บทที่ 1

บทนำ

1.1 ปัญหาและที่มาของการศึกษา

การสื่อสารคือทักษะที่สำคัญของมนุษย์ ซึ่งการพูดเป็นการสื่อสารรูปแบบหนึ่งที่มนุษย์มีความจำเป็นที่ต้องใช้อย่างมากในชีวิตประจำวัน เพื่อแสดงถึงความต้องการและใช้ในการติดต่อสื่อสาร โดยการพูดจำเป็นต้องสื่อออกมาทางภาษา ภาษาที่มนุษย์ใช้ก็จะแตกต่างกันไปในแต่ละท้องถิ่น การที่มนุษย์ในแต่ละท้องถิ่นสามารถพูดภาษาท้องถิ่นเพื่อใช้ติดต่อสื่อสารและแสดงความต้องการนั้นจำเป็นต้องผ่านการเรียนรู้ ฝึกฝน จนเกิดความคุ้นเคย

แต่ในเด็กทารกแรกเกิดนั้นยังไม่ได้รับการฝึกฝนหรือเรียนรู้เกี่ยวกับการใช้ภาษาในการคิดค่อสื่อสาร รวมถึงการพัฒนาการของร่างกายในด้านการออกเสียงยังไม่พร้อมที่จะสื่อสารออกมาเป็นคำพูดที่คนทั่วไปเข้าใจได้ เสียงที่เด็กทารกสามารถสื่อสารออกมาได้จึงเป็นเพียงเสียงร้องที่คนทั่วไปไม่เข้าใจ Priscilla Dunstan [1] ผู้เชี่ยวชาญในด้านการดูแลเด็กทารก ได้ทำการศึกษาค้นคว้าเสียงร้องของเด็กทารกและทำการจำแนกเสียงร้องตามความต้องการของเด็กทารกไว้ เสียงเด็กทารกที่ร้องออกมาเพื่อแสดงความต้องการนั้น จำเป็นต้องอาศัยความคุ้นเคยและมีความสามารถในการฟังที่ดีเพื่อจะได้เข้าใจว่าเด็กทารกมีความต้องการอย่างไร สำหรับผู้ที่ไม่มีความคุ้นเคยหรือพ่อแม่มือใหม่นั้นอาจไม่สามารถแยกแยะเสียงร้องของเด็กทารกได้ว่ามีความต้องการอะไร

วิทยานิพนธ์นี้จึงมีความสนใจในการพัฒนาอัลกอริทึม (Algorithm) เพื่อใช้ในการรู้จำเสียง (Speech Recognition) และการจำแนกเสียงร้อง (Speech Classification) โดยใช้เสียงตัวอย่างจากผู้เชี่ยวชาญซึ่งได้มีการระบุความหมายไว้แล้ว เพื่อเป็นข้อมูลในการพัฒนาอัลกอริทึม งานวิจัยด้านนี้มักเป็นงานวิจัยที่ได้รับความนิยมในการทำวิจัย งานวิจัยที่ทำขึ้นจึงสามารถนำไปใช้ประโยชน์ได้ในหลายๆด้านเช่น การระบุตัวตนด้วยเสียงเพื่อใช้ในงานด้านความปลอดภัย การใช้เสียงพูดสั่งอุปกรณ์อิเล็กทรอนิกส์ เป็นต้น

ในการทำวิทยานิพนธ์นี้มีแนวคิดในการประยุกต์ใช้ความรู้ด้านการทำเหมืองข้อมูล (Data mining) [2] ใช้ในการรู้จำและจำแนกเสียง ในการทำเหมืองข้อมูลนั้นมักถูกนำไปใช้ในงานสามส่วน คือ งานด้านการหาความสัมพันธ์ (Relation) เป็นการหาว่าข้อมูลหนึ่งๆ มีความเกี่ยวข้องกับข้อมูลอื่นๆ มากน้อยเพียงใดโดยอาศัยกฎความสัมพันธ์ (Association Rule) หรือทฤษฎีความน่าจะเป็นแบบเบย์ (Bayes Theorem) ส่วนที่สองเป็นงานด้านการแบ่งกลุ่ม (Clustering) เป็นการนำข้อมูลที่มีอยู่มาพิจารณาว่าควรแบ่งออกเป็นทั้งหมดกี่กลุ่ม และงานด้านการจำแนก (Classification) คือการนำข้อมูลมาระบุว่าข้อมูลนั้นๆอยู่กลุ่มใด โดยการแบ่งกลุ่มและการจำแนกนั้นสามารถใช้อัลกอริทึมต่างๆเช่น โครงข่ายประสาทเทียม (Neural Network) ฟัซซีโลจิก (Fuzzy Logic) ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm : GA) ความฉลาดแบบกลุ่ม (Swarm Intelligence) มาช่วยในการแก้ปัญหา

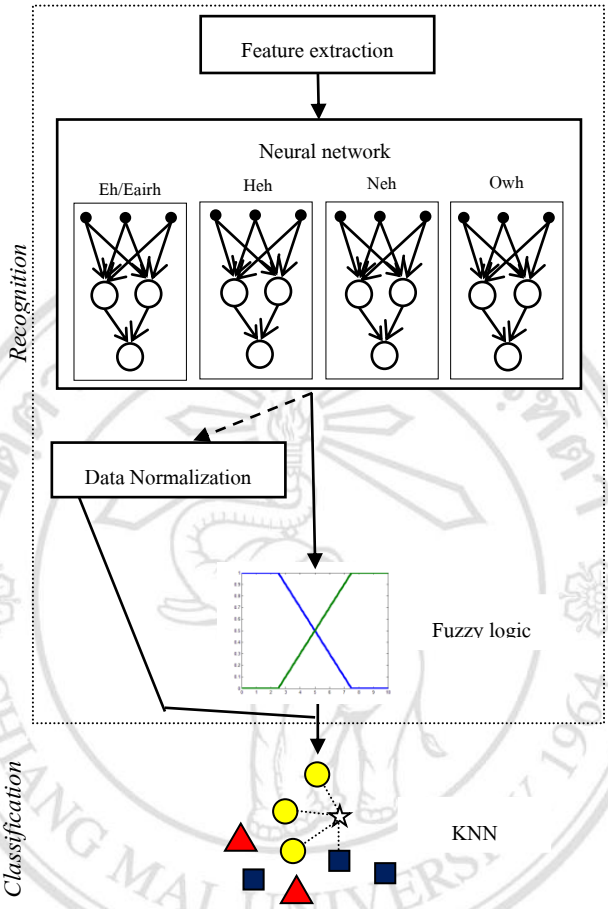
เสียงร้องของเด็กทารกที่พูดออกมาคำแรกนั้นตามที่ Priscilla Dunstan ผู้เชี่ยวชาญด้านการดูแลเด็ก ได้จำแนกไว้มีอยู่ 5 คำคือ Eh, Eairh Heh Neh และ Owh ที่สามารถสื่อความหมายได้แต่ละคำ ยากที่จะฟังเข้าใจว่าเป็นคำว่าอะไรอีกทั้งยังไม่มีความหมายทางภาษาพูด ดังนั้นวิทยานิพนธ์นี้จึงจะเป็นการหาอัลกอริทึมที่มีประสิทธิภาพในสกัดค่าคุณลักษณะ (Feature Extraction) ร่วมกับการนำโครงข่ายประสาทเทียมและตรรกศาสตร์คลุมเครือที่มีความง่ายในการเรียนรู้มาประยุกต์เพื่อจำแนกเสียงโดยอาศัยการทดลองต่างๆ เพื่อพัฒนาโครงสร้างให้เหมาะสม สามารถช่วยให้คนดูแลหรือพ่อแม่ของเด็กทารกสามารถเข้าใจได้ว่า เบื้องต้นเด็กมีความต้องการอะไรและดูแลอย่างเหมาะสมต่อไป

1.2 แนวทางการแก้ปัญหา

ในการทำงานของอัลกอริทึมเพื่อใช้แก้ปัญหานั้นแบ่งออกเป็น 2 ขั้นตอนคือการรู้จำและการจำแนกเสียงโดยเป็นการนำอัลกอริทึม [3] มาแก้ไขปรับปรุงเพื่อให้มีความเหมาะสมกับการใช้งานด้านเสียงร้องของเด็กทารก

ขั้นตอนการรู้จำ จากรูปที่ 1.1 จะเห็นได้ว่าส่วนของการรู้จำประกอบไปด้วยส่วนย่อย 4 ส่วน โดยมีขั้นตอนการทำงานคือ เมื่อเสียงผ่านกระบวนการการสกัดค่าคุณลักษณะจะได้ชุดข้อมูลตัวเลขออกมา จากนั้นทำการแบ่งข้อมูลออกเป็น 2 ส่วนคือข้อมูลสำหรับใช้สอนระบบ (Training Data) และข้อมูลสำหรับทดสอบประสิทธิภาพของระบบ (Testing Data) โดยการแบ่งข้อมูลนั้นจะทำการตรวจสอบความสมเหตุสมผลแบบไขว้ (Cross Validation) จากนั้นนำข้อมูลสำหรับใช้สอนระบบมาผ่านโครงข่ายประสาทเทียมโดยทำการสอนจนโครงข่ายประสาทเทียมมีความเสถียร หลังจากนั้นทำการเก็บค่าถ่วงน้ำหนัก (Weight) ไว้เพื่อนำไปใช้สำหรับข้อมูลทดสอบ ส่วนค่าเอาต์พุต (Output) ที่ได้จากโครงข่ายประสาทเทียมถูกนำไปใช้ในขั้นตอนต่อไปโดยกำหนดให้โครงข่ายประสาทเทียม 1

โครงข่ายแทนเสียง 1 เสียง ดังนั้นในงานวิจัยนี้จะมีโครงข่ายประสาทเทียมทั้งสิ้น 4 โครงข่ายโดยกำหนดให้โครงข่ายประสาทเทียม 1 โครงข่ายแทนเสียง 1 เสียง ดังนั้นในงานวิจัยนี้จะมีโครงข่ายประสาทเทียมทั้งสิ้น 4 โครงข่ายดังรูปที่ 1.1



รูปที่ 1.1 ภาพรวมของอัลกอริทึมที่ใช้

ขั้นตอนต่อไปนำค่าเอาต์พุตจากโครงข่ายประสาทเทียมไปผ่านกระบวนการนอร์มัลไลเซชันข้อมูล (Data Normalization) เพื่อให้ได้ข้อมูลชุดใหม่ออกมา แล้วนำข้อมูลชุดใหม่ไปใช้ฟัซซีโลจิกโดยใช้วิธีของแมมดานิ (Mamdani) จนได้เอาต์พุตออกมาเพื่อนำไปใช้ในขั้นตอนการจำแนก

ขั้นตอนการจำแนกข้อมูลนั้นจะใช้เคเนียเรสเนเบอร์ (K-nearest Neighbor : KNN) ซึ่งจะใช้ข้อมูลจากเอาต์พุตของฟัซซีโลจิก โดยนำข้อมูลสำหรับใช้ทดสอบมาผ่านโครงสร้างที่สร้างไว้ในข้างต้นแล้วนำมาทดสอบเพื่อหาความแม่นยำในการจำแนก

1.3 สรุปสาระสำคัญจากเอกสารที่เกี่ยวข้อง

งานในด้านการรู้จำเสียงและการจำแนกเสียงนั้นเป็นที่นิยมในการทำวิจัยของนักวิจัยจำนวนมาก ซึ่งมีวัตถุประสงค์ที่แตกต่างกัน แต่จะเป็นการสร้างอัลกอริทึมขึ้นมาเพื่อใช้ระบุว่าเสียงนั้นๆ หมายถึงจากนั้นนำผลลัพธ์ที่ได้ไปต่อยอดตามวัตถุประสงค์ของงานวิจัยนั้นๆ

K. Saetern และ N. Eiamkanitchat [3] นำเสนอผลงานวิจัยเรื่อง "An Ensemble K-Nearest Neighbor with Neuro-fuzzy Method for Classification " งานวิจัยนี้เป็นการปรับปรุงอัลกอริทึมนิวโรฟัซซี โดยใช้ K-NN เพื่อเพิ่มประสิทธิภาพในการจำแนกให้มากขึ้น โดยใช้ชุดข้อมูลจาก UCI Machine Learning Repository [4] ในการทดสอบ โครงสร้างของอัลกอริทึมนี้แบ่งเป็นสองส่วนคือ ส่วนการแลปงชุดข้อมูลเดิมเป็นชุดข้อมูลใหม่ (Transformed dataset) และส่วนการจำแนกโดยใช้ KNN ส่วนแรกนั้นแบ่งออกเป็น 2 ส่วนย่อยคือการใช้โครงข่ายประสาทเทียมแบบ MLP ในการหาค่าผิดพลาด และการนำค่าผิดพลดาที่ได้จากส่วนย่อยแรกไปใช้ฟัซซีโลจิกโดยวิธีของ Mandani ในส่วนของการจำแนกนั้นจะนำค่าที่ได้จากวิธีของ Mandani ไปใช้ K-NN ในการจำแนก ผลการทดลองพบว่าอัลกอริทึมที่นำเสนอมีความแม่นยำในการจำแนกสูงเมื่อเทียมกับอัลกอริทึมที่ได้รับความนิยมอื่นๆ

R. Hidayati และคณะ [5] นำเสนองานวิจัยเรื่อง "The extraction of acoustic features of infant cry for emotion detection based on pitch and formants" ซึ่งเป็นงานวิจัยเกี่ยวกับการระบุความหมายเสียงร้องของเด็กทารกที่มีอายุตั้งแต่ 1 วันถึง 6 เดือน ซึ่งเสียงร้องของเด็กนั้นผู้เขียนได้ทำการเก็บรวบรวมเสียงจากโรงพยาบาลและคลีนิก ทำการแบ่งเสียงร้องของเด็กออกเป็น 5 รูปแบบคือ เจ็บปวด เสียใจ หิว กลัว และโกรธ งานวิจัยนี้จะแบ่งออกเป็น 2 ส่วนคือการสร้างโมเดลเสียงและส่วนการตัดสินใจว่าเสียงนั้นมีความหมายว่าอะไร โดยส่วนของโมเดลเสียงนั้นมีขั้นตอนการทำงานแบ่งออกเป็น 2 ส่วนย่อยคือ ส่วนของการตรวจหาพิทซ์ (Pitch Detection) ใช้อัลกอริทึมคือการคูณค่าฮาร์โมนิคในสเปกตรัม (Harmonic Product Spectrum: HPS) และส่วน Format Detection ใช้อัลกอริทึมการประมาณพันธะเชิงเส้น (Linear predictive coding: LPC) ซึ่งทั้งสองส่วนย่อยนี้จะทำงานไปพร้อมๆกัน ส่วนการตัดสินใจนั้นใช้อัลกอริทึมเคมีน (K-means algorithm) ระยะทางแบบยุคลิด (Euclidian distance) ในการระบุความหมายของเสียงร้อง ผลจากการวิจัยพบว่า เสียงร้องที่ที่มีความหมายว่าเจ็บปวดให้ค่าความแม่นยำสูงสุดที่ 90 % ส่วนเสียงร้องที่แทนความหมายว่าโกรธให้ค่าความแม่นยำต่ำสุดที่ 58 % และมีค่าเฉลี่ยของทุกเสียงอยู่ที่ 75 %

A. Zabidi และคณะ [6] นำเสนอผลงานวิจัยเรื่อง "Classification of Infant Cries with Asphyxia Using Multilayer Perceptron Neural Network" ซึ่งเป็นงานวิจัยที่นำเสนอการพิจารณาเสียง

ร้องของเด็กทารกเป็นโรค Asphyxia หรือไม่ เนื่องจากโรคนี้จะส่งผลกับเสียงร้องของเด็ก ข้อมูลเสียงร้องของเด็กทารกที่ใช้นั้นนำมาจาก University of Milano–Bicocca ในขั้นตอนการทำงานนั้นแบ่งออกเป็น 2 ส่วนคือการสกัดค่าคุณลักษณะและการจำแนก การสกัดค่าคุณลักษณะของเสียงนั้นใช้อัลกอริทึม MFCC ร่วมกับวิธีกำลังสองน้อยที่สุดแบบตั้งฉาก (Orthogonal Least Square: OLS) เพราะว่าจากการทดลองใช้ OLS และไม่ใช้ OLS พบว่า การใช้ OLS ร่วมกับ MFCC ให้ความแม่นยำที่สูงกว่า นอกจากนั้นยังมีการทดลองในส่วนของ MFCC เพื่อหาจำนวนชุดตัวกรอง (Filter Bank) ที่เหมาะสมกับเสียงโดยใช้ชุดตัวกรอง จำนวน 30 40 50 และ 60 ส่วนการจำแนกนั้นได้ใช้โครงข่ายประสาทเทียมแบบ MLP ซึ่งได้มีการทดลองเพื่อหาโครงสร้าง MLP ที่เหมาะสมที่สุด โดยความแม่นยำสูงสุดของงานวิจัยนี้อยู่ที่ 94% โดยใช้ 40 ชุดตัวกรองได้จำนวนค่าคุณลักษณะจาก MFCC 12 ค่าคุณลักษณะและส่วนของโครงข่ายประสาทเทียมใช้ 15 โหนดซ่อน

M.S. Daud และคณะ [7] นำเสนอผลงานวิจัยเรื่อง “Investigation of MFCC Feature Representation for Classification of Spoken Letters using Multi-Layer Perceptrons ” งานวิจัยนี้นำเสนอการแยกเสียงพูดตัวอักษร ‘A’ และ ‘S’ งานวิจัยนี้นำข้อมูลเสียงมาจากคนจำนวน 12 คนพูด ‘A’ และ ‘S’ ตัวอักษรละ 3 ครั้ง การสกัดค่าคุณลักษณะใช้ MFCC และการจำแนกใช้โครงข่ายประสาทเทียมแบบ MLP ซึ่งมีการทดลองเพื่อหาโครงสร้าง MLP ที่ดีที่สุดสำหรับข้อมูลนี้จากการกำหนดโหนด (Nodes) ของชั้นซ่อน (Hidden Layer) ที่แตกต่างกัน คือ 2 4 6 8 10 12 และ 14 โหนด ผลการวิจัยพบว่าส่วนการเรียนรู้ของโครงสร้างมีความแม่นยำ100% และการทดสอบโครงสร้างนั้นมีความแม่นยำที่ 93% โดยใช้จำนวนโหนดของชั้นซ้อนจำนวน 2 โหนด เพราะว่าเมื่อเพิ่มจำนวนโหนดไปความแม่นยำยังเท่าเดิม การมีจำนวนโหนดมากทำให้เสียเวลามากขึ้นในการทำงาน ในงานวิจัยนี้สาเหตุที่ความแม่นยำสูงมากมาจากการแยกเสียงพูดตัวอักษร ‘A’ และ ‘S’ นั้น มีความแตกต่างกันของเสียงค่อนข้างชัดเจน ทำให้ง่ายต่อการจำแนก

B.B. Vachhani และ H.A. Patil [8] นำเสนอผลงานวิจัยเรื่อง “Use of PLP Cepstral Features for Phonetic Segmentation” งานวิจัยนี้จะเป็นการเปรียบเทียบอัลกอริทึมในส่วนการสกัดค่าคุณลักษณะกับงานด้านการแบ่งหมวดตามการออกเสียง (Phonetic Segmentation) ซึ่งงานด้านนี้นำไปใช้สำหรับแอปพลิเคชันเกี่ยวกับ Text-to-Speech (TTS) และการรู้จำเสียงพูดอัตโนมัติ (Automatic Speech Recognition: ASR) ในงานวิจัยนี้นำฐานข้อมูลที่ใช้วิจัยมาจาก TIMIT โดยอัลกอริทึมที่เลือกมาเปรียบเทียบคือ MFCC PLP และ RASTA ผลจากการทดลองพบว่า PLP สามารถทำงานได้ดีกว่าอัลกอริทึมอื่นๆ เพราะว่า PLP มีการใช้ Critical Band Masking Curves และ Equal Loudness Pre-emphasis ในการสกัดค่าคุณลักษณะ

R. Cohen และ Y. Lavner [9] นำเสนอผลงานวิจัยเรื่อง “Infant Cry Analysis and Detection” งานวิจัยนี้เป็นการสร้างอัลกอริทึมสำหรับแอปพลิเคชันในการตรวจว่าเด็กทารกร้องไห้อยู่หรือไม่ โดยนำข้อมูลมาจาก Gyorgy Varallyay ซึ่งส่วนการสกัดค่าคุณลักษณะใช้ MFCC และ ส่วนการจำแนกใช้ K-NN โดยมีขั้นตอนการทำงาน 3 ขั้นตอนหลักเพื่อตรวจสอบความถูกต้องคือ Voice Activity Detector (VAD) เพื่อตรวจสอบการร้องของเด็ก Classification เพื่อจำแนกว่าเด็กร้องไห้หรือไม่ และ Post-processing เพื่อตรวจสอบผลการจำแนกอีกครั้งและลดความผิดพลาดที่เกิดขึ้น ในการทดลองนั้นมีการตรวจสอบอัลกอริทึมกับสภาพแวดล้อมที่มีเสียงรบกวน (noise) ประเภทต่างๆด้วย ผลการทดลองพบว่าการทำงานให้ได้ประสิทธิภาพที่สูงนั้นต้องมี Signal-to-noise ratio (SNR) ของเสียงที่สูงจะทำให้มีความแม่นยำสูงเกือบถึง 100 %

Yu-Min Zeng และคณะ [10] นำเสนอผลงานวิจัยเรื่อง “Robust GMM Based Gender Classification using Pitch and RASTA-PLP Parameters of Speech” งานวิจัยนี้เป็นการวิจัยเพื่อแยกแยะเสียงของผู้ชายและผู้หญิง โดยมีการทดลองที่ควบคุมโดยปัจจัยรูปแบบต่างๆคือ เสียงปกติ เสียงมีเสียงรบกวนรูปแบบต่างๆ และภาษาที่แตกต่างกัน โดยส่วนการสกัดค่าคุณลักษณะนั้นใช้ RASTA ร่วมกับ Pitch เพราะ Pitch เป็นหนึ่งในปัจจัยสำคัญในการระบุเพศของผู้พูด ส่วนการจำแนกใช้เกาส์เซียนมิกซ์เจอร์โมเดล ในการทดลองนั้นได้ทดลองในสภาวะที่มี SNR แตกต่างกันรวมทั้งการทดลองที่ใช้เพียงแค่ RASTA หรือ Pitch เพียงอย่างเดียวในการสกัดค่าคุณลักษณะอีกด้วย ผลการทดลองพบว่าอัลกอริทึมที่สร้างขึ้นมีความแม่นยำในการจำแนกสูงมากถึง 98 % สำหรับเสียงปกติ และ 95 % สำหรับเสียงที่มีเสียงรบกวนซึ่งถือว่ามีความทนทานต่อเสียงรบกวน นอกจากนั้นยังสามารถใช้งานได้ถึง 4 ภาษาจากการทดลองคือภาษาอังกฤษ ภาษาเยอรมัน ภาษาญี่ปุ่น ภาษาอิตาลี

X. Zhang และคณะ [11] นำเสนอผลงานวิจัยเรื่อง “Aircraft Type Recognition Based on Shortwave Speech Communication with PLP” งานวิจัยนี้นำเสนอการแยกแยะประเภทของเครื่องบินจากเสียงโดยอาศัยการสกัดค่าคุณลักษณะด้วย RASTA-PLP-HOC และการจำแนกเสียงโดยใช้ SVM ผลการทดสอบพบว่า PLP มีความแม่นยำมากกว่า MFCC และ RASTA ปกติพอสมควร

1.4 วัตถุประสงค์ของการศึกษา

ประยุกต์ใช้นิวโรฟัซซีในการรู้จำและจำแนกเสียงร้องของเด็กทารกได้อย่างมีประสิทธิภาพ

1.5 ประโยชน์ที่ได้รับจากการศึกษาเชิงทฤษฎีและ/หรือเชิงประยุกต์

ได้อัลกอริทึมในการจำแนกเสียงร้องของเด็กทารกที่มีประสิทธิภาพ และสามารถนำอัลกอริทึมที่พัฒนาขึ้นไปประยุกต์ใช้ในงานวิจัยเกี่ยวกับการรู้จำเสียงที่เกี่ยวข้องได้

1.6 ขอบเขตการทำวิจัย

1.6.1 งานวิจัยนี้พัฒนาโดยใช้โปรแกรมแมตแล็บ (Matlab) ในการทำการทดลองต่างๆและใช้โปรแกรมไมโครซอฟท์เอกซ์เซล (Microsoft Excel) ในการจัดเก็บผลการทดลอง

1.6.2 เสียงเด็กทารกนั้นอ้างอิงจากการภาษาทารกของดันส์ตัน โดยเสียงที่นำมาใช้ทำการทดลองนั้นตัดมาจากคลิปเสียงที่มีการระบุความหมายของเสียงไว้เรียบร้อยแล้ว

1.6.3 ในการวัดประสิทธิภาพการจำแนกเสียงร้องของเด็กทารกนั้นใช้วิธีเปรียบเทียบผลการจำแนกกับอัลกอริทึมอื่นๆที่ได้รับความนิยมในการจำแนก

1.6.4 การประยุกต์ใช้กับเสียงร้องอื่นๆนั้น เสียงที่นำมาทดสอบประสิทธิภาพของอัลกอริทึมนั้นใช้จำนวนผู้ทดลองคือ ผู้ชาย 7 คน และ ผู้หญิง 7 คนโดยใช้จำนวนคำในการทดลอง 9 คำ โดยทดลองกับปัจจัยต่างๆ 3 ปัจจัยคือ จำนวนพยางค์ การออกเสียงรูปแบบต่างๆ และเสียงรบกวน

1.7 วิธีการทำวิจัย

1.7.1 ศึกษาเทคนิคการสกัดค่าคุณลักษณะออกมาจากเสียงด้วยเทคนิคต่างๆ

1.7.2 ศึกษาเทคนิคการจำแนกข้อมูลโดยการรวมกันของเทคนิคต่างๆคือโครงข่ายประสาทเทียม ฟัซซีโลจิก และเคเนียเรสเนเบอร์

1.7.3 ค้นหาข้อมูลเกี่ยวกับเสียงร้องของเด็กทารกและความหมาย

1.7.4 ทำการทดลองเพื่อสร้างอัลกอริทึมที่สามารจำแนกเสียงร้องของเด็กทารกได้อย่างมีประสิทธิภาพ โดยทดลองทีละส่วนย่อยๆ คือการหาเทคนิคสกัดค่าคุณลักษณะที่เหมาะสมที่สุด การสร้างโครงสร้างของโครงข่ายประสาทเทียม การเตรียมข้อมูลก่อนการประมวลผล การเลือกเทคนิคเคเนียเรสเนเบอร์

1.7.5 เปรียบเทียบผลการทดลองกับอัลกอริทึมที่ได้รับความนิยมอื่นๆ เพื่อสรุปผลการทดลอง

1.7.6 ออกแบบการทดลองเพื่อทดสอบอัลกอริทึมที่พัฒนาขึ้นว่าสามารถนำไปใช้จำแนกเสียงอื่นๆได้อย่างมีประสิทธิภาพ

1.7.7 เก็บรวบรวมเสียงที่ใช้ทดสอบประสิทธิภาพของอัลกอริทึม โดยมีการกำหนดปัจจัยในการอัดเสียง

1.7.8 ทำการทดลองเสียงที่เก็บรวบรวมมากับอัลกอริทึมที่สร้างขึ้นด้วยการทดลองย่อยๆ

1.7.9 สรุปและวิเคราะห์ผลการทดลอง และจัดทำวิทยานิพนธ์

1.8 ระยะเวลาดำเนินงานวิจัย

ตารางที่ 1.1 ระยะเวลาในการดำเนินการวิจัย

การดำเนินการวิจัย	ระยะเวลาศึกษา (เดือน)							
	1	2	3	4	5	6	7	8
ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง	■							
เก็บรวมรวมเสียงร้องของเด็กทารกเพื่อใช้ในการทดลอง	■	■						
ทำการทดลองและปรับปรุงอัลกอริทึมเพื่อให้ได้ประสิทธิภาพของการจำแนกเสียงสูงสุด		■	■	■				
เก็บรวมรวมเสียงอื่นๆเพื่อทดสอบประสิทธิภาพของอัลกอริทึมที่พัฒนาขึ้นมา				■	■			
ทดลองอัลกอริทึมที่สร้างขึ้นมากับเสียงร้องอื่นๆ						■	■	
วิเคราะห์ สรุปผลและจัดทำวิทยานิพนธ์								■