของอัลกอริทึมขั้นสูงในเทคโนโลยีการศึกษา ผลการวิจัยชี้ให้เห็นว่าการพัฒนาและการนำระบบอัจฉริยะไปใช้ต่อไปอาจปฏิวัติการเรียนรู้ภาษาด้วยการจัดหาเครื่องมือของผู้เรียนที่ปรับให้เข้ากับความต้องการและสภาพแวดล้อมของแต่ละบุคคล ศักยภาพในการศึกษาส่วนบุคคลจะเพิ่มขึ้นอย่างมีนัยสำคัญ

9. ข้อเสนอแนะงานวิจัย (Recommendation)

1) ข้อเสนอแนะทั่วไป การใช้แอปพลิเคชันช่วยในการวิเคราะห์การอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่านของนักเรียนระดับชั้นประถมศึกษาปีที่ 1 เป็นการวิจัยเกี่ยวกับการเทียบความคล้ายของคลื่นเสียงกับเสียงของต้นฉบับ โดยใช้โมเดลในการเทรน และทดสอบ ดังนั้น ควรมีข้อมูลที่ใช้ในการเทรนเสียงยังมีมากเท่าไร ค่าความแม่นยำยิ่งมากขึ้น ควรศึกษาเกี่ยวกับข้อมูลทฤษฎีต่าง ๆ และศึกษาเทคนิคที่หลากหลาย เพื่อเลือกใช้เทคนิคที่เหมาะสมกับการนำมาใช้ในการพัฒนาระบบให้มีความคล้ายของสัญญาณเสียงที่สูงขึ้นแอปพลิเคชันช่วยในการวิเคราะห์การอ่านออกเสียงภาษาไทย ควรได้รับการพัฒนาให้มีความหลากหลายและเหมาะสมกับระดับชั้นของนักเรียนมากยิ่งขึ้น นอกจากนี้ ควรมีการวิจัยเพื่อศึกษาผลการใช้แอปพลิเคชันช่วยในการวิเคราะห์การอ่านออกเสียงภาษาไทยตามแบบทดสอบการอ่านคำที่มักอ่านผิดในภาษาไทยกับนักเรียนระดับชั้นอื่น ๆ ด้วย

2) ข้อเสนอแนะในการวิจัยครั้งต่อไป หากนำเทคนิคการวิจัยนี้ไปประยุกต์ใช้ในงานด้านอื่น ควรคำนึงถึงความแม่นยำและความเหมาะสมของเทคนิคที่ใช้วิเคราะห์เนื่องจากเทคนิคการประมวลผลสัญญาณเสียงมีหลากหลายรูปแบบ แต่ละเทคนิคมีข้อดีข้อเสียแตกต่างกันไป ดังนั้น ควรศึกษาข้อมูลเกี่ยวกับเทคนิคแต่ละประเภทอย่างละเอียด เพื่อให้สามารถเลือกเทคนิคที่เหมาะสมกับงานวิจัยที่ต้องการพัฒนา

10. เอกสารอ้างอิง (References)

Angsuphan, A. (2019). *Speech Consistency Test for Cerebral Palsy Children with Dysarthric*. [Bachelor Thesis, Chulalongkorn University]. Chulalongkorn University Intellectual Repository.
<http://cuir.car.chula.ac.th/handle/123456789/78486>. (In Thai)

- Bashori, M., Van Hout, R., Strik, H., & Cucchiari, C. (2024). I Can Speak : Improving English Pronunciation through Automatic Speech Recognition-Based Language Learning Systems. *Innovation in Language Learning and Teaching*, 1–19. <https://doi.org/10.1080/17501229.2024.2315101>.
- Boyi, H., & Guangliang, L. (2024, January 29-30). Analysis of English Speech Learning Quality based on Speech Recognition Technology. *2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC)*, 1–4. <https://doi.org/10.1109/ICOCWC60930.2024.10470570>.
- Charoendee, M. (2016). *Feature Selection and Extraction for Thai Emotional Speech Classification*. [Master dissertation, Chulalongkorn University]. Chulalongkorn University Intellectual Repository. <http://cuir.car.chula.ac.th/handle/123456789/54912>. (In Thai)
- Jingning, L. (2024). Speech Recognition Based on Mobile Sensor Networks Application in English Education Intelligent Assisted Learning System. *Measurement: Sensors*, 32, 101084. <https://doi.org/10.1016/j.measen.2024.101084>.
- Klangbungrong, K. (2012). *Web Application to Assist Mothers in Monitoring Health and Recording Child Development*. [Unpublished Bachelor Thesis]. King Mongkut's University of Technology Thonburi. (In Thai)
- Malangpoo, P., Philuek, W., & Pomsamrit, N. (2022). Using Speech Recognition System for Enhancing Chinese Pronunciation. *International Journal of Mechanical Engineering*, 7(1), 3442-3451.
- NovaBizz. (n.d.). *Biometrics*. <https://www.novabizz.com/CDC/System/Biometrics.htm>. (In Thai)
- Philuek, W., & Puttasem, D. (2023). Teaching Thai Language Literacy: Proposed of Using Speech Recognition Technology Techniques to Detect Read Aloud in Thai Tonal Conjugation for Primary Education Students. *Shanlax International Journal of Education*, 11(4), 77–84. <https://doi.org/10.34293/education.v11i4.5935>.
- Pinno, K., Khahakitkoson, A., Attanatwong, A., & Chaisangjan, A. (2011). *Language and Communication*. (2nd ed.). Chen Printing. (In Thai)
- Pitaksirianant, N. (2011). *Evaluation of Automatic Speech Intelligibility in Cleft Lip and Palate Patients*. [Master dissertation, Chulalongkorn University]. Chulalongkorn University Intellectual Repository. <http://cuir.car.chula.ac.th/handle/123456789/52521>. (In Thai)
- Rabiner, L., & Juang, B. H. (1993). *Fundamentals of Speech Recognition*. Englewood Cliffs.
- Rockkittichareon, W. (2011). *Acoustic Parameters for Manner of Articulation Classification in Thai Continuous Speech*. [Master dissertation, Chulalongkorn University]. Chulalongkorn University Intellectual Repository. <http://cuir.car.chula.ac.th/handle/123456789/31730>. (In Thai)
- Samah, N. (2024). L'Impact de L'Utilisation de La Reconnaissance Vocale Dans L'Enseignement du FLE à L'Oral Dans Les Classes du Secondaire en Algérie. *International Journal of Early Childhood Special Education*, 16(2), 80–88. <https://doi.org/10.48047/intjecse/v16i2.9>.
- Sharma, R., Menon, A., Khairnar, S., Bhagat, S., & Rawat, S. (2024). Pronunciation Learning Using Automatic Speech Recognition. *EPRA International Journal of Multidisciplinary Research*, 10(4), 565–568. <https://doi.org/10.36713/epra16634>.