อุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน

DIALOGUE BOX FOR HEARING IMPAIRMENT

โดย

นางสาวสุดารัตน์ คำยันต์

นายเสฏฐวุฒิ ชุมแสง

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2561

อุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน

DIALOGUE BOX FOR HEARING IMPAIRMENT

โดย

นางสาวสุดารัตน์ คำยันต์ 58011336

นายเสฏฐวุฒิ ชุมแสง 58011378

อาจารย์ที่ปรึกษา

ผศ. ดร.ตุลยา ลิมปิติ

อาจารย์ที่ปรึกษาร่วม

รศ. ดร.จีรสุดา โกษียาภรณ์

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2561

ผ่านการตรวจรูปเล่มแล้ว

(.....)

อาจารย์ที่ปรึกษา

23/5/62

KMITL วิศวกรรมโทรคมนาคม Telecommunications Engineering

ผ่านการตรวจชิ้นงานแล้ว

(.....)

กรรมการผู้ตรวจชิ้นงาน

23/5/62

KMITL วิศวกรรมโทรคมนาคม Telecommunications Engineering

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2561

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง อุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน

DIALOGUE BOX FOR HEARING IMPAIRMENT

ผู้จัดทำ

1. นางสาวสุดารัตน์ คำยันต์ 58011336
2. นายเสฏฐวุฒิ ชุมแสง 58011378

..... อาจารย์ที่ปรึกษา

(ผศ. ดร.ตุลยา ลิมปิติ)

..... อาจารย์ที่ปรึกษาร่วม

(รศ. ดร.จีรสุดา โกษียาภรณ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

การดำเนินปริญญานิพนธ์เรื่อง “อุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน” จะไม่สามารถสำเร็จลุล่วงไปได้ด้วยดี หากไม่ได้รับความช่วยเหลือ และความอนุเคราะห์อย่างดียิ่งจาก ผศ. ดร.ตุลยา ลิมปิติ อาจารย์ที่ปรึกษาปริญญานิพนธ์, รศ. ดร.จีรสุดา โกษียาภรณ์ อาจารย์ที่ปรึกษาร่วม และ ผศ. ดร.ณัฐกานต์ พุทธรักษ์ ที่กรุณาให้คำแนะนำ คำปรึกษา และแนวทางการแก้ไขปัญหาที่เป็นประโยชน์ต่อการศึกษา ค้นคว้าวิจัยให้ปริญญานิพนธ์นี้สำเร็จสมบูรณ์ยิ่งขึ้น รวมถึงสนับสนุนสถานที่ เครื่องมือ และอุปกรณ์ต่างๆ ที่ใช้ระหว่างการจัดทำปริญญานิพนธ์ ขอขอบพระคุณทุกท่านเป็นอย่างสูงในความเอาใจใส่ และความปรารถนาดีที่มีให้แก่ผู้จัดทำเสมอมา

ขอขอบคุณคณาจารย์และเจ้าหน้าที่ประจำภาควิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกท่าน ที่ได้อบรมสั่งสอนประสิทธิ์ประสาทวิชา ความรู้ และประสบการณ์ให้แก่ผู้จัดทำ

ขอบคุณเพื่อนทุกคนที่ให้ความช่วยเหลือ และคำแนะนำที่เป็นประโยชน์ในการจัดทำปริญญานิพนธ์ ตลอดจนขอกราบขอบพระคุณ บิดา มารดา และครอบครัว ที่เป็นกำลังใจสำคัญและให้โอกาสการศึกษาอันมีค่ายิ่ง

ขอขอบคุณผู้เขียนหนังสือ รวมทั้งบทความต่างๆ ที่ช่วยสร้างแนวความคิด และให้ความรู้ที่เป็นประโยชน์แก่ผู้จัดทำ ตลอดจนผู้มีส่วนเกี่ยวข้องที่ไม่ได้กล่าวถึงทุกท่านที่คอยให้การสนับสนุนช่วยเหลือ ให้ความรู้เพิ่มเติม จนสามารถทำให้ปริญญานิพนธ์นี้สำเร็จลุล่วงไปได้ด้วยดี

น.ส.สุดารัตน์ คำยันต์
นายเสฏฐวุฒิ ชุมแสง
ผู้จัดทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน
DIALOGUE BOX FOR HEARING IMPAIRMENT

โดย นางสาวสุดารัตน์ คำยันต์ 58011336
นายเสฏฐวุฒิ ชุมแสง 58011378

อาจารย์ที่ปรึกษา ผศ. ดร.ตุลยา ลิมปิติ
อาจารย์ที่ปรึกษาร่วม รศ. ดร.จีรสุดา โกษียาภรณ์

บทคัดย่อ

การสนทนาของผู้พิการทางการได้ยินกับบุคคลทั่วไปมีความยากลำบาก เพราะบุคคลทั่วไปส่วนมากไม่รู้ภาษามือ เพื่อการสื่อสารที่ง่ายขึ้นและลดความไม่เท่าเทียมของผู้พิการ ปริญญานิพนธ์นี้จึงออกแบบอุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยินแบบพกพาให้สนทนาโดยใช้ภาษามือกับบุคคลทั่วไปได้โดยไม่ต้องพึ่งล่าม และเป็นการลดความแตกต่าง ทำให้ผู้พิการรู้สึกเท่าเทียมกับบุคคลปกติ โดยระบบที่ออกแบบและพัฒนาขึ้นประกอบไปด้วย 3 ส่วน คือ ส่วนการรับภาษามือจากผู้พิการทางการได้ยินผ่านกล้อง, ส่วนการตอบกลับโดยรับเสียงจากคู่สนทนาด้วยไมโครโฟน ซึ่งทั้งสองส่วนประมวลผลโดยใช้ไมโครคอมพิวเตอร์ และส่วนสุดท้ายคือส่วนการเก็บบันทึกการสนทนาและการแสดงผลการสนทนาผ่านจอแสดงผล

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ABSTRACT

Conversations between a person with hearing impairment and a normal person is difficult because normal people who know sign language are rare. Therefore, to facilitate the conversation, this project aims at developing a portable dialogue box that can be used without an interpreter. The system consists of three parts. The first component detects sign language of the hearing impaired through a camera. The vocal reply from a normal person is received through a microphone. Both parts are processed by a microcomputer to produce texts of the conversations. The last part of the system is the display of the dialogue on a display screen. The system can also record the conversation logs afterward.

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

		หน้า
กิตติกรรมประกาศ		I
บทคัดย่อ		II
สารบัญ		IV
สารบัญรูป		VIII
สารบัญตาราง		XIV
บทที่ 1	**บทนำ**	
	1.1 ความเป็นมาและความสำคัญของปัญหา	1
	1.2 วัตถุประสงค์	1
	1.3 ขอบเขตของปริญญานิพนธ์	1
บทที่ 2	**ทฤษฎีและหลักการที่เกี่ยวข้อง**	
	2.1 ภาษามือ	2
	2.2 ไมโครคอมพิวเตอร์	3
	2.3 ภาษาไพธอน (Python)	5
	2.4 OpenCV	6
	2.5 การแปลเสียงพูด	7
	2.6 การสร้างฐานข้อมูล	8
	2.7 ปริภูมิสี RGB และปริภูมิสี HSV	9
	2.8 Histogram of Oriented Gradients (HOG)	11
	2.9 Support Vector Machine (SVM)	12
	2.9.1 SVM แบ่งแยกกลุ่มข้อมูลแบบเชิงเส้น (Linear kernel)	13
	2.9.2 SVM แบ่งแยกกลุ่มข้อมูลแบบไม่เป็นเชิงเส้น	15

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

		หน้า
บทที่ 3	**การออกแบบและการจัดทำปริญญานิพนธ์**	
	3.1 การออกแบบ	17
	3.1.1 การออกแบบการทำงานของระบบ	17
	3.1.2 การออกแบบส่วนประมวลภาพของโปรแกรมแปลภาษามือเป็นตัวอักษร	17
	3.1.3 การออกแบบการทำงานของอัลกอริทึม SVM สำหรับการเรียนรู้ของโปรแกรมแปลภาษามือเป็นตัวอักษร	22
	3.1.4 การออกแบบการทำงานของระบบบันทึกข้อความสนทนาจากผู้พิการทางการได้ยิน	25
	3.1.5 การออกแบบการทำงานของระบบบันทึกข้อความสนทนาจากคนปกติ	26
	3.1.6 การออกแบบการทำงานของระบบเรียกดูประวัติการสนทนา	27
	3.1.7 การออกแบบตารางฐานข้อมูลในการจัดเก็บข้อมูล	28
	3.1.8 การแปลเสียงพูดเป็นข้อความ	30
	3.1.9 การออกแบบส่วนต่อประสานกราฟิกกับผู้ใช้ (GUI)	32
	3.1.10 การออกแบบกล่องอุปกรณ์	36
	3.2 เครื่องมือที่ใช้ในการทดลอง	38
	3.2.1 ไมโครคอมพิวเตอร์	38
	3.2.2 ไมโครโฟน	38
	3.2.3 จอแสดงผล	39
	3.2.4 กล้องเว็บแคม	39
	3.3 การจัดเก็บผลการทดลอง	40
	3.3.1 การทดสอบข้อมูลสำหรับการเรียนรู้ของโปรแกรมแปลภาษามือ	40

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

		หน้า
	3.3.2 การทดสอบพารามิเตอร์สำหรับการเรียนรู้ของโปรแกรมแปลภาษามือเป็นตัวอักษร	40
	3.3.3 ความแม่นยำในการจำแนกตัวอักษรของ Linear SVM แบบ multi-class	40
	3.3.4 การทดสอบส่วนต่อประสานกราฟิกกับผู้ใช้และฐานข้อมูล	40
	3.3.5 การทดสอบผลลัพธ์โปรแกรมแปลภาษามือเป็นตัวอักษร	41
	3.3.6 การทดสอบผลลัพธ์โปรแกรมแปลเสียงเป็นตัวอักษร	41
	3.3.7 การทดสอบการทำงานของระบบรวม	41
บทที่ 4	**ผลการทดลอง**	
	4.1 การทดสอบข้อมูลสำหรับการเรียนรู้ของโปรแกรมแปลภาษามือ	42
	4.2 การทดสอบพารามิเตอร์สำหรับการเรียนรู้ของโปรแกรมแปลภาษามือเป็นตัวอักษร	49
	4.2.1 ทดสอบพารามิเตอร์ของ Binary และ Multi-class SVM แบบ Linear kernel	49
	4.2.2 ทดสอบพารามิเตอร์ของ Binary SVM แบบ RBF kernel	55
	4.3 ความแม่นยำในการจำแนกตัวอักษรที่ของ Linear SVM แบบ multi-class	58
	4.4 การทดสอบส่วนต่อประสานกราฟิกกับผู้ใช้และฐานข้อมูล	60
	4.4.1 การทดสอบหน้าต่าง Login และลงทะเบียนเข้าใช้งานกับฐานข้อมูล	60
	4.4.2 การทดสอบหน้าต่าง Add Friend กับฐานข้อมูล และการแสดงรายชื่อผู้ติดต่อในหน้าต่าง Dialogue Box	63

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

		หน้า
	4.4.3 การทดสอบหน้าต่าง Exit และการบันทึกข้อมูลการสนทนาเป็นไฟล์ .txt	65
	4.4.4 การทดสอบหน้าต่าง History กับฐานข้อมูล	66
	4.5 การทดสอบผลลัพธ์โปรแกรมแปลภาษามือเป็นตัวอักษร	67
	4.5.1 การทดสอบผลลัพธ์การแปลภาษามือของตัวอักษร A-Z และเครื่องหมายวรรคตอนต่างๆ	67
	4.5.2 การทดสอบการลบตัวอักษรด้วยภาษามือ	72
	4.5.3 การทดสอบการส่งตัวอักษรด้วยภาษามือกับฐานข้อมูล	73
	4.5.4 การทดสอบการแสดงผลลัพธ์โปรแกรมแปลภาษามือเป็นตัวอักษรบนส่วนต่อประสานกราฟิกกับผู้ใช้	75
	4.6 การทดสอบผลลัพธ์โปรแกรมแปลเสียงเป็นตัวอักษร	78
	4.7 การทดสอบการทำงานของระบบรวม	81
	4.7.1 การทดสอบการทำงานของระบบรวมบนคอมพิวเตอร์	81
	4.7.2 การทดสอบการทำงานของระบบรวมบน Raspberry Pi	85
บทที่ 5	**สรุปผลและข้อเสนอแนะ**	
	5.1 สรุปผล	89
	5.2 ข้อเสนอแนะ	89
บรรณานุกรม		**90**

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่		หน้า
2.1	ภาษามือของตัวอักษร A-Z และเครื่องหมายวรรคตอน	3
2.2	Raspberry Pi 3 Model B+	3
2.3	ไดอะแกรมพอร์ต GPIO ของ Raspberry Pi Model B+	4
2.4	หน้าต่างโปรแกรม PyCharm	6
2.5	ขั้นตอนของระบบแยกแยะเสียงพูด (Speech Recognition)	8
2.6	หน้าต่างโปรแกรม phpMyAdmin ในการจัดการฐานข้อมูล MySQL	9
2.7	ปริภูมิสี RGB	10
2.8	ปริภูมิสี HSV	10
2.9	การแบ่งภาพของ HOG	11
2.10	บล็อกไดอะแกรมของขั้นตอน HOG	11
2.11	การกระจายตัวของข้อมูลคลาส A และ B	13
2.12	โมเดลเส้นตรง	13
2.13	SVM แบบเชิงเส้น	14
2.14	ค่า C ที่มีผลต่อระยะ Margin	15
2.15	การเปลี่ยนแปลงของ gamma ที่มีผลต่อขอบการตัดสินใจ	16
3.1	บล็อกไดอะแกรมของอุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน	17
3.2	ขั้นตอนการประมวลผลภาพภาษามือ	17
3.3	การรับภาพจากกล้องเว็บแคมและสร้างกรอบที่สนใจ	18
3.4	ภาพที่ถูกลบพื้นหลัง	18
3.5	การทำให้บริเวณสีผิวเป็นสีขาว	19
3.6	การลบจุดสีดำ	19
3.7	การสร้างเกณฑ์ให้รูปภาพเป็นรูปขาวดำ	20
3.8	ปรับขนาดภาพให้บริเวณขอบมือพอดีกับขนาด 64x64 พิกเซล	20
3.9	การหาเกรเดียนท์ของภาพ	21

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่		หน้า
3.10	ขั้นตอนการหา HOG	21
3.11	การใช้ผลลัพธ์การทำนายรูปภาพกับข้อมูลที่ถูกเรียนรู้ด้วยอัลกอริทึม SVM ในการสนทนา	21
3.12	แผนผังการทำงานของระบบบันทึกข้อความสนทนา	26
3.13	แผนผังการทำงานของระบบบันทึกข้อความสนทนาจากคนปกติ	27
3.14	แผนผังการทำงานของระบบเรียกดูประวัติการสนทนา	28
3.15	การสร้างฐานข้อมูล project4a	28
3.16	ตาราง account	29
3.17	ตาราง friends	29
3.18	ตาราง gui	30
3.19	ตาราง text_sign	30
3.20	โปรเจคที่สร้างบน Google Cloud Platform	31
3.21	Cloud Speech API	31
3.22	ข้อความที่แสดงบนส่วนต่อประสานกราฟิกกับผู้ใช้	31
3.23	หน้าต่าง Login และลงทะเบียน	32
3.24	การแจ้งเตือนเมื่อกรอก Username หรือ Password ไม่ถูกต้อง	32
3.25	หน้าต่าง Dialogue box	33
3.26	หน้าต่าง Add Friend	34
3.27	หน้าต่าง Exit	34
3.28	ข้อความที่ถูกบันทึกในรูปไฟล์ .txt	35
3.29	หน้าต่าง Dialogue box สำหรับคนปกติ	35
3.30	หน้าต่าง History	36
3.31	การออกแบบกล่องอุปกรณ์	36
3.32	ภาพฉายด้านหน้า ด้านข้าง และด้านบนของต้นแบบกล่องอุปกรณ์	37

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่		หน้า
3.33	กล่องอุปกรณ์ที่จัดทำขึ้น	37
3.34	Raspberry Pi 3 Model B+	38
3.35	มินิโมโครโฟน USB 2.0	39
3.36	จอแสดงผลขนาด 5 นิ้ว	39
3.37	กล้องเว็บแคม Logitech C170	40
4.1	ผลการตรวจสอบสัญลักษณ์ที่ใช้ในการเรียนรู้	45
4.2	จำนวนรูปภาพทั้งหมดที่ใช้เรียนรู้	45
4.3	ผลการตรวจสอบสัญลักษณ์ที่ใช้ในการทดสอบ	48
4.4	จำนวนรูปภาพทั้งหมดที่ใช้ทดสอบ	48
4.5	กราฟความสัมพันธ์ของ C และ f1-score ของข้อมูล A-B สำหรับข้อมูลที่ใช้ฝึกสอน	49
4.6	กราฟความสัมพันธ์ของ C และ Accuracy ของข้อมูล A-B สำหรับข้อมูลที่ใช้ฝึกสอน	49
4.7	กราฟความสัมพันธ์ของ C และ f1-score ของข้อมูล A-B สำหรับข้อมูลที่ใช้ทดสอบ	50
4.8	กราฟความสัมพันธ์ของ C และ Accuracy ของข้อมูล A-B สำหรับข้อมูลที่ใช้ทดสอบ	50
4.9	กราฟความสัมพันธ์ของ C และ f1-score ของข้อมูล A-N สำหรับข้อมูลที่ใช้ฝึกสอน	52
4.10	กราฟความสัมพันธ์ของ C และ Accuracy ของข้อมูล A-N สำหรับข้อมูลที่ใช้ฝึกสอน	53
4.11	กราฟความสัมพันธ์ของ C และ f1-score ของข้อมูล A-N สำหรับข้อมูลที่ใช้ทดสอบ	53

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่		หน้า
4.12	กราฟความสัมพันธ์ของ C และ Accuracy ของข้อมูล A-N สำหรับข้อมูลที่ใช้ทดสอบ	53
4.13	กราฟความสัมพันธ์ของ C และ f1-score ของข้อมูลทุกตัวอักษรสำหรับข้อมูลที่ใช้ฝึกสอน	54
4.14	กราฟความสัมพันธ์ของ C และ Accuracy ของข้อมูลทุกตัวอักษรสำหรับข้อมูลที่ใช้ฝึกสอน	54
4.15	กราฟความสัมพันธ์ของ C และ f1-score ของข้อมูลทุกตัวอักษรสำหรับข้อมูลที่ใช้ทดสอบ	55
4.16	กราฟความสัมพันธ์ของ C และ Accuracy ของข้อมูลทุกตัวอักษรสำหรับข้อมูลที่ใช้ทดสอบ	55
4.17	กราฟความสัมพันธ์ของ C, gamma และ f1-score ของข้อมูล A-B สำหรับข้อมูลที่ใช้ฝึกสอน	56
4.18	กราฟความสัมพันธ์ของ C, gamma และ Accuracy ของข้อมูล A-B สำหรับข้อมูลที่ใช้ฝึกสอน	56
4.19	กราฟความสัมพันธ์ของ C, gamma และ f1-score ของข้อมูล A-B สำหรับข้อมูลที่ใช้ทดสอบ	57
4.20	กราฟความสัมพันธ์ของ C, gamma และ Accuracy ของข้อมูล A-B สำหรับข้อมูลที่ใช้ทดสอบ	58
4.21	หน้าต่าง Login สำหรับเข้าสู่หน้าต่างการสนทนา	60
4.22	ฐานข้อมูลสำหรับเก็บข้อมูล Username และ Password	61
4.23	การแจ้งเตือนหากกรอก Username และ Password ไม่ถูกต้อง	61
4.24	หน้าต่างลงทะเบียนผู้ใช้งาน	62
4.25	ข้อความยืนยันการลงทะเบียนสำเร็จ	62
4.26	การบันทึกข้อมูลการลงทะเบียนลงในฐานข้อมูล	63
4.27	การกรอกรายชื่อผู้ติดต่อใหม่ในหน้าต่าง Add Friend	63

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่		หน้า
4.28	ข้อความยืนยันการเพิ่มรายชื่อผู้ติดต่อสำเร็จ	64
4.29	รายชื่อผู้ติดต่อใหม่บนหน้าต่าง Dialogue box	64
4.30	ฐานข้อมูลสำหรับเก็บข้อมูลรายชื่อผู้ติดต่อ	65
4.31	การบันทึกข้อมูลการสนทนา	65
4.32	ข้อมูลการสนทนาที่ถูกบันทึกในไฟล์ .txt	66
4.33	การเรียกดูประวัติการสนทนา	66
4.34	ฐานข้อมูลประวัติการสนทนา	67
4.35	ประวัติการสนทนาจากการค้นหาคำ	67
4.36	ผลลัพธ์ในการแปลภาษามือตัวอักษร a ถึง g	68
4.37	ผลลัพธ์ในการแปลภาษามือตัวอักษร h ถึง n	69
4.38	ผลลัพธ์ในการแปลภาษามือตัวอักษร o ถึง u	70
4.39	ผลลัพธ์ในการแปลภาษามือตัวอักษร v ถึง z	71
4.40	ผลลัพธ์ในการแปลภาษามือเครื่องหมายวรรคตอนต่างๆ	72
4.41	ตัวอย่างการลบตัวอักษรที่ผิด	73
4.42	ตัวอย่างการส่งข้อความ	74
4.43	ข้อความ hi giftfy. ถูกบันทึกลงในฐานข้อมูล	75
4.44	ภาษามือคำว่า “hi giftfy.” ที่แสดงผ่านหน้าจอ GUI	75
4.45	ผลลัพธ์ในการแปลภาษามือคำว่า How are you?	76
4.46	ข้อความใหม่ว่า “How are you?” บน Textbox ของ GUI เมื่อส่งคำว่า go	76
4.47	ข้อความที่ถูกบันทึกลงฐานข้อมูล	76
4.48	ตัวอย่างการอัพเดตข้อความใหม่	77
4.49	ตัวอย่างการบันทึกบทสนทนา	78
4.50	หน้าต่างสำหรับแปลเสียงพูดเป็นตัวอักษร	79

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่		หน้า
4.51	ข้อความที่แปลได้จากเสียงพูด	79
4.52	ข้อความที่ถูกบันทึกลงฐานข้อมูล	80
4.53	ข้อความถูกแสดงบนพื้นที่การสนทนา	80
4.54	หน้าต่างการสนทนา	81
4.55	ผลลัพธ์การแปลเสียงเป็นตัวอักษร	81
4.56	ผลลัพธ์การทำท่าทางภาษามือคำว่า “Sorry, I am not available.”	82
4.57	ภาษามือที่ถูกบันทึกลงในฐานข้อมูล	82
4.58	คำว่า “Sorry, I am not available.” บนส่วนต่อประสานกราฟิกกับผู้ใช้	82
4.59	การส่งคำว่า “Sorry, I am not available.” บนส่วนต่อประสานกราฟิกกับผู้ใช้	83
4.60	การบันทึกข้อความสนทนา	83
4.61	ประวัติการสนทนาเมื่อเลือกคู่สนทนา	84
4.62	ประวัติการสนทนาเมื่อเลือกวันที่	84
4.63	ประวัติการสนทนาเมื่อเลือกเวลา	85
4.64	ประวัติการสนทนาที่มีคำว่า can	85
4.65	หน้าต่าง Dialogue box ที่ได้ปรับปรุง	86
4.66	ผลลัพธ์ในการแปลภาษามือคำว่า Where are you from?	86
4.67	ข้อความ “Where are you from?” บน Textbox	87
4.68	ข้อความ “Where are you from?” บนส่วนต่อประสานกราฟิกกับผู้ใช้	87
4.69	ข้อความ “I am from Thailand” จากเสียงพูดบน Textbox	87
4.70	ข้อความ “I am from Thailand” บนส่วนต่อประสานกราฟิกกับผู้ใช้	88

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

หน้า

ตารางที่

3.1 จำนวนรูปภาพที่ใช้ในการเรียนรู้ 23

4.1 จำนวนรูปภาพที่ใช้ในการเรียนรู้อักษร A-N 51

4.2 จำนวนรูปภาพที่ใช้ในการทดสอบอักษร A-N 52

4.3 Accuracy ของแต่ละตัวอักษร 59

4.4 Accuracy ของเครื่องหมายวรรคตอนต่างๆ 60

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ผู้พิการทางการได้ยินมักมีปัญหาด้านการสื่อสารกับบุคคลอื่น เนื่องจากบุคคลที่สื่อสารกับผู้พิการมักไม่มีความรู้ด้านภาษามือ จึงทำให้ยากต่อการสื่อสารและอาจทำให้สื่อสารผิดความหมายไป อีกทั้งล่ามภาษามือมีไม่เพียงพอต่อจำนวนผู้พิการทางการได้ยิน และไม่สามารถอยู่กับผู้พิการได้ตลอดเวลา ปัญหาการสื่อสารนี้ทำให้เกิดความไม่เท่าเทียมระหว่างผู้พิการทางการได้ยินกับคนปกติ ด้วยเหตุนี้ปริญญานิพนธ์นี้จึงได้จัดทำอุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน ซึ่งช่วยให้ผู้พิการทางการได้ยินสื่อสารได้ง่าย สามารถสื่อสารได้ทุกเวลา รวมทั้งอุปกรณ์มีขนาดเล็กที่พกพาง่าย อุปกรณ์นี้จึงช่วยทำให้ผู้พิการทางการได้ยินสื่อสารได้โดยไม่ต้องพึ่งพาล่ามภาษามือ ทำให้ผู้พิการทางการได้ยินรู้สึกเท่าเทียมกับบุคคลอื่น

1.2 วัตถุประสงค์

1) เพื่อศึกษาการเขียนโปรแกรมแปลภาษามือและเสียงพูดเป็นตัวอักษร
2) เพื่อศึกษาการเขียนโปรแกรมสำหรับจอแสดงผล
3) เพื่อศึกษาและพัฒนาการสร้างฐานข้อมูล
4) เพื่อพัฒนาอุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน

1.3 ขอบเขตของปริญญานิพนธ์

ปริญญานิพนธ์นี้ได้ทำการออกแบบและสร้างอุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน โดยระบบประกอบไปด้วย 3 ส่วน คือ 1) การรับภาษามือจากผู้พิการทางการได้ยินผ่านกล้อง 2) การตอบกลับโดยรับเสียงคู่สนทนาที่เป็นคนปกติจากไมโครโฟน ซึ่งทั้งสองส่วนประมวลผลโดยใช้ไมโครคอมพิวเตอร์ 3) ส่วนฐานข้อมูลที่เก็บประวัติการสนทนาและแสดงผลผ่านส่วนต่อประสานกราฟิกกับผู้ใช้ (GUI) ผ่านจอแสดงผล ซึ่งการทำงานเริ่มต้นจากผู้พิการทางการได้ยินสนทนาโดยใช้ภาษามือผ่านกล้อง และส่งข้อความไปถึงคู่สนทนาผ่านส่วนต่อประสานกราฟิกกับผู้ใช้ คู่สนทนาตอบกลับโดยใช้เสียงผ่านไมโครโฟน ข้อความตอบกลับจะแสดงผ่านส่วนต่อประสานกราฟิกกับผู้ใช้เช่นกัน และสามารถดูประวัติการสนทนาย้อนหลังผ่านส่วนต่อประสานกราฟิกกับผู้ใช้ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและหลักการที่เกี่ยวข้อง

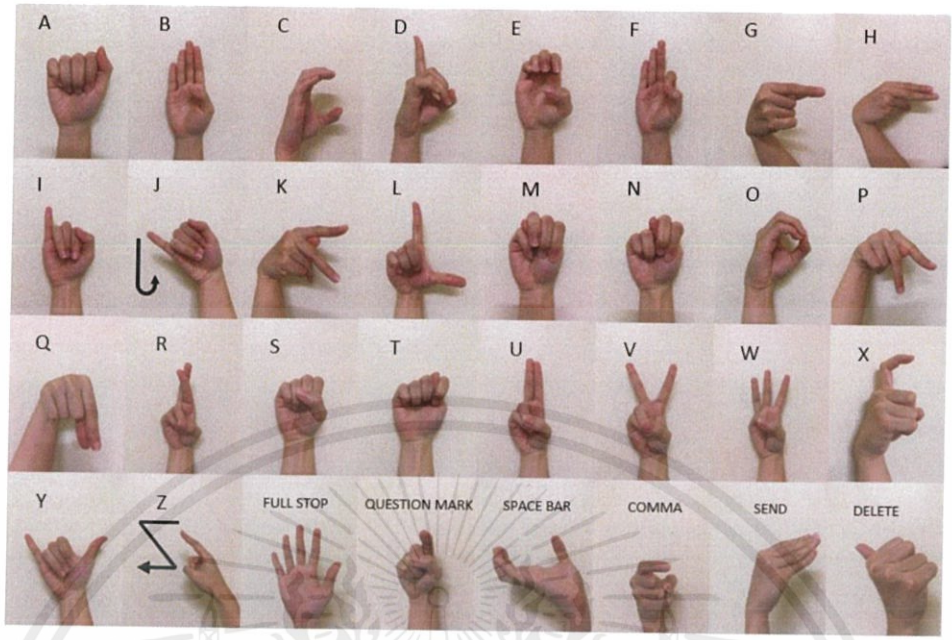
ปริญญานิพนธ์เรื่อง “อุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน” ได้ทำการออกแบบอุปกรณ์ที่ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน ซึ่งอุปกรณ์นี้ผู้พิการทางการได้ยินสามารถใช้ภาษามือในการสนทนา และบุคคลที่สื่อสารกับผู้พิการสามารถใช้เสียงในการสนทนาโดยระบบใช้ไมโครคอมพิวเตอร์ในการประมวลผล และข้อความที่สนทนาจะเก็บไว้ที่ฐานข้อมูลและแสดงผลบนส่วนต่อประสานกราฟิกกับผู้ใช้ผ่านหน้าจอแสดงผล ดังนั้นปริญญานิพนธ์จึงมีหลักการที่เกี่ยวข้องดังต่อไปนี้

2.1 ภาษามือ

ภาษามือ คือภาษาสำหรับผู้พิการทางการได้ยินโดยใช้มือ สีหน้า และกิริยาท่าทางประกอบในการสื่อความหมาย และถ่ายทอดอารมณ์แทนการพูด ภาษามือของแต่ละชาติมีความแตกต่างกัน เช่น ภาษามือจีน ภาษามืออเมริกัน และภาษามือไทย เป็นต้น ภาษามือเป็นภาษาหนึ่งที่ใช้สำหรับติดต่อสื่อสารระหว่างผู้พิการทางการได้ยินด้วยกัน และระหว่างผู้พิการทางการได้ยินกับคนปกติ โดยปริญญานิพนธ์นี้ออกแบบสำหรับภาษามือแบบอเมริกัน เนื่องจากเป็นสากล และไม่ซับซ้อนเท่ากับภาษามือไทย

ภาษามือแบบอเมริกัน ใช้การเคลื่อนย้ายมือและการแสดงออกทางสีหน้าและท่าทางเป็นภาษาหลักของชาวอเมริกาเหนือจำนวนมากที่เป็นผู้พิการทางการได้ยิน และเป็นหนึ่งในตัวเลือกการสื่อสารของผู้พิการทางการได้ยิน โดยปริญญานิพนธ์นี้ศึกษาภาษามือของตัวอักษรที่ใช้ในการสะกดคำในภาษาอังกฤษและเครื่องหมายวรรคตอน คือเครื่องหมายมหัพภาค (.) เครื่องหมายปรัศนีย์ (?) เครื่องหมายเว้นวรรค เครื่องหมายจุลภาค (,) ภาษามือแสดงการส่งข้อความ และภาษามือสำหรับลบตัวอักษร ซึ่งเครื่องหมายที่ไม่มีท่าทางภาษามือจะใช้ท่าทางที่สื่อถึงเครื่องหมายนั้นแทนเช่น ภาษามือเครื่องหมายมหัพภาคอ้างอิงมาจากภาษามือคำว่า Stop [1], ภาษามือเครื่องหมายเว้นวรรคอ้างอิงมาจากสัญลักษณ์เว้นวรรคที่คีย์บอร์ด, ภาษามือแสดงการส่งข้อความอ้างอิงมาจากภาษามือคำว่า Go, ภาษามือแสดงการลบตัวอักษรอ้างอิงมาจากภาษามือคำว่า Delete ดังรูปที่ 2.1

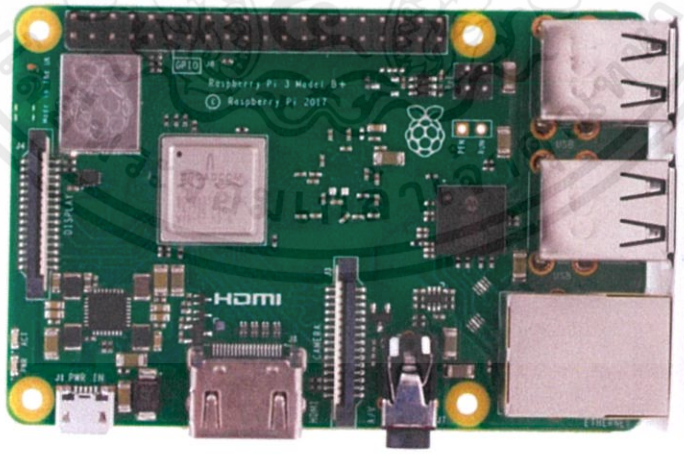
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.1 ภาษามือของตัวอักษร A-Z และเครื่องหมายวรรคตอน

2.2 ไมโครคอมพิวเตอร์

ไมโครคอมพิวเตอร์ (Microcomputer) คืออุปกรณ์ควบคุมขนาดเล็กซึ่งบรรจุความสามารถที่คล้ายคลึงกับระบบคอมพิวเตอร์ โดยในไมโครคอมพิวเตอร์มีส่วนซีพียู, หน่วยความจำ และพอร์ต ซึ่งเป็นส่วนประกอบหลักสำคัญของระบบคอมพิวเตอร์ โดยในปริญญานิพนธ์นี้ใช้ Raspberry Pi ดังรูปที่ 2.2



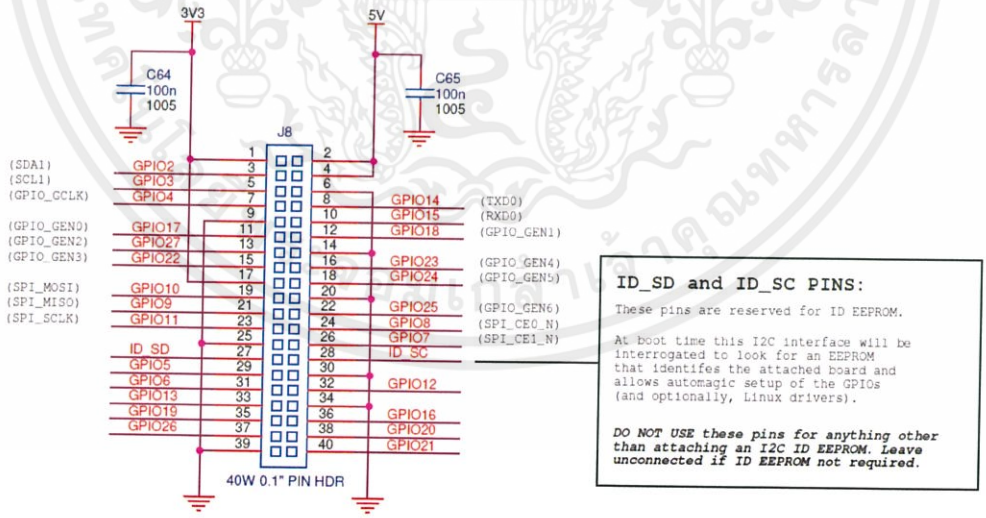
รูปที่ 2.2 Raspberry Pi 3 Model B+ [2]

Raspberry Pi เป็นเครื่องคอมพิวเตอร์ขนาดเล็ก มีราคาถูกเมื่อเทียบกับคอมพิวเตอร์ตั้งโต๊ะ สามารถต่อเข้ากับจอคอมพิวเตอร์ผ่านพอร์ต HDMI ต่อเมาส์ และคีย์บอร์ดผ่านพอร์ต USB ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แหล่งจ่ายไฟของ Raspberry Pi คือสาย Mini USB สำหรับชาร์จโทรศัพท์มือถือ Raspberry Pi 3 Model B+ มีคุณสมบัติดังนี้

- หน่วยประมวลผล Broadcom BCM2837B0, Cortex-A53 (ARMv8) 64-bit SoC @ 1.4GHz
- แรม 1GB LPDDR2 SDRAM
- 2.4GHz and 5GHz IEEE 802.11.b/g/n/ac wireless LAN, Bluetooth 4.2, BLE
- Gigabit Ethernet ผ่าน USB 2.0 (ความเร็วสูงสุด 300 Mbps)
- พิน GPIO 40 ขา
- ช่องต่อจอแสดงผล Full-size HDMI
- USB 2.0 4 ช่อง
- พอร์ตกล้อง CSI สำหรับเชื่อมต่อกล้อง Raspberry Pi
- พอร์ตแสดง DSI สำหรับเชื่อมต่อจอแสดงผลหน้าจอสัมผัส Raspberry Pi
- เอาต์พุตสเตอริโอ 4 ขาและพอร์ตวิดีโอคอมโพสิต
- พอร์ต Micro SD สำหรับโหลดระบบปฏิบัติการและจัดเก็บข้อมูล
- กำลังไฟ DC 5V/2.5A
- การสนับสนุน Power-over-Ethernet (PoE) (ต้องใช้ PoE HAT แยกต่างหาก)
- โดยแผนผังของขาเชื่อมต่อต่างๆ แสดงดังรูปที่ 2.3



รูปที่ 2.3 ไดอะแกรมพอร์ต GPIO ของ Raspberry Pi Model B+ [2]

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 ภาษาไพธอน (Python)

ไพธอน [3] เป็นภาษาคอมพิวเตอร์ยุคใหม่ หรือที่เรียกกันว่า ภาษาสคริปต์ ที่สามารถทำงานได้ทั้งรูปแบบโครงสร้างแบบเก่าและการเขียนเชิงวัตถุ (dynamic object oriented) และยังสามารถทำงานได้ทั้งแบบ Desktop และบนเครือข่ายอินเทอร์เน็ต หรือ Web application ต่างๆ สามารถทำงานได้ทุกระบบปฏิบัติการ เช่น Linux Windows และ FreeBSD รวมทั้งระบบปฏิบัติการ Mac/OS และจุดเด่นคือยังเป็น Open Source ภาษาไพธอนออกแบบมาเพื่อให้ทำงานได้กับ Web Application ที่ลักษณะคล้ายกับภาษา Perl, PHP, JAVA และ ASP เนื่องจากภาษาไพธอนเป็นภาษาที่ใหม่จึงมีคุณสมบัติที่ดีดังต่อไปนี้

1) สามารถใช้ได้ทุกแพลตฟอร์ม กล่าวคือสามารถทำงานได้บนหลายๆ ระบบปฏิบัติการ ผู้เขียนโปรแกรมอาจเขียนจากแพลตฟอร์มใดๆ แล้วนำโปรแกรมที่ได้ไปทำงานต่างแพลตฟอร์มกันได้

2) ไม่ต้องเสียค่าใช้จ่ายในการจัดซื้อโปรแกรมต้นฉบับ โดยปกติแล้วโปรแกรมภาษาทั่วๆ ไปจะต้องจัดซื้อโปรแกรมต้นฉบับเพื่อนำมาติดตั้งในราคาสูง แต่โปรแกรมภาษาไพธอนสามารถดาวน์โหลดจาก www.python.org ได้โดยตรง นำมาติดตั้ง และศึกษาการใช้งานด้วยตนเองได้ เนื่องจากเป็นโปรแกรมประเภท Open Source

3) ภาษาไพธอนได้นำเอาข้อดีของโปรแกรมในอดีตเข้ามาไว้ด้วยกัน เช่น ภาษา C, C++, Java และ Perl เป็นต้น

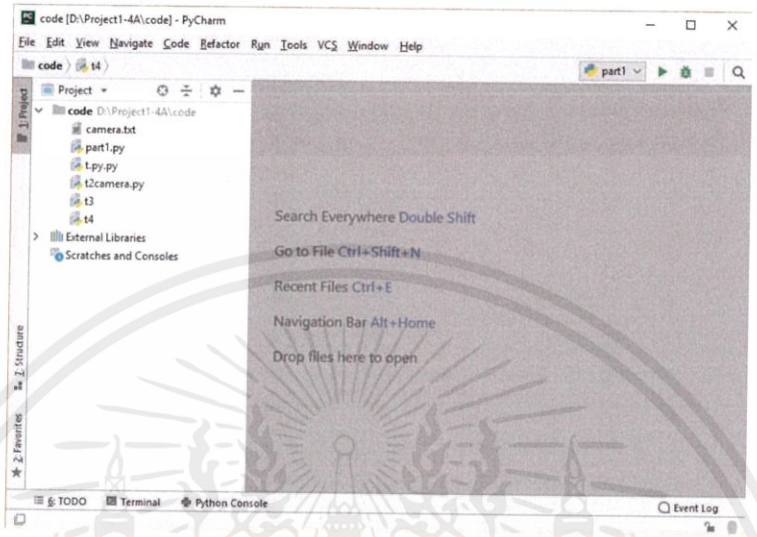
4) มีความปลอดภัยสูง เนื่องจากภาษาไพธอนทำงานอยู่ทางด้าน Server เป็นหลัก เมื่อมีการร้องขอจากเครื่อง Client จะประมวลผลที่เครื่อง Server ทำให้ผู้ใช้ทั่วไปไม่สามารถเข้าถึงเครื่อง Server ได้โดยตรง

5. ใช้ในการพัฒนา Web Service ซึ่งในปัจจุบันการพัฒนาซอฟต์แวร์ได้เน้นที่มีการแลกเปลี่ยนข้อมูลซึ่งกันและกันทั้งในองค์กรเดียวกันหรือแม้แต่ต่างองค์กรกัน ทำให้เกิดความสะดวกสบาย ไม่ต้องใช้ซอฟต์แวร์อื่นๆ มาแปลงข้อมูลเพื่อให้เข้ากันได้ เรียนรู้ได้เร็วกว่าโปรแกรมภาษาอื่นๆ เพราะมีโครงสร้างภาษาที่ไม่ซับซ้อน ซึ่งโครงสร้างภาษาคล้ายคลึงกับภาษา C แต่จะมีความกระชับและสั้นกว่าภาษา C

โปรแกรม PyCharm [4] ดังรูปที่ 2.4 เป็น Integrated Development Environment (IDE) ใช้เขียนโปรแกรมคอมพิวเตอร์ในภาษาไพธอน สามารถใช้งานได้ง่ายแม้ไม่ชำนาญในการเขียน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โค้ด และเป็นเครื่องมือสำหรับการพัฒนาโค้ดในไพธอน เนื่องจากมีฟังก์ชันต่างๆ มากมาย เช่น การนำทางที่ง่าย, การประมวลผลโค้ดอัจฉริยะ และการแก้ไขข้อผิดพลาดอย่างรวดเร็ว



รูปที่ 2.4 หน้าต่างโปรแกรม PyCharm

2.4 OpenCV

OpenCV [5] เป็น Open Source Computer Vision Library สำหรับทำงานด้าน image processing OpenCV เปิดตัวภายใต้ใบอนุญาต BSD และให้ใช้ฟรีสำหรับการใช้งานเชิงวิชาการและเชิงพาณิชย์ มีอินเตอร์เฟสหลายภาษา เช่น C++, Java, Python เป็นต้น และ OpenCV รองรับ Window, Linux, Android, และ Mac ซึ่ง OpenCV ได้รับการออกแบบมาเพื่อประสิทธิภาพในการคำนวณและเน้นการใช้งานแบบเรียลไทม์ โดย Toolbox และฟังก์ชันที่ใช้ในปริญญานิพนธ์นี้มีดังนี้

- imread ใช้ในการรับรูปภาพ
- imshow ใช้ในการแสดงรูปภาพ
- rgb2gray ใช้แปลงรูป RGB เป็นรูป grayscale
- imresize ใช้เปลี่ยนขนาดรูปภาพ
- threshold ใช้ในการสร้างเกณฑ์รูปภาพให้เป็นรูปขาวดำ
- dilate ใช้ในการขยายภาพ
- erode ใช้ในการกร่อนขนาดเพื่อลบค่า noise
- canny ใช้ในการหาขอบรูปภาพ
- CvtColor ใช้ในการแปลงปริภูมิสีของรูปภาพ
- createBackgroundSubtractorMOG2 ใช้ในการลบพื้นหลัง

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- GaussianBlur ใช้ในการเบลอรูปภาพ
- inRange ใช้ในการตรวจจับวัตถุจากการติดตามสี

2.5 การแปลเสียงพูด

Cloud Speech API [6] คือเทคโนโลยีการรู้จำเสียงพูด (speech recognition) ที่ใช้งานในผลิตภัณฑ์ต่างๆ ของ Google เช่น Google Search, Google Now, Google Assistant ที่ได้รับการปรับให้เหมาะสมกับความต้องการของผู้ที่ต้องการใช้งาน ข้อดีคือเป็นระบบการรู้จำเสียงพูดที่ถูกทดสอบ และพิสูจน์แล้วว่าสามารถใช้งานได้จริง สามารถใช้งานได้แม้ในที่ที่มีเสียงรบกวนมาก สามารถรองรับทั้งการแปลเสียงแบบเรียลไทม์ และแปลงไฟล์เสียงที่บันทึกมาแล้ว รองรับรูปแบบไฟล์ที่ถูกขยายแล้วรวมถึง WAV, Opus และ Speex สามารถแปลเสียงพูดเป็นตัวอักษรได้อย่างแม่นยำในเวลาใกล้เคียงกับเวลาจริง ซึ่ง Cloud Speech API สามารถรู้จำเสียงพูดของมนุษย์ได้กว่า 80 ภาษา และรองรับการแปลเสียงภาษาไทย ซึ่งประเภทของระบบแยกแยะเสียงพูดแบ่งออกเป็น 3 ประเภท ดังนี้

1) การแยกแยะเสียงพูดจากสัญญาณเสียงพูดแบบคำโดด (isolated speech) เป็นการฝึกฝนให้ระบบแยกแยะเสียงพูดเป็นคำสั้นๆ หรือแยกแยะคำสั่งบางคำ

2) การแยกแยะเสียงพูดจากสัญญาณเสียงพูดแบบต่อเนื่อง (continuous speech) เป็นการฝึกฝนให้ระบบแยกแยะทุกคำจากเสียงพูดแบบต่อเนื่อง และพิจารณาว่าสัญญาณเสียงนั้นมีคำสั่งใดบ้าง

3) การแยกแยะเสียงพูดจากสัญญาณเสียงพูดแบบคำอุทาน (spontaneous speech) เป็นการฝึกฝนให้ระบบแยกแยะเสียงที่มีคำอุทาน เพื่อให้ระบบรู้จำและเข้าใจเนื้อหาสำคัญที่อยู่ในสัญญาณเสียง

ขั้นตอนการสร้างระบบแยกแยะเสียงพูดดังรูปที่ 2.5 มีขั้นตอนดังนี้

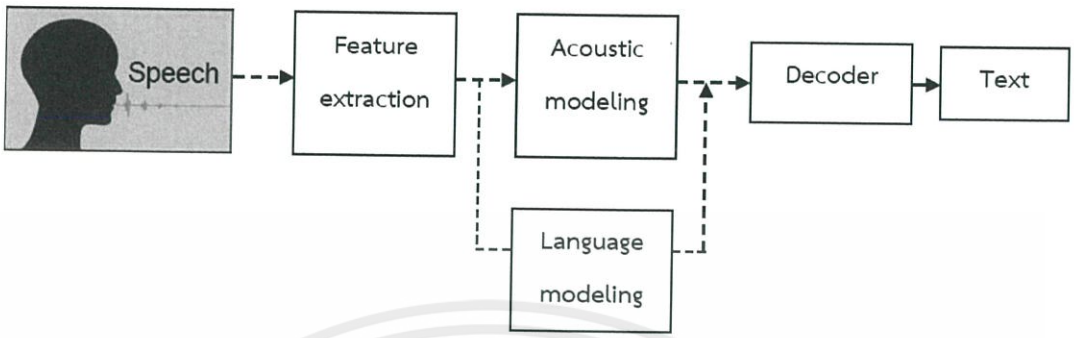
1) การสกัดคุณลักษณะ (Feature extraction) ของสัญญาณเสียงที่เข้ามาในระบบ

2) การกำหนดแบบจำลองเสียง (Acoustic modeling) ซึ่งทำหน้าที่วิเคราะห์สัญญาณเสียงที่เข้ามาในระบบ

3) การกำหนดแบบจำลองภาษา (Language modeling) ทำหน้าที่วิเคราะห์ความน่าจะเป็นในการเกิดขึ้นของคำจากลำดับคำที่กำหนด

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4) การถอดรหัส (Decoder) ทำหน้าที่ตีความสัญญาณเสียงเป็นข้อความ



รูปที่ 2.5 ขั้นตอนของระบบแยกแยะเสียงพูด (Speech Recognition)

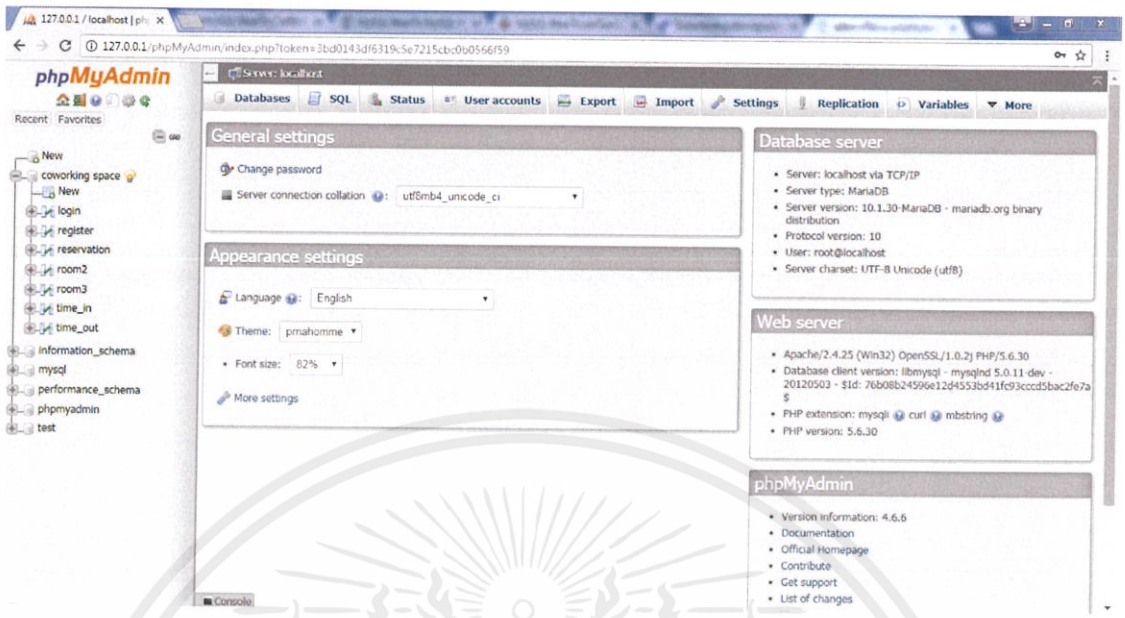
2.6 การสร้างฐานข้อมูล

ฐานข้อมูล [7] เป็นการจัดเก็บข้อมูลอย่างเป็นระบบ ทำให้ผู้ใช้สามารถใช้ข้อมูลที่เกี่ยวข้องในระบบงานต่างๆ ร่วมกันได้โดยที่ไม่เกิดความซ้ำซ้อนของข้อมูล ข้อมูลในระบบมีความถูกต้อง และเป็นมาตรฐานเดียวกัน โดยจะมีการกำหนดระบบความปลอดภัยของข้อมูล

ระบบฐานข้อมูล หมายถึง โครงสร้างสารสนเทศที่ประกอบด้วยรายละเอียดของข้อมูลที่เกี่ยวข้องกันเพื่อนำข้อมูลมาใช้ในระบบต่างๆ ร่วมกัน ซึ่งผู้ใช้สามารถจัดการกับข้อมูลเช่น การเพิ่ม การแก้ไข การลบ และการเรียกดูข้อมูลได้ โดยปริญญานิพนธ์นี้ใช้โปรแกรมระบบจัดการฐานข้อมูลคือ MySQL

MySQL คือโปรแกรมระบบจัดการฐานข้อมูล ที่พัฒนาโดยบริษัทมายเอสคิวแอลเอบี (MySQL AB) มีหน้าที่เก็บข้อมูลอย่างเป็นระบบ รองรับคำสั่ง SQL เป็นเครื่องมือสำหรับเก็บข้อมูลที่ต้องใช้ร่วมกับเครื่องมือหรือโปรแกรมอื่นเพื่อให้ได้ระบบงานที่รองรับความต้องการของผู้ใช้ โปรแกรมถูกออกแบบให้สามารถทำงานได้บนระบบปฏิบัติการที่หลากหลาย เช่น UNIX OS/2, MAC OS และ Windows และเป็นระบบฐานข้อมูล Open Source ที่ถูกนำไปใช้งานมาก ซึ่งสามารถใช้งานร่วมกับภาษา C, C++ , Java, Perl, PHP และ Python ได้ โดยมีโปรแกรมที่ใช้ในการจัดการฐานข้อมูล MySQL ผ่านบราวเซอร์คือ phpMyAdmin ดังรูปที่ 2.6

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 หน้าต่างโปรแกรม phpMyAdmin ในการจัดการฐานข้อมูล MySQL

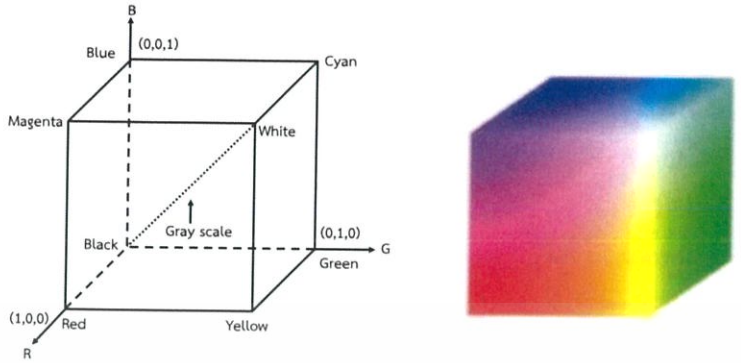
phpMyAdmin [8] คือโปรแกรมที่ถูกพัฒนาโดยใช้ภาษา PHP เพื่อใช้ในการบริหารจัดการฐานข้อมูล MySQL แทนการคีย์คำสั่ง phpMyAdmin ถือเป็นส่วนต่อประสานที่สร้างโดยภาษา PHP ซึ่งใช้จัดการฐานข้อมูล MySQL ผ่านเว็บเบราว์เซอร์ โดยสามารถที่จะทำการสร้างฐานข้อมูลใหม่ หรือทำการสร้างตารางข้อมูล และยังมีฟังก์ชันที่ใช้สำหรับการทดสอบการ Query ข้อมูลด้วยภาษา SQL พร้อมกันนั้นยังสามารถทำการ insert delete update หรือสามารถใช้คำสั่งต่างๆ เหมือนกับการใช้ภาษา SQL ในการสร้างตารางข้อมูล

phpMyAdmin เป็นโปรแกรมประเภท MySQL Client ตัวหนึ่งที่ใช้ในการจัดการข้อมูล MySQL ผ่าน web browser ได้โดยตรง phpMyAdmin จะทำงานบน Web server เป็น PHP Application ที่ใช้ควบคุมจัดการ MySQL Server

2.7 ปริภูมิสี RGB และปริภูมิสี HSV

ปริภูมิสี RGB เป็นปริภูมิสีที่ใช้ในระบบภาพดิจิทัล โดยแทนค่าสีของจุดภาพแต่ละจุดภาพด้วยเวกเตอร์สามมิติ ซึ่งแทนค่าสีปฐมภูมิได้แก่ สีแดง (R), สีเขียว (G) และสีน้ำเงิน (B) โดยมีช่วงของค่าสีระหว่าง 0-255 ซึ่งปริภูมิสี RGB จะมีลักษณะเป็นรูปลูกบาศก์ดังรูปที่ 2.7

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



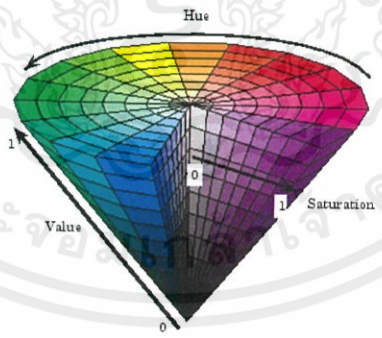
รูปที่ 2.7 ปริภูมิสี RGB [9]

ปริภูมิสี HSV ใช้หลักการแยกความสว่างออกจากเนื้อสีของจุดภาพ โดยปริภูมิสี HSV องค์ประกอบหลักสามส่วนดังนี้

2.7.1 Hue (H) คือค่าของแม่สีหลัก (สีแดง สีเขียว และสีน้ำเงิน) โดยค่าของสีจะอยู่ระหว่าง 0-1 ซึ่งถ้า Hue มีค่าเท่ากับ 0 จะแทนที่ด้วยสีแดง และเมื่อ Hue มีค่าเพิ่มขึ้นสีจะเปลี่ยนแปลงไปตามสเปคตรัมของสีจนถึงค่า 1 จึงจะกลับมาเป็นสีแดงอีกครั้ง

2.7.2 Saturation (S) คือค่าความบริสุทธิ์ของสี ซึ่งค่า Saturation มีค่าเท่ากับ 0 คือเป็นสีขาวล้วน แต่ถ้าค่า Saturation มีค่าเท่ากับ 1 แสดงว่าไม่มีแสงสีขาวผสมอยู่เลย

2.7.3 Value (V) คือค่าความสว่างของสี ซึ่งสามารถวัดได้จากค่าความเข้มของความสว่างของแต่ละสีที่ประกอบกัน ซึ่งปริภูมิสี HSV จะมีลักษณะเป็นรูปกรวยดังรูปที่ 2.8

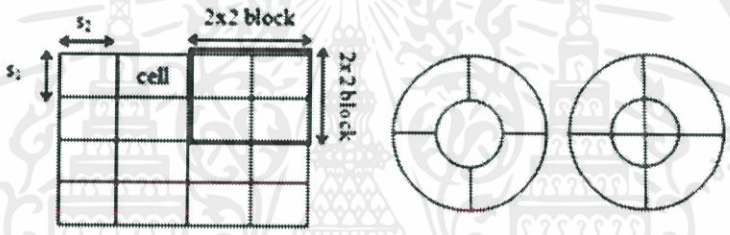


รูปที่ 2.8 ปริภูมิสี HSV [9]

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.8 Histogram of Oriented Gradients (HOG)

Histogram of Oriented Gradients (HOG) คือการสกัดคุณลักษณะของภาพด้วยรูปร่าง (Feature Extraction) ซึ่งใช้วิธีการหาขอบแล้วแบ่งภาพเป็นภาพย่อย วิธีการแบ่งเป็นภาพย่อยหรือเรียกว่าเซลล์ (Cells) มี 2 แบบ คือแบบสี่เหลี่ยมและแบบวงกลมดังรูปที่ 2.9 โดยแต่ละเซลล์จะรวบรวมฮีสโทแกรมของทิศทางเกรเดียนท์ หรือทิศทางของขอบภายในเซลล์ที่มีขนาดหนึ่งมิติ (1-D) เพื่อแสดงถึงคุณลักษณะเฉพาะของวัตถุที่สนใจ เพื่อประสิทธิภาพของความถูกต้องที่เพิ่มมากขึ้น ทำการนอร์มัลไลซ์ด้วยการคำนวณตัวชี้วัดของค่าความเข้มทั่วทั้งพื้นที่ขนาดใหญ่ของภาพ โดยผลลัพธ์จากการนอร์มัลไลซ์ คือการทำให้ผลกระทบจากการเปลี่ยนแปลงของแสงสว่างและเงาน้อยลง และสามารถหาคุณลักษณะเฉพาะของวัตถุได้ดีมากขึ้น ขั้นตอนของ HOG แสดงดังรูปที่ 2.10 โดยมีรายละเอียดดังนี้



(ก) (ข)

รูปที่ 2.9 การแบ่งภาพของ HOG [10]
(ก) การแบ่งภาพย่อยแบบสี่เหลี่ยม (ข) การแบ่งภาพย่อยแบบวงกลม

รูปที่ 2.10 บล็อกไดอะแกรมของขั้นตอน HOG

กำหนดให้ $f(x,y)$ คือ พิกเซลของภาพที่ตำแหน่ง (x,y)

1) กำหนดขนาดรูปภาพและคำนวณเกรเดียนต์ของภาพดังสมการที่ (2.1)

$$G_H(x,y) = f(x+1,y) - f(x-1,y)$$
$$G_V(x,y) = f(x,y+1) - f(x,y-1) \qquad (2.1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) แบ่งภาพเป็นภาพย่อย (block) และกำหนดจำนวนเซลล์ในแต่ละภาพย่อย

3) ทำการคำนวณขนาดและทิศทางของภาพดังสมการที่ (2.2)

$$|G(x,y)| = \sqrt{G_H(x,y)^2 + G_V(x,y)^2}$$
$$\theta_G(x,y) = \tan^{-1}\left(\frac{G_V(x,y)}{G_H(x,y)}\right) \quad (2.2)$$

4) หาผลรวมของแต่ละทิศทางที่กำหนดเป็นจำนวน n bins ดังสมการที่ (2.3)

$$C_b = \sum_{i=1}^{n} \theta(x,y)_i \quad (2.3)$$

โดย n คือ จำนวนตำแหน่ง (x,y) ในแต่ละภาพย่อย (block)

b คือ ทิศทางที่พิจารณา

C_b คือ ผลรวมของแต่ละทิศทาง

5) คำนวณคุณลักษณะเด่นของทุกภาพย่อยและถ่วงน้ำหนักด้วยขนาดของเกรเดียนท์ดังสมการที่ (2.4)

$$V_k = \sum_{i=1}^{n} \left(|G(x,y)| * C_b\right)\theta(x,y) \quad (2.4)$$

6) ค่านอร์มัลไลซ์คุณลักษณะของภาพย่อยด้วย L_1 - norm ดังสมการที่ (2.5)

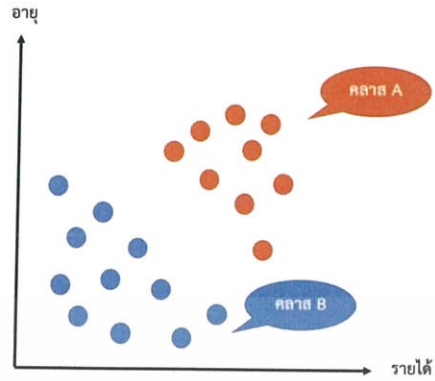
$$V_k = \frac{V_k}{\|V_k\|_1 + \varepsilon} \quad (2.5)$$

เมื่อ ε คือเป็นค่าคงที่เป็นบวกน้อยๆ เพื่อป้องกันการหารด้วยศูนย์

2.9 Support Vector Machine (SVM)

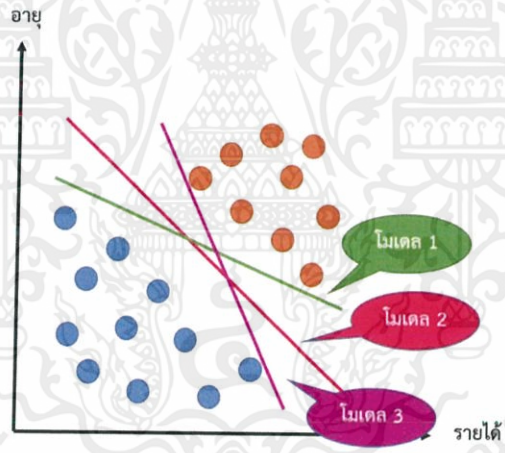
Support Vector Machine (SVM) [11] คืออัลกอริทึมการเรียนรู้ที่ใช้ในการจำแนกประเภท โดยในอัลกอริทึมนี้จะคำนวณจุดข้อมูลให้อยู่ในพื้นที่ n มิติ โดย n คือจำนวนคุณลักษณะของข้อมูล จากนั้นทำการจำแนกโดยการหาไฮเปอร์เพลน (hyperplane) ที่แยกความแตกต่างของข้อมูลได้ดีที่สุด ตัวอย่างเช่น ข้อมูลความสัมพันธ์ระหว่างอายุและรายได้ มีข้อมูล 2 คลาส คือคลาส A และคลาส B ดังรูปที่ 2.11

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.11 การกระจายตัวของข้อมูลคลาส A และ B [11]

โดยข้อมูลสามารถแบ่งออกเป็น 2 กลุ่มได้อย่างชัดเจนโดยใช้ไฮเปอร์เพลน (hyperplane) สำหรับข้อมูล 2 มิติซึ่งก็คือเส้นตรง โดยเลือกเส้นตรงได้หลากหลาย ซึ่งไฮเปอร์เพลนที่ดีที่สุด คือโมเดลที่จำแนกประเภทได้ดี แต่ไม่ขึ้นกับข้อมูลที่ใช้ในการเรียนรู้มากเกินไป ดังนั้นในตัวอย่างพบว่าโมเดลที่ 2 เป็นโมเดลที่ดีที่สุดดังรูปที่ 2.12



รูปที่ 2.12 โมเดลเส้นตรง [11]

อัลกอริทึม SVM แบ่งออกเป็น 2 แบบหลักๆ คือแบ่งแยกกลุ่มข้อมูลที่มีลักษณะเป็นเชิงเส้น และแบ่งแยกกลุ่มข้อมูลที่มีลักษณะไม่เป็นเชิงเส้น

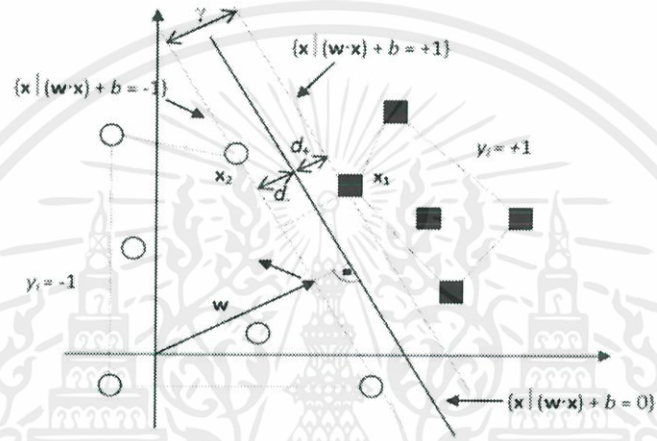
2.9.1 SVM แบ่งแยกกลุ่มข้อมูลแบบเชิงเส้น (Linear kernel)

หลักการของอัลกอริทึม SVM ซึ่งใช้การแบ่งแยกกลุ่มข้อมูลแบบเชิงเส้น คือการเลือกไฮเปอร์เพลนที่ให้ Margin หรือผลรวมระยะห่างของเส้นตรงที่เป็นไฮเปอร์เพลนถึงเส้นตรงที่ผ่านอินพุตที่ใกล้ที่สุดหรือซัพพอร์ตเวกเตอร์ (Support Vector) สูงสุด โดยในรูปที่ 2.13 ข้อมูลอินพุต $\mathbf{x}_1$ และ $\mathbf{x}_2$ คือซัพพอร์ตเวกเตอร์ ระยะ d+ และ d- คือระยะ Margin และสมการไฮเปอร์เพลนเป็นดังสมการที่ (2.6)

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned} \mathbf{w}^{\mathrm{T}}\mathbf{x}+\mathrm{b} &\geq +1; \mathrm{y}_{\mathrm{i}} = +1 \\ \mathbf{w}^{\mathrm{T}}\mathbf{x}+\mathrm{b} &\leq -1; \mathrm{y}_{\mathrm{i}} = -1 \\ \mathrm{y}_{\mathrm{i}}(\mathbf{w}^{\mathrm{T}}\mathbf{x}+\mathrm{b})-1 &\geq 0; \forall_{\mathrm{i}} \end{aligned} \tag{2.6}$$

โดย $\mathbf{x}_i$ คือข้อมูลอินพุต, $\mathbf{w}$ เป็นเวกเตอร์ที่ตั้งฉากกับไฮเปอร์เพลน, b เป็นค่าคงที่ซึ่งกำหนดตำแหน่งของเวคเตอร์ที่สัมพันธ์กับตำแหน่งดั้งเดิมใน input space และ $y = \{+1, -1\}$ โดย +1 คือข้อมูลคลาส 1 และ -1 คือข้อมูลคลาส 0 สำหรับข้อมูล 2 มิติ



รูปที่ 2.13 SVM แบบเชิงเส้น [12]

ในกรณีที่ไม่สามารถแยกแยะข้อมูลได้ถูกต้องทั้งหมด จะต้องมีการกำหนดตัวแปรเพื่อยอมรับค่าความผิดพลาด โดยทำการเพิ่มตัวแปร Slack Variable $\xi > 0$ ดังสมการที่ (2.7)

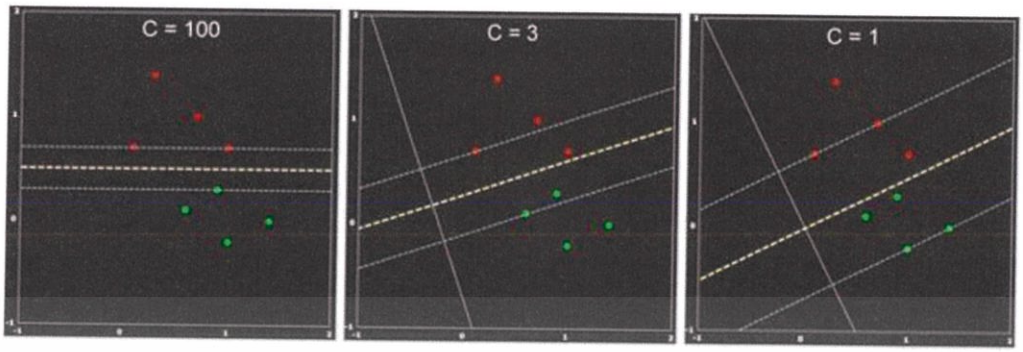
$$\begin{aligned} \mathbf{w}^{\mathrm{T}}\mathbf{x}+\mathrm{b} &\geq +1-\xi_{\mathrm{i}}; \mathrm{y}_{\mathrm{i}} = +1 \\ \mathbf{w}^{\mathrm{T}}\mathbf{x}+\mathrm{b} &\leq -1+\xi_{\mathrm{i}}; \mathrm{y}_{\mathrm{i}} = -1 \end{aligned} \tag{2.7}$$

พารามิเตอร์ที่สำคัญของ SVM แบบเชิงเส้นคือ C ดังสมการที่ (2.8) ซึ่งเป็นค่าตัวแปรที่ผู้ใช้สามารถกำหนดค่าได้เองเพื่อปรับความสมดุลระหว่างการเพิ่มระยะแยกแยะสูงสุด หรือค่าความผิดพลาดที่ต้องการให้ต่ำที่สุด โดย $\|\mathbf{w}\|$ แปรผกผันกับค่า d ของระยะ Margin ทำให้ได้สมการ optimization เป็น

$$\underset{\mathbf{w},\mathrm{b},\xi}{\mathrm{Minimize}}\left[\frac{1}{2}\|\mathbf{w}\|^2 + \mathrm{C}\sum_{\mathrm{i}=1}^{\mathrm{N}}\xi_{\mathrm{i}}\right] \tag{2.8}$$

โดยพารามิเตอร์ C แปรผกผันกับระยะ Margin นั่นคือ ถ้าค่า C น้อยจะทำให้ระยะ Margin มากและถ้าค่า C มากจะทำให้ระยะ Margin น้อยแสดงดังรูปที่ 2.14

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.14 ค่า C ที่มีผลต่อระยะ Margin [13]

2.9.2 SVM แบ่งแยกกลุ่มข้อมูลแบบไม่เป็นเชิงเส้น

SVM ที่ใช้การแบ่งแยกกลุ่มข้อมูลแบบไม่เป็นเชิงเส้น มีการแปลงกลุ่มข้อมูลตัวอย่างไปสู่ปริภูมิมิติที่สูงขึ้นซึ่งเรียกว่า feature space โดย kernel ของ SVM สำหรับแบ่งแยกกลุ่มข้อมูลแบบไม่เป็นเชิงเส้นที่ได้ศึกษามี 3 แบบดังนี้

2.9.2.1 Polynomial kernel

Polynomial kernel เป็นฟังก์ชันที่ใช้กับ SVM ซึ่งแสดงเวกเตอร์ที่ใช้ในการเรียนรู้ในปริภูมิมิติที่สูงขึ้นจากตัวแปรเดิม ทำให้สามารถเรียนรู้โมเดลที่ไม่เป็นเชิงเส้นได้ โดยสมการที่ใช้คำนวณไฮเปอร์เพลนของ Polynomial kernel เป็นดังสมการที่ (2.9)

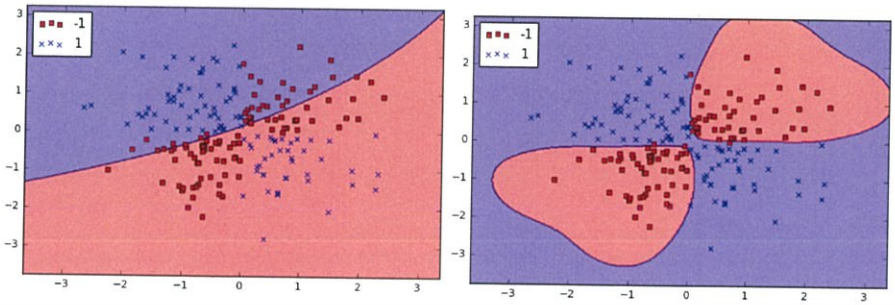
$$k(\mathbf{x}_i, \mathbf{x}_j) = (1 + \mathbf{x}_i \cdot \mathbf{x}_j)^d \tag{2.9}$$

2.9.2.2 Radial Basis Function kernel (RBF)

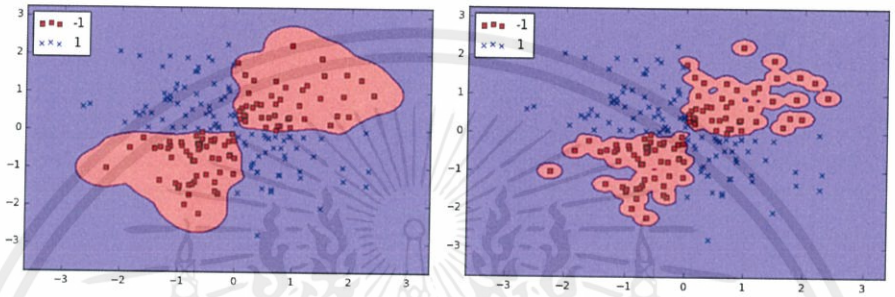
สมการที่ใช้คำนวณไฮเปอร์เพลนของ RBF kernel เป็นดังสมการที่ (2.10) โดยมีพารามิเตอร์ที่สำคัญคือ gamma (γ) ซึ่ง gamma เป็นส่วนของการตัดสินใจ คือถ้า gamma มีค่าน้อย จะทำให้เส้นโค้งของขอบการตัดสินใจต่ำ ทำให้ขอบเขตการตัดสินใจกว้าง และถ้า gamma มีค่ามาก จะทำให้เส้นโค้งของขอบการตัดสินใจสูง ซึ่งจะสร้างขอบเขตของการตัดสินใจรอบๆ จุดข้อมูลดังตัวอย่างในรูปที่ 2.15

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2) \tag{2.10}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก) (ข)



(ค) (ง)

รูปที่ 2.15 การเปลี่ยนแปลงของ gamma ที่มีผลต่อขอบการตัดสินใจ [14]
(ก) $\gamma = 0.01$ (ข) $\gamma = 1$ (ค) $\gamma = 10$ (ง) $\gamma = 100$

2.9.2.3 Sigmoid kernel

Sigmoid kernel [15] เป็นไปตามทฤษฎีบทของ Mercer ที่กำหนดว่า kernel ต้องเป็นบวกเสมอ Sigmoid kernel ยังไม่เป็นที่ยอมรับแน่นอนสำหรับค่าพารามิเตอร์บางค่า ดังนั้นต้องเลือกพารามิเตอร์ที่เหมาะสม มิเช่นนั้นผลลัพธ์อาจผิดพลาด โดยสมการที่ใช้คำนวณไฮเปอร์เพลนของ Sigmoid kernel เป็นดังสมการที่ (2.11)

$$k(\mathbf{x}_i, \mathbf{x}_j) = \tanh(k\mathbf{x}_i \cdot \mathbf{x}_j) - \delta \qquad (2.11)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

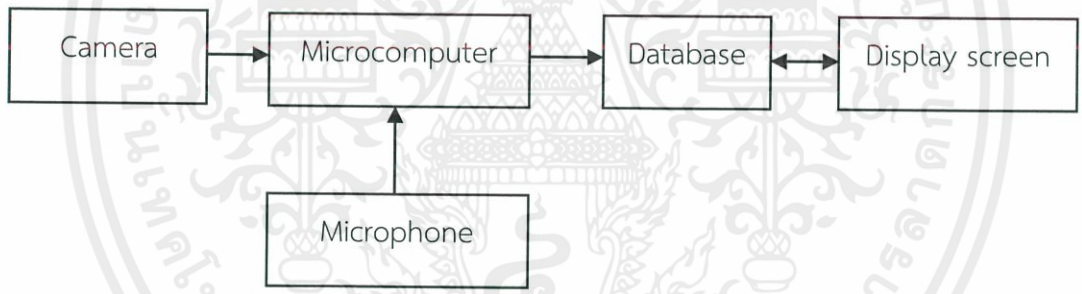
บทที่ 3

การออกแบบและการจัดทำปริญญานิพนธ์

3.1 การออกแบบ

3.1.1 การออกแบบการทำงานของระบบ

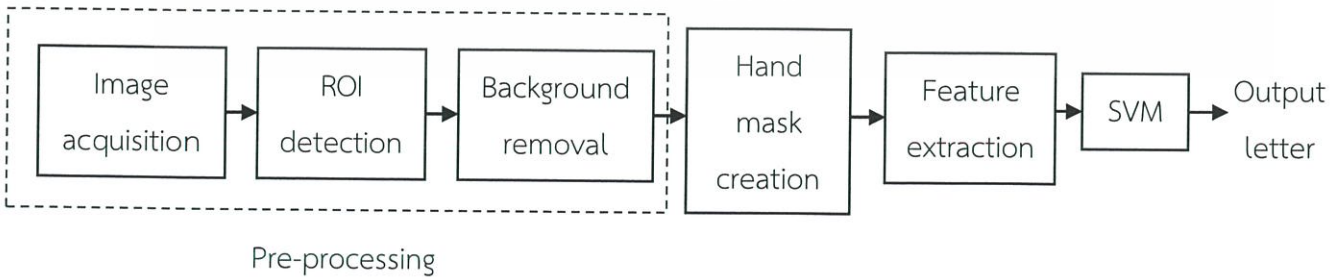
อุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยินได้ถูกออกแบบเพื่อให้ผู้พิการสามารถสนทนาโดยใช้ภาษามือกับคนปกติได้ โดยระบบประกอบไปด้วย 3 ส่วน คือการรับภาษามือจากผู้พิการทางการได้ยินผ่านกล้อง โดยภาษามือที่ใช้ คือภาษามือแบบอเมริกัน เนื่องจากภาษามือแต่ละประเทศมีท่าทางที่ไม่เหมือนกันจึงเลือกใช้การทำท่าทางภาษามือทีละตัวอักษรเป็นภาษาอังกฤษซึ่งเป็นภาษาสากล และคนส่วนมากใช้ภาษาอังกฤษในการสื่อสาร ส่วนที่สองคือการตอบกลับโดยรับเสียงคู่สนทนาที่เป็นคนปกติจากไมโครโฟน ซึ่งทั้งสองส่วนประมวลผลโดยใช้ไมโครคอมพิวเตอร์ และส่วนสุดท้ายคือฐานข้อมูลที่เก็บประวัติการสนทนาและแสดงผลผ่านจอแสดงผล โดยบล็อกไดอะแกรมของระบบแสดงดังรูปที่ 3.1



รูปที่ 3.1 บล็อกไดอะแกรมของอุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยิน

3.1.2 การออกแบบส่วนประมวลผลภาพของโปรแกรมแปลภาษามือเป็นตัวอักษร

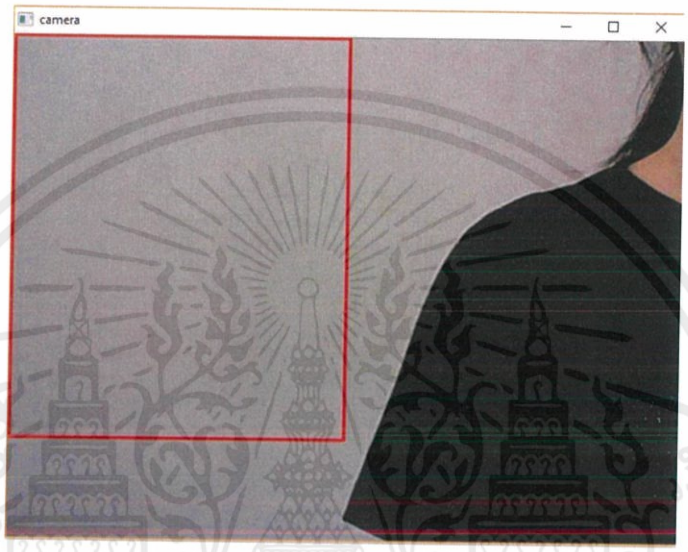
การประมวลผลภาพภาษามือจากกล้องมีขั้นตอนดังแสดงในรูปที่ 3.2 [16]



รูปที่ 3.2 ขั้นตอนการประมวลผลภาพภาษามือ

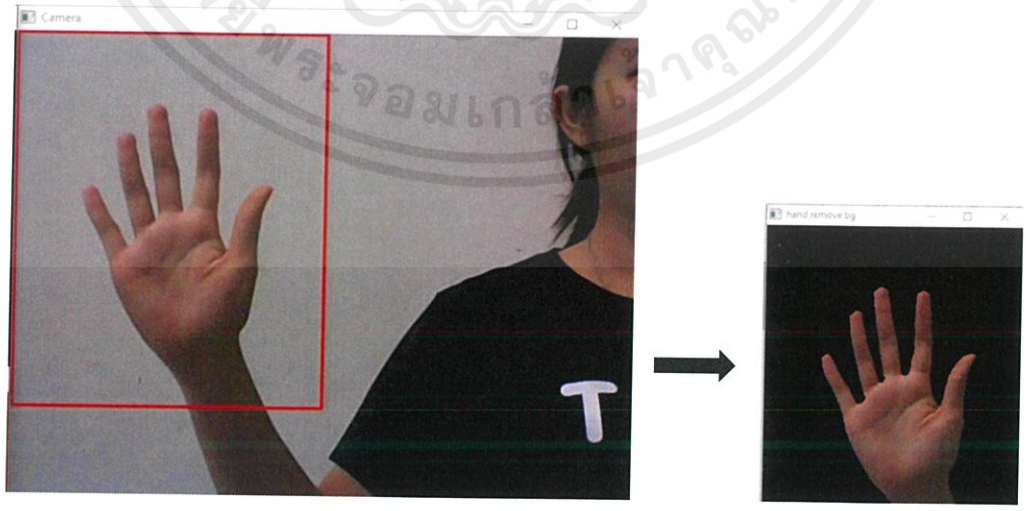
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) การเตรียมภาพ (Pre-processing) ประกอบด้วย การรับภาพจากกล้องเว็บแคม, การสร้างกรอบที่สนใจ (Region of interest หรือ ROI) เพื่อแสดงท่าทางภาษามือ และการลบภาพพื้นหลังเพื่อสะดวกต่อการประมวลผลภาพ โดยภาพที่ได้จากเว็บแคมมีขนาด 384 x 320 พิกเซล อยู่ในโหมดสี RGB จากนั้นทำการแปลงให้อยู่ในโหมดสี HSV การสร้างกรอบที่สนใจแสดงดังรูปที่ 3.3 ซึ่งการแสดงท่าทางภาษามือต้องแสดงในกรอบที่สนใจ



รูปที่ 3.3 การรับภาพจากกล้องเว็บแคมและสร้างกรอบที่สนใจ (สีแดง)

จากนั้นทำการลบพื้นหลังโดยการนำภาพพื้นหลังจาก ROI ในรูปที่ 3.3 มาหักลบออกจากภาพใหม่ เมื่อแสดงท่าทางภาษามือภายในกรอบที่สนใจจะได้ภาพที่ถูกลบพื้นหลังดังรูปที่ 3.4

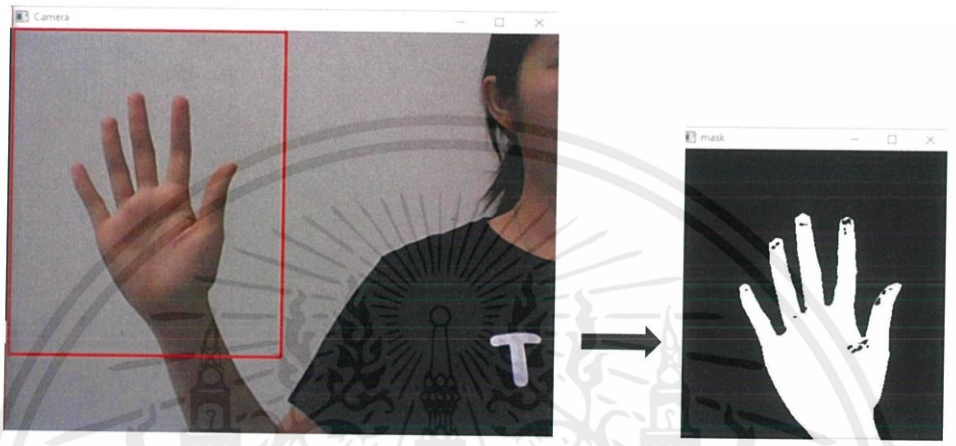


รูปที่ 3.4 ภาพที่ถูกลบพื้นหลัง

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) การค้นหาภาพภาษามือ (Hand mask creation)

จากนั้นทำให้สีผิวบริเวณมือเป็นสีขาวแสดงในหน้าต่าง mask ดังรูปที่ 3.5 โดยเริ่มจากการค้นหา pixel ที่มีสีเท่ากับสีผิว คือช่วงที่สีของภาพมีค่าเป็น [0, 130, 101] ถึง [255, 170, 133] จากนั้นแปลงค่าสีที่ตำแหน่ง pixel นั้นๆ ให้เป็นสีขาว



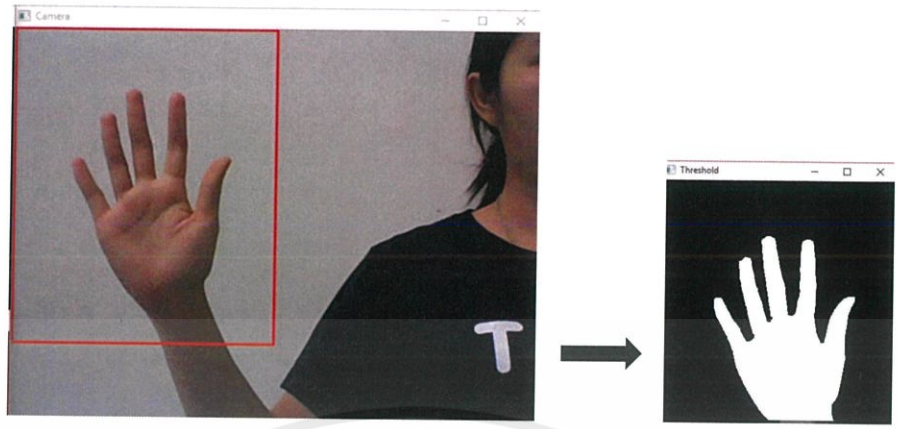
รูปที่ 3.5 การทำให้บริเวณสีผิวเป็นสีขาว

จัดการภาพมือให้สมบูรณ์ขึ้นดังแสดงในหน้าต่าง im_out ในรูปที่ 3.6 โดยทำการหาจุดสีดำเล็กๆ ในมือ จากนั้นทำการเติมสีขาวลงที่จุดสีดำนั้น

รูปที่ 3.6 การลบจุดสีดำ

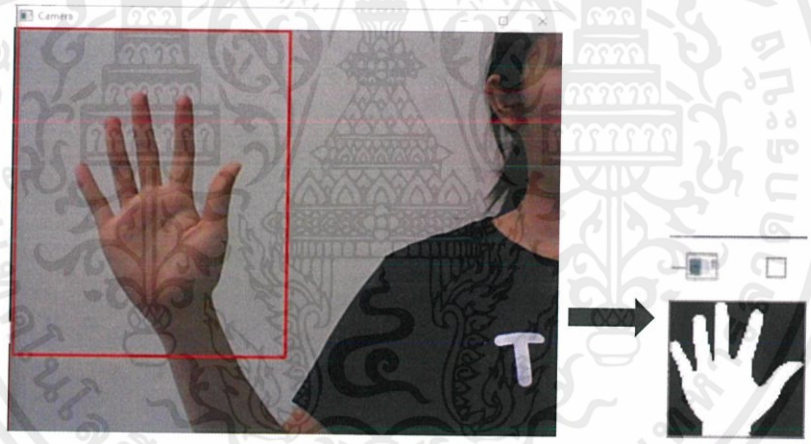
ทำการสร้างเกณฑ์ให้รูปภาพเป็นรูปขาวดำโดยให้ตำแหน่ง pixel ที่อยู่ในบริเวณมือมีค่าเท่ากับ 255 เป็นสีขาว ส่วนตำแหน่ง pixel อื่นให้มีค่าเท่ากับ 0 เป็นสีดำคือบริเวณพื้นหลัง และทำการกร่อนขนาด (Erosion) เพื่อลบค่า noise และทำการขยายภาพ (Dilation) เพื่อขยายมือทำให้การตรวจจับแม่นยำขึ้น แสดงในหน้าต่าง Threshold ดังรูปที่ 3.7

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.7 การสร้างเกณฑ์ให้รูปภาพเป็นรูปขาวดำ

จากนั้นทำการหาพื้นที่บริเวณขอบมือ และทำการปรับขนาดภาพให้บริเวณขอบมือพอดีกับขนาด 64x64 พิกเซล ดังรูปที่ 3.8 เพื่อเพิ่มความแม่นยำ เนื่องจากมือไม่ได้อยู่กึ่งกลางทุกรูปอาจทำให้การทำนายผิดพลาด



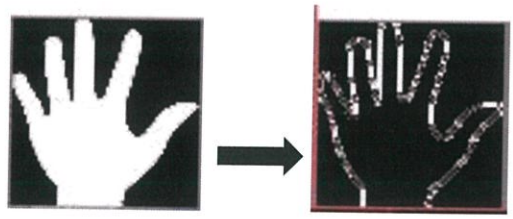
รูปที่ 3.8 ปรับขนาดภาพให้บริเวณขอบมือพอดีกับขนาด 64x64 พิกเซล

3) การหาคุณลักษณะของภาพ (Feature extraction)

Histogram of gradients (HOG) ถูกเลือกมาใช้เป็นคุณลักษณะของภาพ โดยมีขั้นตอนการคำนวณดังนี้

A. ทำการหาเกรเดียนท์ของภาพดังสมการที่ (2.1) และ (2.2) ขนาดของเกรเดียนท์แสดงดังรูปที่ 3.9

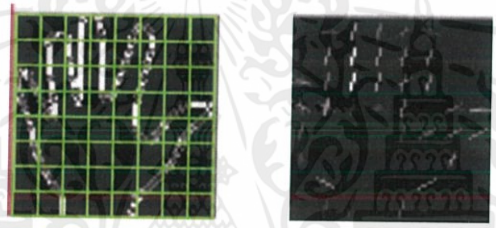
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.9 การหาเกรเดียนท์ของภาพ

B. กำหนดขนาดของภาพย่อย (Block) เท่ากับ 8x8 ดังรูปที่ 3.10 (ก)

C. สร้างฮิสโทแกรมของ gradient orientation ในแต่ละ bin โดยใช้ n=9 orientation bins กำหนด cells per block เท่ากับ 2x2 ดังสมการที่ (2.3) และ (2.4) แล้วทำการนอร์มัลไลซ์ (Normalize) ด้วย L1-norm โดยใช้สมการที่ (2.5) ซึ่งจะได้ผลลัพธ์ดังรูปที่ 3.10 (ข)



(ก) (ข)

รูปที่ 3.10 ขั้นตอนการหา HOG

(ก) รูปภาพที่ถูกกำหนดขนาด block เท่ากับ 8x8 (ข) รูปภาพ HOG หลังผ่านการนอร์มัลไลซ์

4) ทำนายรูปภาพกับข้อมูลที่ถูกเรียนรู้ด้วยอัลกอริทึม SVM

ผลลัพธ์การทำนายรูปภาพกับข้อมูลที่ถูกเรียนรู้ด้วยอัลกอริทึม SVM จะแสดงเป็นตัวอักษรหลังคำว่า Predicted text: โดยหากระบบตัดสินใจภาษามือได้เป็นตัวอักษรเดิมในเวลา 5 วินาที ตัวอักษรนั้นจะปรากฏในบรรทัดถัดไปและเรียงต่อกันเป็นประโยค ดังรูปที่ 3.11

รูปที่ 3.11 การใช้ผลลัพธ์การทำนายรูปภาพกับข้อมูลที่ถูกเรียนรู้ด้วยอัลกอริทึม SVM ในการสนทนา

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.3 การออกแบบการทำงานของอัลกอริทึม SVM สำหรับการเรียนรู้ของโปรแกรมแปลภาษามือเป็นตัวอักษร

เริ่มต้นโดยทำการสร้างชุดข้อมูลสำหรับการเรียนรู้ของโปรแกรมแปลภาษามือเป็นตัวอักษร โดยทำการรับภาพจากฐานข้อมูล [17][18] ที่มีรูปภาพภาษามือแต่ละตัวอักษรเข้ามาเก็บไว้เป็นอาร์เรย์ และทำการสร้างสัญลักษณ์ประจำรูป เช่นรูปภาพภาษามือตัวอักษร A มีสัญลักษณ์ A และรูปภาพภาษามือตัวอักษร B มีสัญลักษณ์ B เป็นต้น จากนั้นบันทึกข้อมูลสำหรับการเรียนรู้ของโปรแกรมแปลภาษามือเป็นไฟล์ .npy

ทำการแบ่งชุดข้อมูลเพื่อใช้ในการเรียนรู้และทดสอบ ซึ่งจะใช้รูปภาพภาษามือตัวอักษร A ถึง Z รวมถึงรูปภาพภาษามือเครื่องหมายจุลภาค (,) เครื่องหมายมหัพภาค (.) เครื่องหมายปรัศนี (?) เครื่องหมายเว้นวรรค ภาษามือสำหรับลบตัวอักษร และภาษามือแสดงการส่งข้อความดังแสดงในรูปที่ 2.1 จำนวนรูปภาพที่ใช้ในการเรียนรู้แต่ละตัวอักษรแสดงดังตารางที่ 3.1 และจำนวนรูปภาพที่ใช้ในการทดสอบแต่ละตัวอักษรคือ 100 รูป

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 จำนวนรูปภาพที่ใช้ในการเรียนรู้

ตัวอักษร	จำนวนรูปภาพที่ใช้ในการเรียนรู้
a	1222
b	1391
c	1077
d	1586
e	1200
f	1409
g	1325
h	1558
i	1523
j	1225
k	1466
l	1378
m	1124
n	1124
o	1382
p	1222
q	1298
r	1012
s	1129
t	1120
u	1105

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 จำนวนรูปภาพที่ใช้ในการเรียนรู้ (ต่อ)

ตัวอักษร	จำนวนรูปภาพที่ใช้ในการเรียนรู้
v	1120
w	1089
x	1257
y	1092
z	1338
เครื่องหมายจุลภาค (,)	1656
เครื่องหมายมหัพภาค (.)	1280
เครื่องหมายปรัศนี (?)	1322
เครื่องหมายเว้นวรรค	1213
สัญลักษณ์สำหรับลบตัวอักษร	1255
สัญลักษณ์แสดงการส่งข้อความ	1133

จากนั้นนำชุดข้อมูลที่ใช้ในการเรียนรู้และการทดสอบมาทำการทดสอบพารามิเตอร์ของอัลกอริทึม SVM แบบ multi-class โดยในปริญญานิพนธ์นี้ทำการทดสอบพารามิเตอร์ในแบบเชิงเส้น (Linear kernel) และแบบไม่เป็นเชิงเส้นโดยใช้ RBF kernel โดยการทดสอบพารามิเตอร์แบบเชิงเส้นกำหนดพารามิเตอร์ C เพื่อใช้ในการทดสอบ และการทดสอบพารามิเตอร์แบบไม่เป็นเชิงเส้น กำหนดพารามิเตอร์ C และ gamma เพื่อใช้ในการทดสอบ โดยทั้ง 2 แบบใช้วิธี SVM ในการจำแนกข้อมูลเพื่อหาค่า f1-score [19] ดังสมการที่ (3.1) ซึ่งเป็นค่าเฉลี่ยฮาร์โมนิคที่ถ่วงน้ำหนักระหว่าง precision และ recall ซึ่ง f1-score ที่ดีที่สุดเท่ากับ 1 จากนั้นนำพารามิเตอร์ที่ดีที่สุดมาใช้เรียนรู้และบันทึกเป็นไฟล์ .xml เพื่อนำไปทำนายในโปรแกรมแปลภาษามือเป็นตัวอักษร

$$\text{f1-score} = 2\frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{3.1}$$

โดย precision คือความแม่นยำของอัตราส่วนดังสมการที่ (3.2)

$$\text{precision} = \frac{t_p}{t_p + f_p} \tag{3.2}$$

recall คือการเรียกคืนของอัตราส่วนดังสมการที่ (3.3)

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\text{recall} = \frac{t_p}{t_p + f_n} \tag{3.3}$$

เมื่อ t_p คือ จำนวนข้อมูลที่ถูกจำแนกอย่างถูกต้องของคลาส 1 (true positive)

f_p คือ จำนวนข้อมูลคลาส 0 ที่ถูกจำแนกว่าเป็นคลาส 1 (false positive)

f_n คือ จำนวนข้อมูลคลาส 1 ที่ถูกจำแนกว่าเป็นคลาส 0 (false negative)

t_n คือ จำนวนข้อมูลที่ถูกจำแนกอย่างถูกต้องของคลาส 0 (true negative)

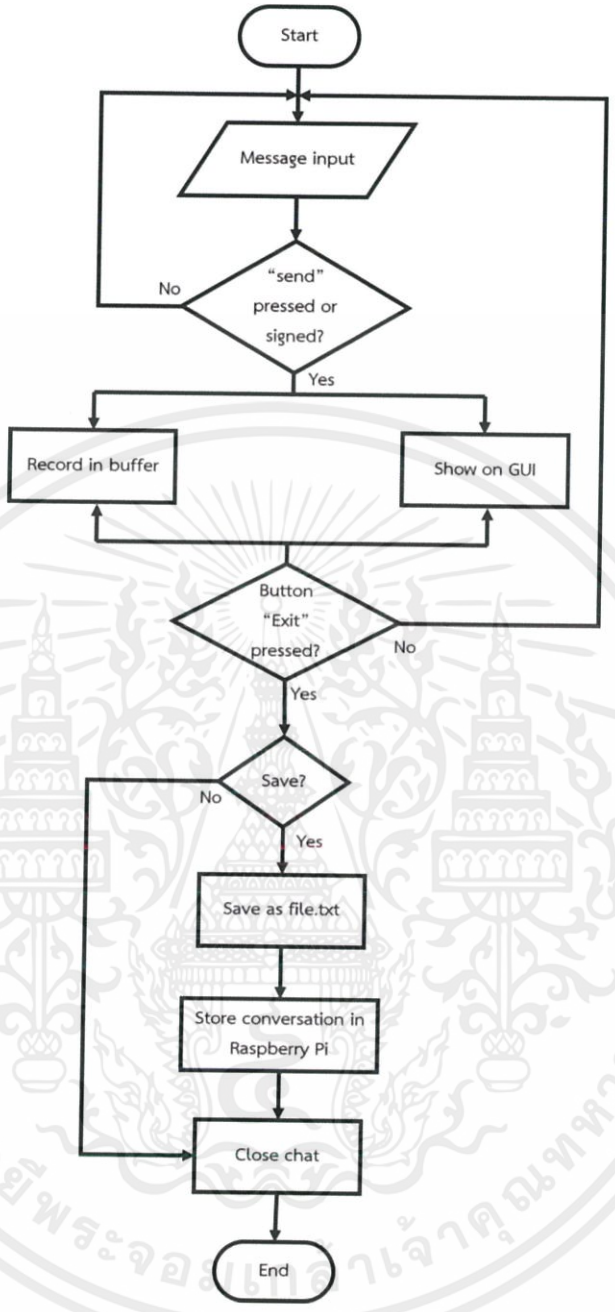
และหา Accuracy ดังสมการที่ (3.4) ซึ่ง Accuracy คือค่าระดับความเที่ยงตรงเมื่อเทียบกับเป้าหมาย ซึ่ง Accuracy ที่ดีที่สุดเท่ากับ 1

$$\text{accuracy} = \frac{t_p + t_n}{t_p + f_p + f_n + t_n} \tag{3.4}$$

3.1.4 การออกแบบการทำงานของระบบบันทึกข้อความสนทนาจากผู้พิการทางการได้ยิน

แผนผังการทำงานของระบบบันทึกข้อความสนทนาจากผู้พิการทางการได้ยินแสดงในรูปที่ 3.12 เริ่มต้นจากการรับค่าข้อความเข้ามาจากโปรแกรมแปลภาษามือเป็นตัวอักษร จากนั้นหากต้องการส่งข้อความให้แสดงภาษามือหรือกดปุ่ม send ข้อความที่ส่งจะถูกเก็บไว้ใน buffer และแสดงบนส่วนต่อประสานกราฟิกกับผู้ใช้ เมื่อกดปุ่ม Exit เพื่อสิ้นสุดการสนทนาสามารถเลือกบันทึกหรือไม่บันทึกข้อความการสนทนาได้ หากเลือกบันทึกข้อความการสนทนา ข้อความจะถูกบันทึกเป็นไฟล์ .txt แล้วเก็บไว้ใน Raspberry Pi หากไม่บันทึกข้อความจะทำการปิดหน้าต่างการสนทนา โดยประวัติการสนทนาจะแสดงค้างไว้ที่หน้าต่างสนทนา และสามารถดูได้จากไฟล์ .txt

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

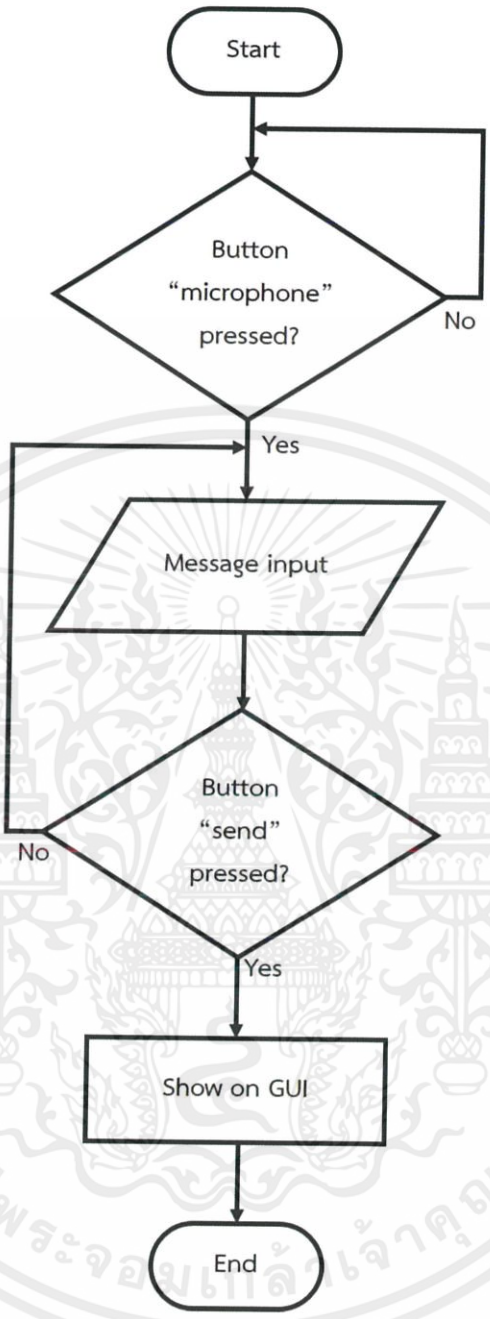


รูปที่ 3.12 แผนผังการทำงานของระบบบันทึกข้อความสนทนา

3.1.5 การออกแบบการทำงานของระบบบันทึกข้อความสนทนาจากคนปกติ

แผนผังการทำงานของระบบบันทึกข้อความสนทนาจากคนปกติแสดงในรูปที่ 3.13 เริ่มต้นจากการกดปุ่มรูปไมโครโฟนเพื่อรับเสียง จากนั้นตัวอักษรที่ได้จากเสียงจะปรากฏบนส่วนต่อประสานกราฟิกกับผู้ใช้ หากต้องการส่งข้อความให้กดปุ่ม send ข้อความที่ส่งจะแสดงบนช่องการสนทนาของส่วนต่อประสานกราฟิกกับผู้ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

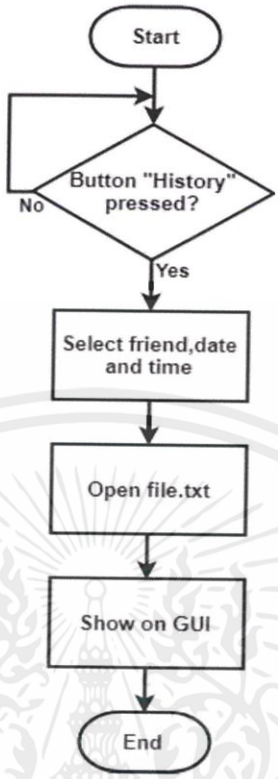


รูปที่ 3.13 แผนผังการทำงานของระบบบันทึกข้อความสนทนาจากคนปกติ

3.1.6 การออกแบบการทำงานของระบบเรียกดูประวัติการสนทนา

การทำงานของระบบเรียกดูประวัติการสนทนา แสดงดังรูปที่ 3.14 เริ่มต้นจากตรวจสอบว่ามีการเรียกดูประวัติการสนทนาหรือไม่ หากมีการเรียกดูประวัติการสนทนาจะมีให้ระบุชื่อผู้ติดต่อ วันที่ และเวลาของการสนทนา และเลือกเปิดไฟล์ตามที่ระบุไว้ข้างต้นเพื่อนำไปแสดงผลบนส่วนต่อประสานกราฟิกกับผู้ใช้

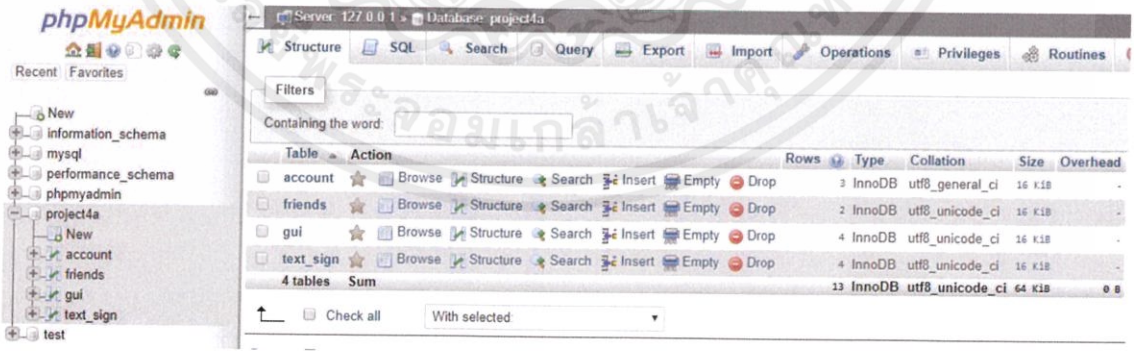
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.14 แผนผังการทำงานของระบบเรียกดูประวัติการสนทนา

3.1.7 การออกแบบตารางฐานข้อมูลในการจัดเก็บข้อมูล

ออกแบบฐานข้อมูลผ่านโปรแกรม phpMyAdmin โดยตั้งชื่อฐานข้อมูลว่า project4a มีตารางฐานข้อมูลจำนวน 4 ตารางคือ account, friends, gui และ text_sign ดังรูปที่ 3.15

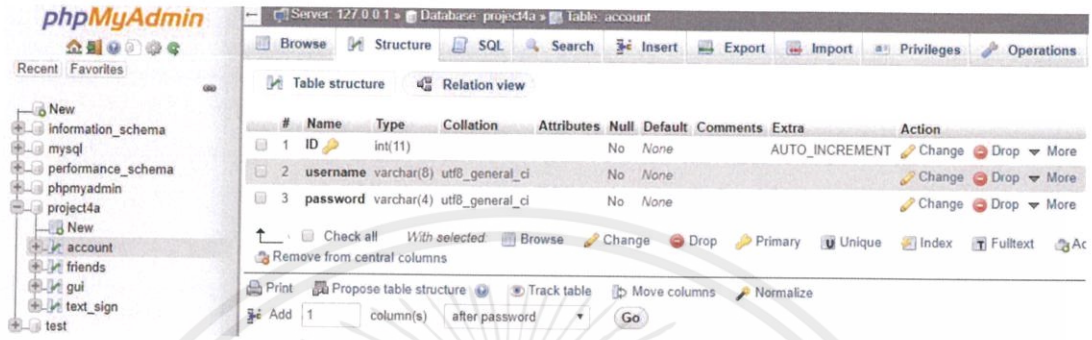


รูปที่ 3.15 การสร้างฐานข้อมูล project4a

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.7.1 ตาราง account

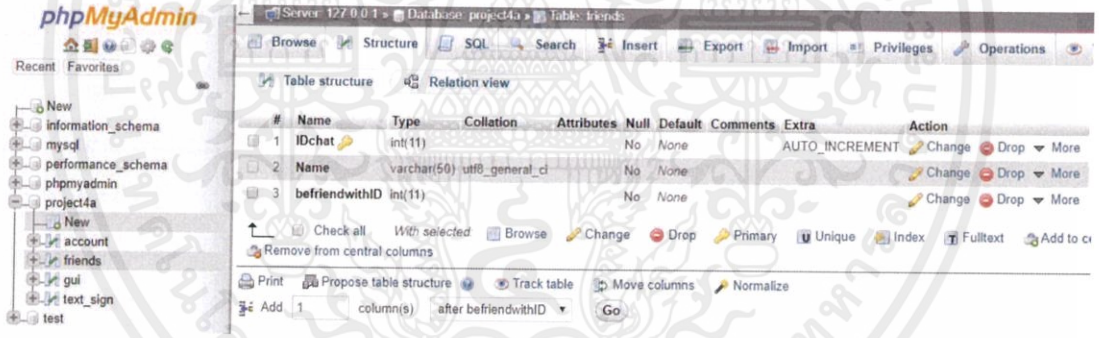
ตาราง account เป็นฐานข้อมูลที่เก็บข้อมูลที่ใช้ในการ Login เข้าสู่ระบบการสนทนา ประกอบด้วย 3 คอลัมน์ได้แก่ ID, username และ password ดังรูปที่ 3.16



รูปที่ 3.16 ตาราง account

3.1.7.2 ตาราง friends

ตาราง friends เป็นฐานข้อมูลที่เก็บรายชื่อผู้ที่ติดต่อกับผู้พิการทางการได้ยิน ประกอบด้วย 3 คอลัมน์ได้แก่ IDchat, Name และ befriendwithID ดังรูปที่ 3.17

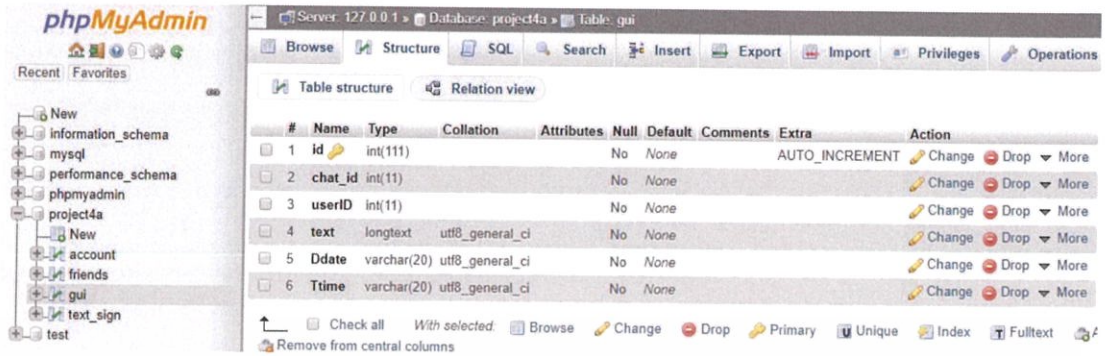


รูปที่ 3.17 ตาราง friends

3.1.7.3 ตาราง gui

ตาราง gui เป็นตารางในฐานข้อมูลที่บันทึกข้อความที่แปลจากภาษามือและแปลจากเสียงที่แสดงบนหน้าต่าง GUI, รหัสผู้ที่สนทนากับผู้พิการทางการได้ยิน, รหัสของผู้พิการทางการได้ยิน, วัน และเวลา ประกอบด้วย 6 คอลัมน์ได้แก่ id, chat_id, userID, text, Ddate และ Ttime ดังรูปที่ 3.18

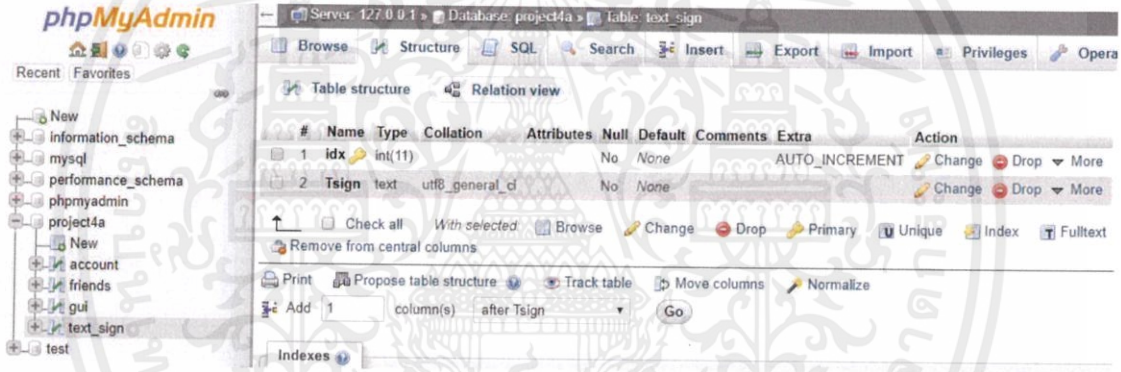
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.18 ตาราง gui

3.1.7.4 ตาราง text_sign

ตาราง text_sign เป็นตารางในฐานข้อมูลที่เก็บข้อความที่ได้จากการแปลภาษามือ ประกอบด้วย 2 คอลัมน์ได้แก่ idx และ Tsign ดังรูปที่ 3.19

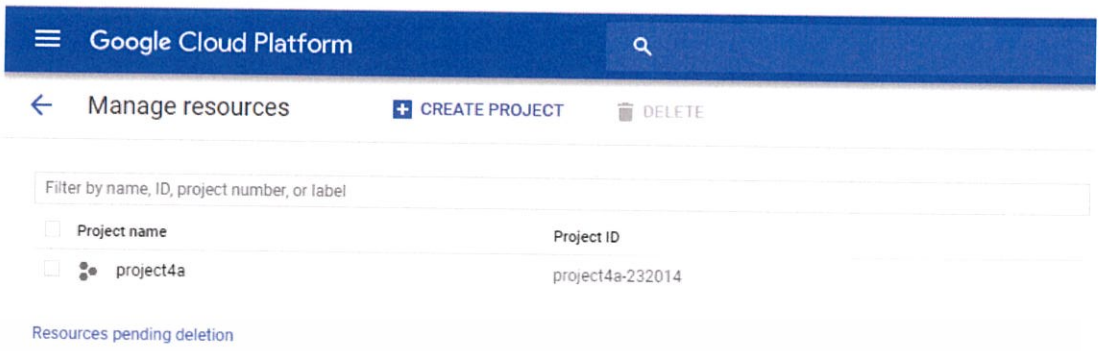


รูปที่ 3.19 ตาราง text_sign

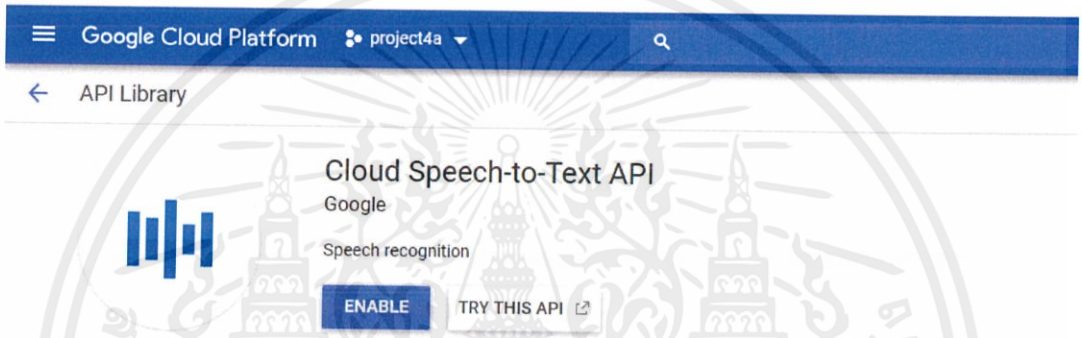
3.1.8 การแปลเสียงพูดเป็นข้อความ

ระบบการแปลเสียงพูดเป็นข้อความเริ่มจากสร้างโปรเจคบน Google Cloud Platform ดังรูปที่ 3.20 และเลือกใช้ Cloud Speech API [20] ดังรูปที่ 3.21 เรียกใช้งานผ่านการเขียนคำสั่งในภาษาไพธอน และใช้โมดูล Speech Recognition โดยการทำงานเริ่มจากรับเสียงพูดผ่านไมโครโฟน จากนั้นส่งไปประมวลผลบน Google Cloud Speech API และส่งผลลัพธ์กลับมาเป็นข้อความแสดงบนส่วนต่อประสานกราฟิกกับผู้ใช้ดังรูปที่ 3.22

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.20 โปรเจคที่สร้างบน Google Cloud Platform



รูปที่ 3.21 Cloud Speech API

รูปที่ 3.22 ข้อความที่แสดงบนส่วนต่อประสานกราฟิกกับผู้ใช้

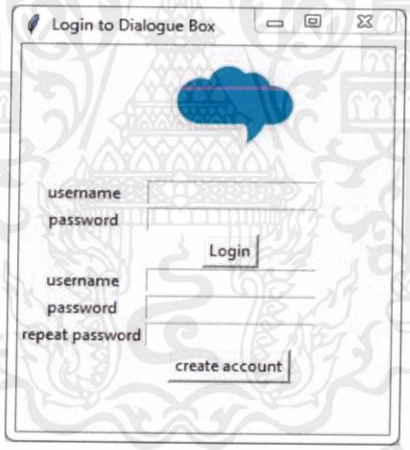
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.9 การออกแบบส่วนต่อประสานกราฟิกกับผู้ใช้ (GUI)

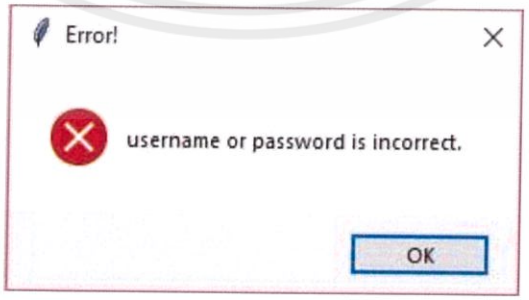
ส่วนต่อประสานกราฟิกกับผู้ใช้ประกอบด้วย หน้าต่าง Login และลงทะเบียน, หน้าต่างการสนทนา, หน้าต่างเพิ่มรายชื่อผู้ติดต่อใหม่, หน้าต่างออกจากการสนทนาและบันทึกข้อความการสนทนา และหน้าต่างแสดงประวัติการสนทนา

3.1.9.1 หน้าต่าง Login และลงทะเบียน

หน้าต่าง Login และลงทะเบียน แสดงดังรูปที่ 3.23 เป็นหน้าต่างสำหรับลงทะเบียนและลงชื่อเพื่อเข้าใช้งานหน้าต่างการสนทนา โดยจะบันทึกข้อมูล username และ password ไว้ในฐานข้อมูล เมื่อต้องการลงชื่อเข้าใช้งานหน้าต่างการสนทนา ผู้ใช้ต้องกรอกข้อมูล username และ password ให้ตรงตามฐานข้อมูลที่ถูกบันทึกไว้ในฐานข้อมูล แล้วกดปุ่ม Login หากเป็นผู้ใช้งานใหม่ต้องทำการสร้าง account ด้วยปุ่ม create account เสียก่อนจึงจะสามารถ Login ได้ หากผู้ใช้กรอกข้อมูล username หรือ password ไม่ตรงตามข้อมูลที่บันทึกไว้ในฐานข้อมูล ระบบจะทำการแจ้งเตือนข้อผิดพลาดนั้นว่า “username or password is incorrect.” ดังรูปที่ 3.24 และไม่สามารถเข้าใช้งานหน้าต่างการสนทนาได้



รูปที่ 3.23 หน้าต่าง Login และลงทะเบียน

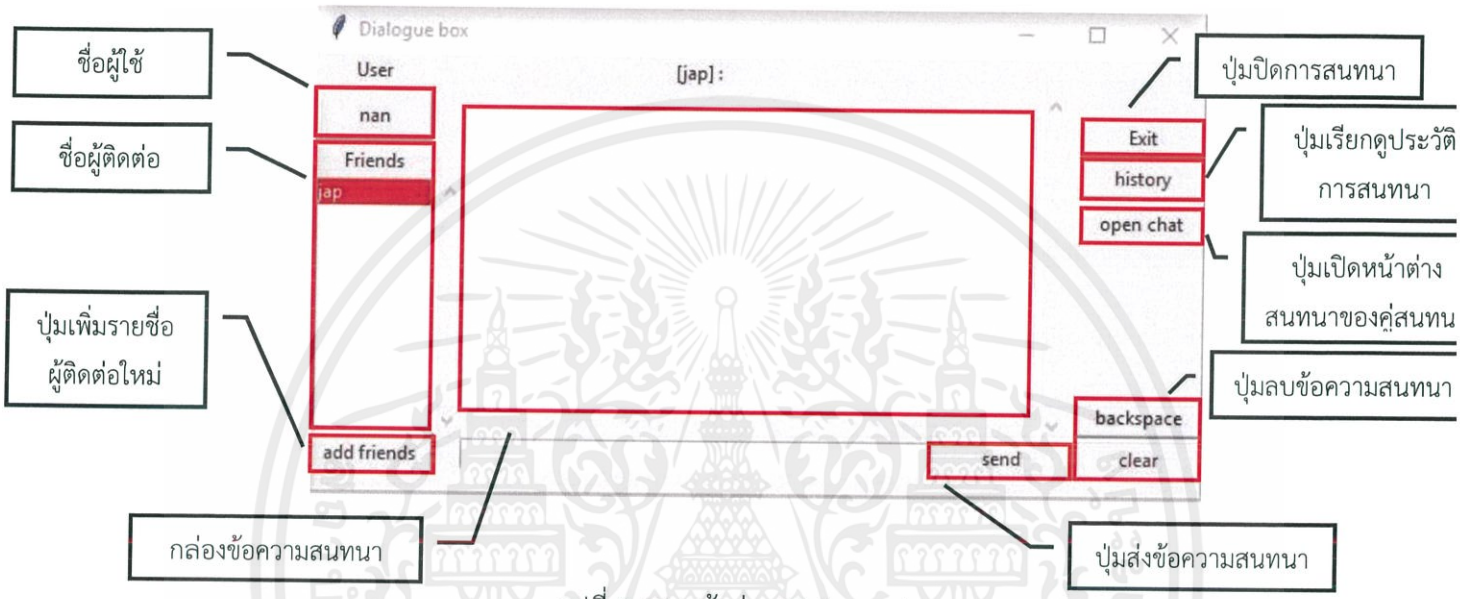


รูปที่ 3.24 การแจ้งเตือนเมื่อกรอก Username หรือ Password ไม่ถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.9.2 หน้าต่างการสนทนา (Dialogue box)

หน้าต่าง Dialogue box เป็นหน้าต่างหลักสำหรับการสนทนา ดังรูปที่ 3.25 ภายในหน้าต่างแสดงชื่อผู้ใช้ ชื่อผู้ติดต่อ กล่องข้อความสนทนา ปุ่มสำหรับส่งและลบข้อความสนทนา ปุ่มสำหรับเพิ่มรายชื่อผู้ติดต่อใหม่ ปุ่มสำหรับออกจากหน้าต่างการสนทนา ปุ่มสำหรับเรียกดูประวัติการสนทนา และปุ่มสำหรับเปิดหน้าต่างสนทนาของคนปกติ

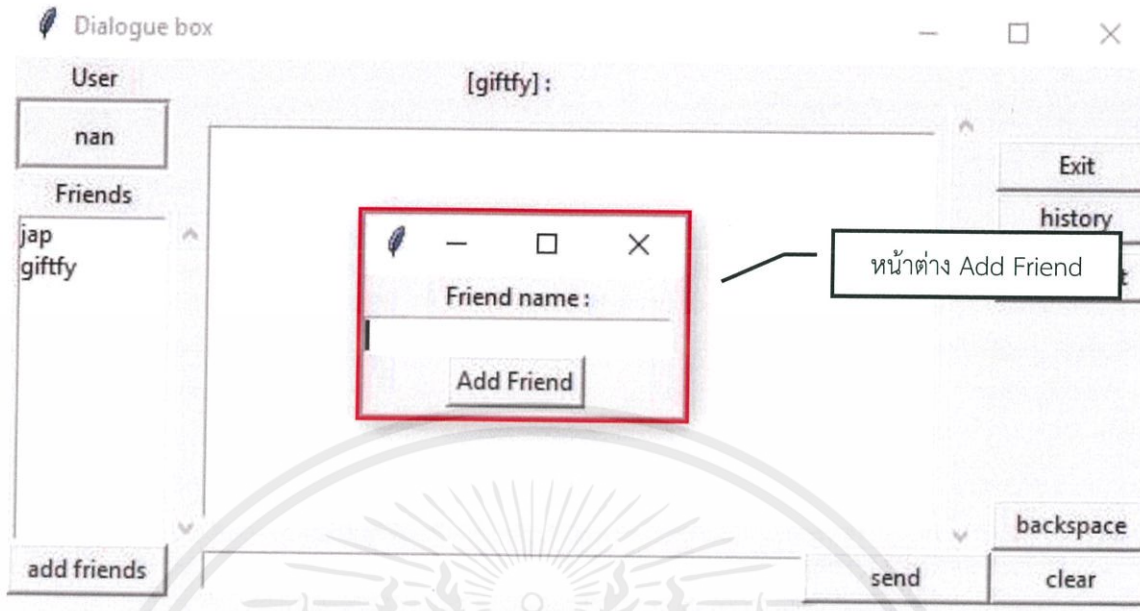


รูปที่ 3.25 หน้าต่าง Dialogue box

3.1.9.3 หน้าต่างเพิ่มรายชื่อผู้ติดต่อใหม่ (Add Friend)

หน้าต่าง Add Friend เป็นหน้าต่างสำหรับเพิ่มรายชื่อผู้ติดต่อใหม่ เมื่อกดปุ่ม add friends จะปรากฏหน้าต่างสำหรับเพิ่มชื่อผู้ติดต่อใหม่ดังรูปที่ 3.26ประกอบด้วยช่องสำหรับกรอกชื่อผู้ติดต่อใหม่ ปุ่ม Enter สำหรับยืนยันรายชื่อผู้ติดต่อใหม่เพื่อบันทึกรายชื่อผู้ติดต่อใหม่ลงในฐานข้อมูลและแสดงรายชื่อผู้ติดต่อในหน้าต่าง Dialogue box และปุ่ม Cancel สำหรับยกเลิกการเพิ่มรายชื่อผู้ติดต่อใหม่

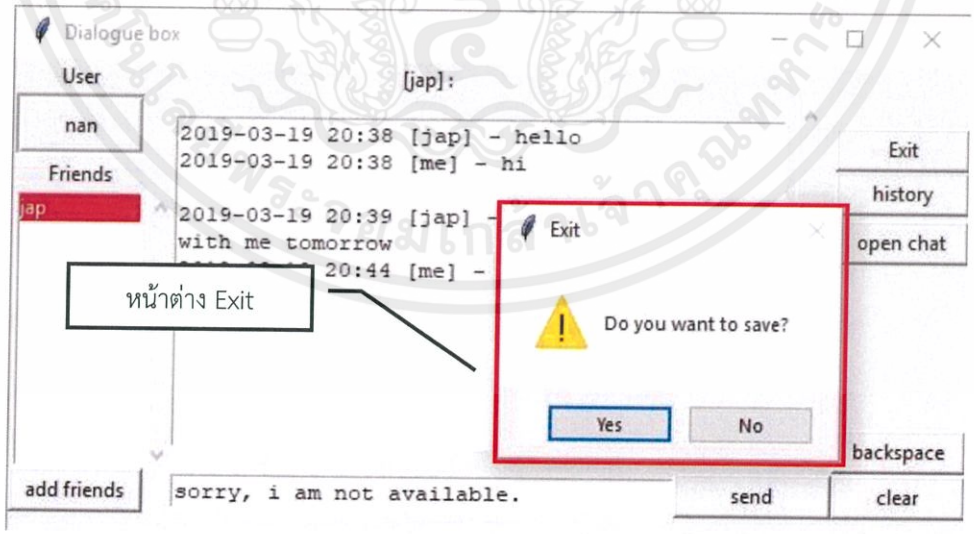
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.26 หน้าต่าง Add Friend

3.1.9.4 หน้าต่างออกจากการสนทนา (Exit)

หน้าต่าง Exit เป็นหน้าต่างสำหรับออกจากหน้าต่างการสนทนา ดังรูปที่ 3.27 เมื่อกดปุ่ม Exit จะปรากฏหน้าต่างให้เลือกบันทึกการสนทนาหรือไม่บันทึกการสนทนา หากเลือกบันทึกการสนทนาจะทำการบันทึกข้อมูลวันที่ เวลา ชื่อผู้ติดต่อ และข้อความการสนทนาลงในไฟล์ .txt ดังรูปที่ 3.28 และหากเลือกไม่บันทึกการสนทนาจะเป็นการปิดหน้าต่างการสนทนา



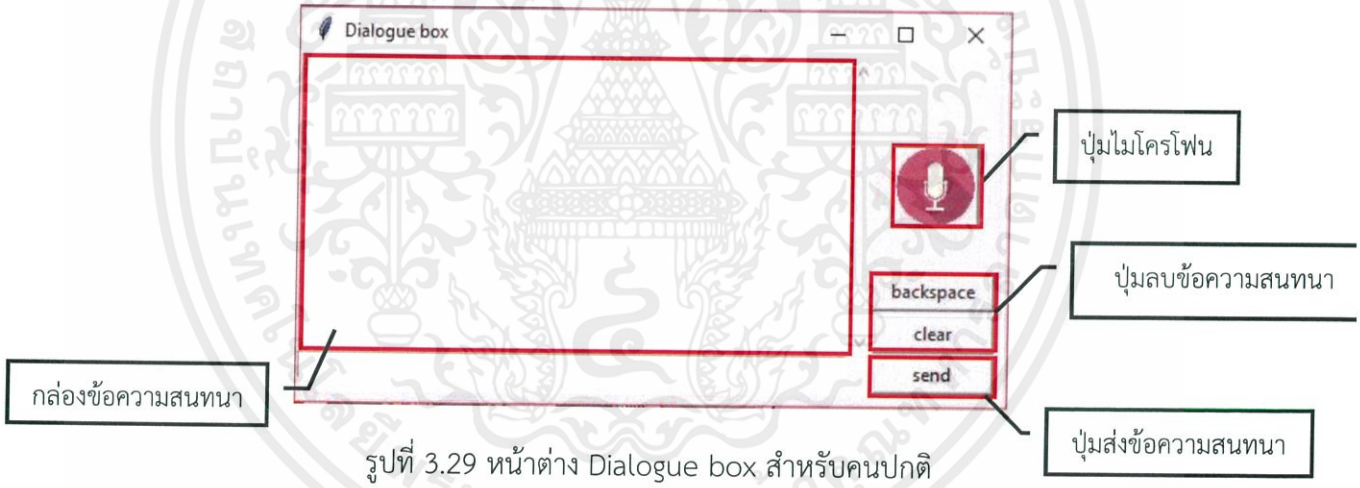
รูปที่ 3.27 หน้าต่าง Exit

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.28 ข้อความที่ถูกบันทึกในรูปไฟล์ .txt

3.1.9.5 หน้าต่าง Dialogue box สำหรับคู่สนทนาที่เป็นคนปกติ

หน้าต่าง Dialogue box สำหรับคนปกติดังรูปที่ 3.29 ภายในหน้าต่างแสดงปุ่มรูปไมโครโฟนโดยใช้กดเมื่อต้องการพูด กล่องข้อความสนทนา ปุ่มสำหรับส่ง และปุ่มสำหรับลบข้อความสนทนา

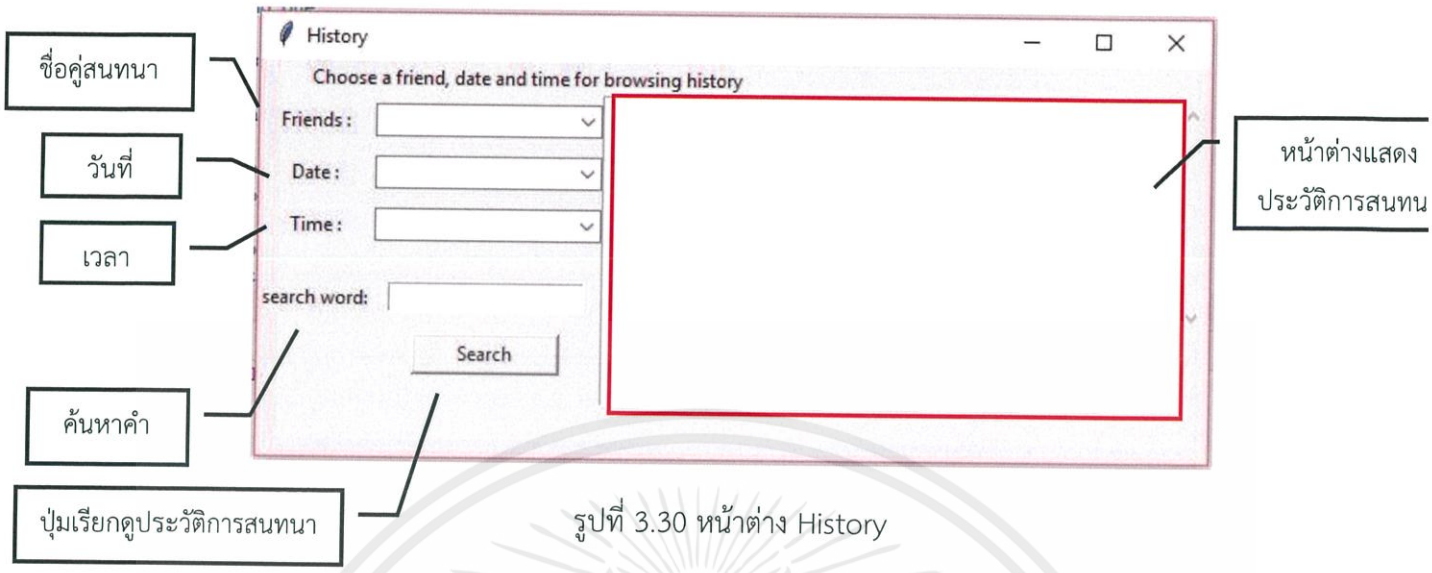


รูปที่ 3.29 หน้าต่าง Dialogue box สำหรับคนปกติ

3.1.9.6 หน้าต่างประวัติการสนทนา (History)

หน้าต่าง History ใช้สำหรับการเรียกดูประวัติการสนทนาดังรูปที่ 3.30 ภายในหน้าต่างแสดงประวัติการสนทนา ช่องสำหรับเลือกชื่อคู่สนทนา วันที่ เวลา ช่อง search word และปุ่ม Search สำหรับค้นหาคำ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



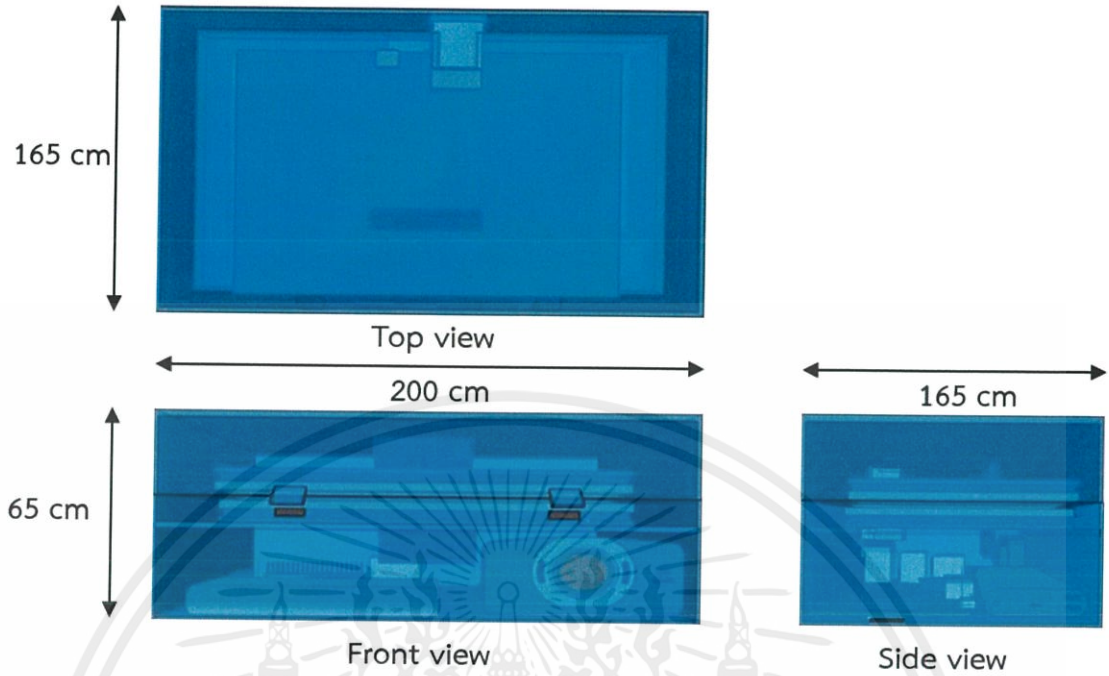
รูปที่ 3.30 หน้าต่าง History

3.1.10 การออกแบบกล่องอุปกรณ์

ผู้จัดทำได้ทำการออกแบบกล่องอุปกรณ์ ขนาดกว้าง 10 เซนติเมตร ยาว 18 เซนติเมตร สูง 6.9 เซนติเมตร ดังรูปที่ 3.31 โดยมีส่วนประกอบ 3 ส่วน หมายเลข 1 คือจอแสดงผลสำหรับผู้พิการทางการได้ยิน หมายเลข 2 คือจอแสดงผลสำหรับบุคคลปกติ ซึ่งสามารถนำออกมาตั้งข้างนอกได้เพื่อง่ายต่อการมองเห็น และหมายเลข 3 คือส่วนเก็บอุปกรณ์ได้แก่ Raspberry Pi ทำหน้าที่ประมวลผลภาพภาษามือและเสียงพูดเป็นตัวอักษร, กล้อง ทำหน้าที่รับภาพภาษามือจากผู้พิการทางการได้ยิน และแบตเตอรี่ โดยภาพฉายด้านหน้า ด้านข้าง และด้านบนของกล่องอุปกรณ์ต้นแบบแสดงดังรูปที่ 3.32 อุปกรณ์ที่จัดทำเสร็จสมบูรณ์แสดงในรูปที่ 3.33

รูปที่ 3.31 การออกแบบกล่องอุปกรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.32 ภาพฉายด้านหน้า ด้านข้าง และด้านบนของกล่องอุปกรณ์ต้นแบบ

รูปที่ 3.33 กล่องอุปกรณ์ที่จัดทำขึ้น

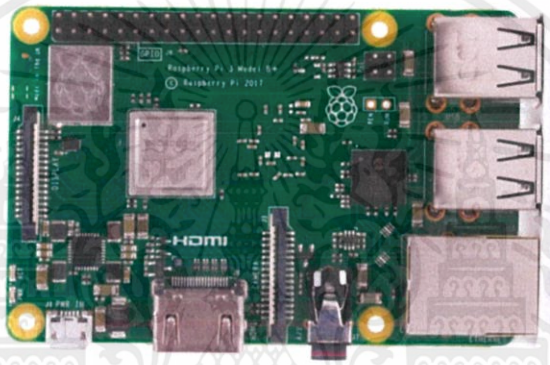
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 เครื่องมือที่ใช้ในการทดลอง

ในปริญญานิพนธ์นี้ มีอุปกรณ์และเครื่องมือที่ใช้ในการทดลอง ดังนี้

3.2.1 ไมโครคอมพิวเตอร์

ไมโครคอมพิวเตอร์ที่เลือกใช้ในปริญญานิพนธ์นี้คือ Raspberry Pi 3 Model B+ ดังรูปที่ 3.34 เป็นอุปกรณ์ควบคุมขนาดเล็ก สามารถต่อเข้ากับจอคอมพิวเตอร์ผ่านพอร์ต HDMI ต่อเมาส์และคีย์บอร์ดผ่านพอร์ต USB ได้ โดยในไมโครคอมพิวเตอร์มีส่วนซีพียู, หน่วยความจำ และพอร์ต ซึ่ง Raspberry Pi 3 Model B+ จะทำหน้าที่ประมวลผลกลางของระบบการทำงาน เชื่อมต่อฐานข้อมูลและอุปกรณ์ต่างๆ ที่ใช้ในปริญญานิพนธ์



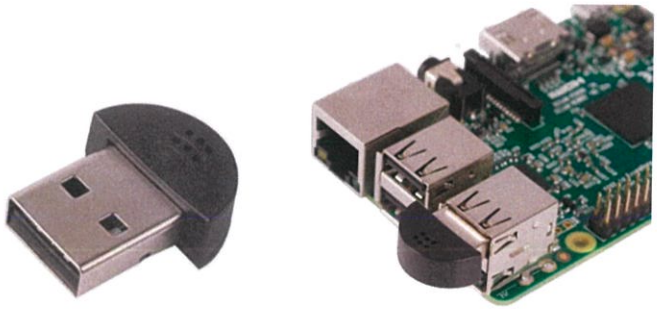
รูปที่ 3.34 Raspberry Pi 3 Model B+ [2]

3.2.2 ไมโครโฟน

มินิโมโครโฟน USB 2.0 ดังรูปที่ 3.35 เป็นไมโครโฟนขนาดเล็กสำหรับ PC, MAC หรือ Raspberry Pi ใช้การเชื่อมต่อแบบ USB ในปริญญานิพนธ์นี้ใช้สำหรับรับเสียงจากคนปกติเพื่อแปลเสียงพูดเป็นตัวอักษรโดยมีคุณสมบัติดังนี้

- ความไว -67 dBV/pBar.-47 dBV/Pascal+/-4dB
- การตอบสนองต่อความถี่ 100-16kHz
- ขนาด 2cm x 2cm x 0.5cm
- ไมโครโฟนมีความสามารถในการตัดเสียงรบกวน
- สวิตช์เพาเวอร์จะสว่างขึ้นเมื่อใช้งานไมโครโฟน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

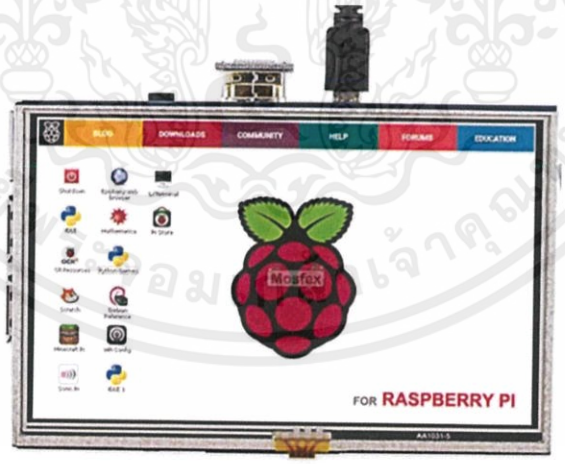


รูปที่ 3.35 มินิโมโครโฟน USB 2.0 [21]

3.2.3 จอแสดงผล

ปริญญานิพนธ์นี้ใช้จอแสดงผลขนาด 5 นิ้ว ดังรูปที่ 3.36 เพื่อใช้แสดงผลส่วนต่อประสานกราฟิกกับผู้ใช้ หน้าจอเป็นแบบ Touch Screen เชื่อมต่อผ่านทางพอร์ต HDMI ทำให้สามารถเชื่อมต่อกับ Raspberry Pi หรือระบบปฏิบัติการอื่นได้ ซึ่งมีคุณสมบัติดังนี้

- หน้าจอสัมผัสขนาด 5 นิ้ว TFT Resistive ความละเอียด 800x480 พิกเซล
- อินพุต HDMI
- แหล่งจ่ายไฟของ USB, 5V @ 1A
- IC ไดรเวอร์จอแอลซีดี: ILI9486L
- อัตราการรีเฟรช: 60HZ
- ขนาด LCD: 121 mm x 78 mm



รูปที่ 3.36 จอแสดงผลขนาด 5 นิ้ว [22]

3.2.4 กล้องเว็บแคม

ปริญญานิพนธ์นี้ใช้กล้องเว็บแคม Logitech C170 ดังรูปที่ 3.37 เพื่อรับภาพภาษามือจากผู้พิการทางการได้ยิน กล้องเว็บแคม C170 มีคุณสมบัติดังนี้

- ความละเอียดสูงสุด 720p/30fps

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ชนิดโฟกัสตายตัว
- พอร์ต USB
- เทคโนโลยีเลนส์มาตรฐาน
- ความยาวของสายเชื่อมต่อ 0.95 m
- มุมมอง 58°

รูปที่ 3.37 กล้องเว็บแคม Logitech C170 [23]

3.3 การจัดเก็บผลการทดลอง

3.3.1 การทดสอบข้อมูลสำหรับการเรียนรู้ของโปรแกรมแปลภาษามือ

ส่วนนี้เป็นการทดสอบข้อมูลสำหรับการเรียนรู้ของโปรแกรมแปลภาษามือ ซึ่งเป็นการทดสอบตำแหน่งการจัดเก็บรูปภาพ และสัญลักษณ์ประจำรูปภาพในอาร์เรย์

3.3.2 การทดสอบพารามิเตอร์สำหรับการเรียนรู้ของโปรแกรมแปลภาษามือเป็นตัวอักษร

3.3.2.1 ทดสอบพารามิเตอร์ของ Binary และ multi-class SVM แบบ Linear kernel

3.3.2.2 ทดสอบพารามิเตอร์ของ Binary SVM แบบ RBF kernel

3.3.3 ความแม่นยำในการจำแนกตัวอักษรที่ของ Linear SVM แบบ multi-class

ทดสอบค่า Accuracy ของพารามิเตอร์ที่เลือกใช้กับข้อมูลทีละตัวอักษร ที่ถูกกำหนดโดย Linear SVM แบบ multi-class

3.3.4 การทดสอบส่วนต่อประสานกราฟิกกับผู้ใช้และฐานข้อมูล

3.3.4.1 ทดสอบหน้าต่าง Login และลงทะเบียนเข้าใช้งานกับฐานข้อมูล

3.3.4.2 ทดสอบหน้าต่าง Add Friend กับฐานข้อมูล และการแสดงรายชื่อผู้ติดต่อในหน้าต่าง Dialogue box

3.3.4.3 ทดสอบหน้าต่าง Exit และการบันทึกข้อมูลการสนทนาเป็นไฟล์ .txt

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.4.4 การทดสอบหน้าต่าง History กับฐานข้อมูล

3.3.5 การทดสอบผลลัพธ์โปรแกรมแปลภาษามือเป็นตัวอักษร

3.3.5.1 ทดสอบผลลัพธ์การแปลภาษามือของตัวอักษร A-Z และเครื่องหมายวรรคตอนต่างๆ

3.3.5.2 ทดสอบการลบตัวอักษรด้วยภาษามือ

3.3.5.3 ทดสอบการส่งตัวอักษรด้วยภาษามือกับฐานข้อมูล

3.3.5.4 ทดสอบการแสดงผลลัพธ์โปรแกรมแปลภาษามือเป็นตัวอักษรบนส่วนต่อประสานกราฟิกกับผู้ใช้

3.3.6 การทดสอบผลลัพธ์โปรแกรมแปลเสียงเป็นตัวอักษร

ทดสอบผลลัพธ์โปรแกรมแปลเสียงเป็นตัวอักษรบนส่วนต่อประสานกราฟิกกับผู้ใช้

3.3.7 การทดสอบการทำงานของระบบรวม

3.3.7.1 ทดสอบการทำงานของระบบรวมบนคอมพิวเตอร์

3.3.7.2 ทดสอบการทำงานของระบบรวมบน Raspberry Pi

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

ผู้จัดทำได้ทำการเก็บผลการทำงานของระบบ โดยแบ่งการทดลองและจัดเก็บผลการทดลองเป็นส่วนๆ ดังต่อไปนี้

4.1 การทดสอบข้อมูลสำหรับการเรียนรู้ของโปรแกรมแปลภาษามือ

ตรวจสอบข้อมูลสำหรับการเรียนรู้ของโปรแกรมแปลภาษามือ โดยเป็นรูปภาพภาษามือตัวอักษร A ถึง Z รวมถึงรูปภาพภาษามือเครื่องหมายจุลภาค (,) เครื่องหมายมหัพภาค (.) เครื่องหมายปรัศนี (?) เครื่องหมายเว้นวรรค ภาษามือสำหรับลบตัวอักษร และภาษามือแสดงการส่งข้อความ มีจำนวนรูปภาพดังตารางที่ 3.1 เมื่อทำการตรวจสอบสัญลักษณ์ประจำรูปภาพที่ใช้เรียนรู้ ผลการตรวจสอบสัญลักษณ์เป็นดังนี้

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 0 พบว่าเป็น A เนื่องจาก A ถูกเก็บที่ตำแหน่ง 0 ถึง 1221
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1222 พบว่าเป็น B เนื่องจาก B ถูกเก็บที่ตำแหน่ง 1222 ถึง 2612
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2613 พบว่าเป็น C เนื่องจาก C ถูกเก็บที่ตำแหน่ง 2613 ถึง 3689
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 3690 พบว่าเป็น D เนื่องจาก D ถูกเก็บที่ตำแหน่ง 3690 ถึง 5275
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 5276 พบว่าเป็น E เนื่องจาก E ถูกเก็บที่ตำแหน่ง 5276 ถึง 6475
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 6476 พบว่าเป็น F เนื่องจาก F ถูกเก็บที่ตำแหน่ง 6476 ถึง 7884
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 7885 พบว่าเป็น G เนื่องจาก G ถูกเก็บที่ตำแหน่ง 7885 ถึง 9209
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 9210 พบว่าเป็น H เนื่องจาก H ถูกเก็บที่ตำแหน่ง 9210 ถึง 10767

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 10768 พบว่าเป็น I เนื่องจาก I ถูกเก็บที่ตำแหน่ง 10768 ถึง 12290

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 12291 พบว่าเป็น J เนื่องจาก J ถูกเก็บที่ตำแหน่ง 12291 ถึง 13515

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 13516 พบว่าเป็น K เนื่องจาก K ถูกเก็บที่ตำแหน่ง 13516 ถึง 14981

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 14982 พบว่าเป็น L เนื่องจาก L ถูกเก็บที่ตำแหน่ง 14982 ถึง 16359

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 16360 พบว่าเป็น M เนื่องจาก M ถูกเก็บที่ตำแหน่ง 16360 ถึง 17483

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 17484 พบว่าเป็น N เนื่องจาก N ถูกเก็บที่ตำแหน่ง 17484 ถึง 18607

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 18608 พบว่าเป็น O เนื่องจาก O ถูกเก็บที่ตำแหน่ง 18608 ถึง 19989

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 19990 พบว่าเป็น P เนื่องจาก P ถูกเก็บที่ตำแหน่ง 19990 ถึง 21211

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 21212 พบว่าเป็น Q เนื่องจาก Q ถูกเก็บที่ตำแหน่ง 21212 ถึง 22509

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 22510 พบว่าเป็น R เนื่องจาก R ถูกเก็บที่ตำแหน่ง 22510 ถึง 23521

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 23522 พบว่าเป็น S เนื่องจาก S ถูกเก็บที่ตำแหน่ง 23522 ถึง 24650

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 24651 พบว่าเป็น T เนื่องจาก T ถูกเก็บที่ตำแหน่ง 24651 ถึง 25770

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 25771 พบว่าเป็น U เนื่องจาก U ถูกเก็บที่ตำแหน่ง 25771 ถึง 26876

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 26876 พบว่าเป็น V เนื่องจาก V ถูกเก็บที่ตำแหน่ง 26876 ถึง 27995

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 27996 พบว่าเป็น W เนื่องจาก W ถูกเก็บที่ตำแหน่ง 27996 ถึง 29084

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 29085 พบว่าเป็น X เนื่องจาก X ถูกเก็บที่ตำแหน่ง 29085 ถึง 30341

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 30342 พบว่าเป็น Y เนื่องจาก Y ถูกเก็บที่ตำแหน่ง 30342 ถึง 31433

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 31434 พบว่าเป็น Z เนื่องจาก Z ถูกเก็บที่ตำแหน่ง 31434 ถึง 32771

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 32772 พบว่าเป็น comma เนื่องจากเครื่องหมายจุลภาค (,) ถูกเก็บที่ตำแหน่ง 32772 ถึง 34427

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 34428 พบว่าเป็น fullstop เนื่องจากเครื่องหมายมหัพภาค (.) ถูกเก็บที่ตำแหน่ง 34428 ถึง 35707

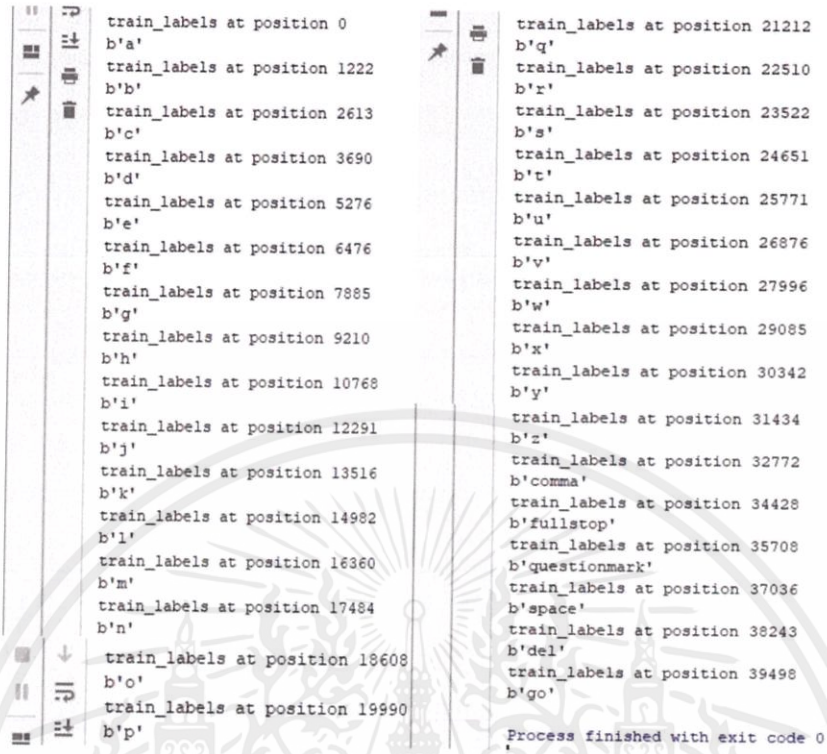
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 35708 พบว่าเป็น questionmark เนื่องจากเครื่องหมายปรัศนี (?) ถูกเก็บที่ตำแหน่ง 35708 ถึง 37029

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 37030 พบว่าเป็น space เนื่องจากเครื่องหมายเว้นวรรคถูกเก็บที่ตำแหน่ง 37030 ถึง 38242

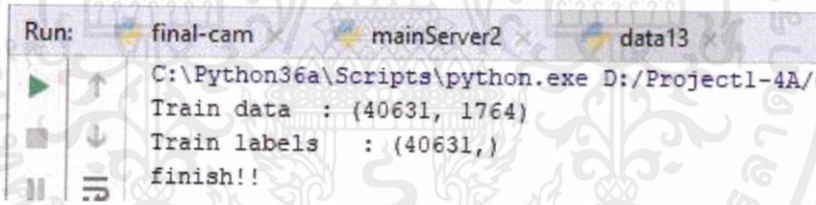
- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 38243 พบว่าเป็น del เนื่องจากสัญลักษณ์สำหรับลบตัวอักษรถูกเก็บที่ตำแหน่ง 38243 ถึง 39497

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 39498 พบว่าเป็น go เนื่องจากสัญลักษณ์สำหรับส่งข้อความถูกเก็บที่ตำแหน่ง 39498 ถึง 40630 โดยผลการตรวจสอบสัญลักษณ์แสดงดังรูปที่ 4.1 และจำนวนรูปภาพที่ใช้เรียนรู้รวม 40631 รูป ดังรูปที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 ผลการตรวจสอบสัญลักษณ์ที่ใช้ในการเรียนรู้



รูปที่ 4.2 จำนวนรูปภาพทั้งหมดที่ใช้เรียนรู้

จากนั้นตรวจสอบข้อมูลสำหรับใช้ทดสอบของโปรแกรมแปลภาษามือ โดยเป็นรูปภาพภาษามือตัวอักษร A ถึง Z รวมถึงรูปภาพภาษามือเครื่องหมายจุลภาค (,) เครื่องหมายมหัพภาค (.) เครื่องหมายปรัศนี (?) เครื่องหมายเว้นวรรค ภาษามือสำหรับลบตัวอักษร และภาษามือแสดงการส่งข้อความตัวอักษรละ 100 รูป เมื่อทำการตรวจสอบสัญลักษณ์ประจำรูปภาพที่ใช้ทดสอบได้ผลลัพธ์ดังนี้

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 0 พบว่าเป็น A เนื่องจาก A ถูกเก็บที่ตำแหน่ง 0 ถึง 99

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 100 พบว่าเป็น B เนื่องจาก B ถูกเก็บที่ตำแหน่ง 100 ถึง 199

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 200 พบว่าเป็น C เนื่องจาก C ถูกเก็บที่ตำแหน่ง 200 ถึง 299

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 300 พบว่าเป็น D เนื่องจาก D ถูกเก็บที่ตำแหน่ง 300 ถึง 399

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 400 พบว่าเป็น E เนื่องจาก E ถูกเก็บที่ตำแหน่ง 400 ถึง 499

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 500 พบว่าเป็น F เนื่องจาก F ถูกเก็บที่ตำแหน่ง 500 ถึง 599

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 600 พบว่าเป็น G เนื่องจาก G ถูกเก็บที่ตำแหน่ง 600 ถึง 699

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 700 พบว่าเป็น H เนื่องจาก H ถูกเก็บที่ตำแหน่ง 700 ถึง 799

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 800 พบว่าเป็น I เนื่องจาก I ถูกเก็บที่ตำแหน่ง 800 ถึง 899

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 900 พบว่าเป็น J เนื่องจาก J ถูกเก็บที่ตำแหน่ง 900 ถึง 999

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1000 พบว่าเป็น K เนื่องจาก K ถูกเก็บที่ตำแหน่ง 1000 ถึง 1099

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1100 พบว่าเป็น L เนื่องจาก L ถูกเก็บที่ตำแหน่ง 1100 ถึง 1199

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1200 พบว่าเป็น M เนื่องจาก M ถูกเก็บที่ตำแหน่ง 1200 ถึง 1299

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1300 พบว่าเป็น N เนื่องจาก N ถูกเก็บที่ตำแหน่ง 1300 ถึง 1399

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1400 พบว่าเป็น O เนื่องจาก O ถูกเก็บที่ตำแหน่ง 1400 ถึง 1499

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1500 พบว่าเป็น P เนื่องจาก P ถูกเก็บที่ตำแหน่ง 1500 ถึง 1599

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1600 พบว่าเป็น Q เนื่องจาก Q ถูกเก็บที่ตำแหน่ง 1600 ถึง 1699

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1700 พบว่าเป็น R เนื่องจาก R ถูกเก็บที่ตำแหน่ง 1700 ถึง 1799

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1800 พบว่าเป็น S เนื่องจาก S ถูกเก็บที่ตำแหน่ง 1800 ถึง 1899

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 1900 พบว่าเป็น T เนื่องจาก T ถูกเก็บที่ตำแหน่ง 1900 ถึง 1999

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2000 พบว่าเป็น U เนื่องจาก U ถูกเก็บที่ตำแหน่ง 2000 ถึง 2099

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2100 พบว่าเป็น V เนื่องจาก V ถูกเก็บที่ตำแหน่ง 2100 ถึง 2199

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2200 พบว่าเป็น W เนื่องจาก W ถูกเก็บที่ตำแหน่ง 2200 ถึง 2299

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2300 พบว่าเป็น X เนื่องจาก X ถูกเก็บที่ตำแหน่ง 2300 ถึง 2399

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2400 พบว่าเป็น Y เนื่องจาก Y ถูกเก็บที่ตำแหน่ง 2400 ถึง 2499

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2500 พบว่าเป็น Z เนื่องจาก Z ถูกเก็บที่ตำแหน่ง 2500 ถึง 2599

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2600 พบว่าเป็น comma เนื่องจากเครื่องหมายจุลภาค (,) ถูกเก็บที่ตำแหน่ง 2600 ถึง 2699

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2700 พบว่าเป็น fullstop เนื่องจากเครื่องหมายมหัพภาค (.) ถูกเก็บที่ตำแหน่ง 2700 ถึง 2799

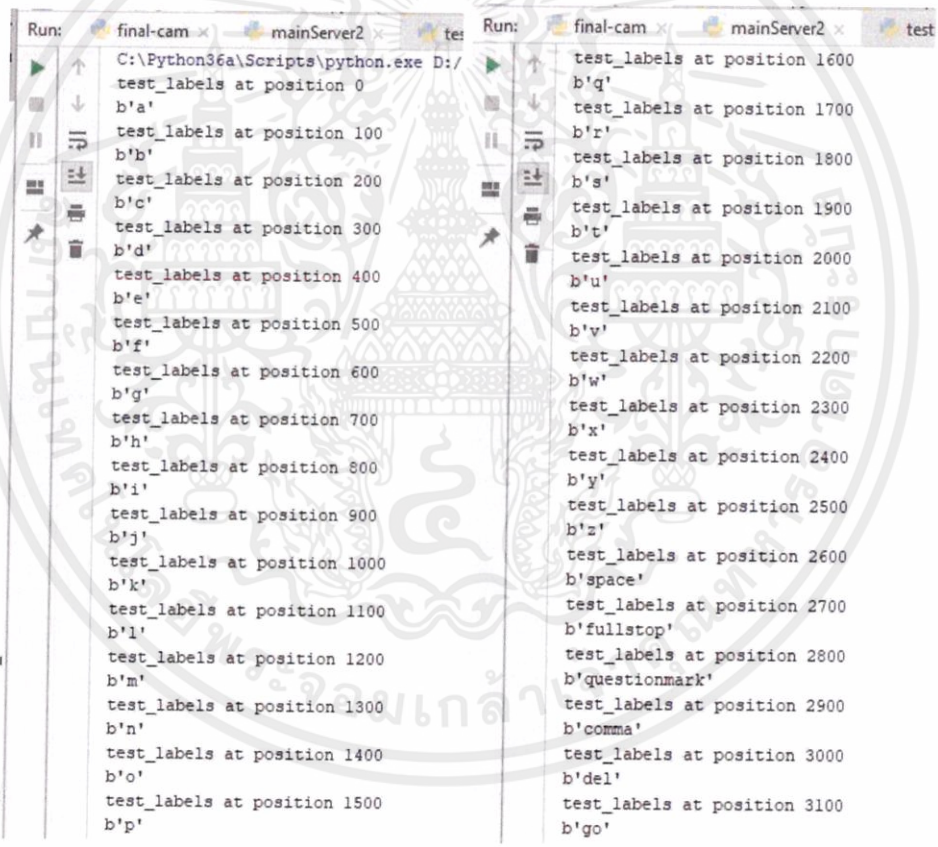
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2800 พบว่าเป็น questionmark เนื่องจากเครื่องหมายปรัศนี (?) ถูกเก็บที่ตำแหน่ง 2800 ถึง 2899

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 2900 พบว่าเป็น space เนื่องจากเครื่องหมายเว้นวรรคถูกเก็บที่ตำแหน่ง 2900 ถึง 2999

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 3000 พบว่าเป็น del เนื่องจากสัญลักษณ์สำหรับลบตัวอักษรถูกเก็บที่ตำแหน่ง 3000 ถึง 3099

- สัญลักษณ์ที่ถูกเก็บไว้ที่ตำแหน่ง 3100 พบว่าเป็น go เนื่องจากสัญลักษณ์แสดงการส่งข้อความถูกเก็บที่ตำแหน่ง 3100 ถึง 3199 โดยผลการตรวจสอบสัญลักษณ์แสดงดังรูปที่ 4.3 และจำนวนรูปภาพที่ใช้ทดสอบรวม 3200 รูป ดังรูปที่ 4.4



รูปที่ 4.3 ผลการตรวจสอบสัญลักษณ์ที่ใช้ในการทดสอบ

```
Test data  : (3200, 1764)
Test labels   : (3200,)
finish!!
```

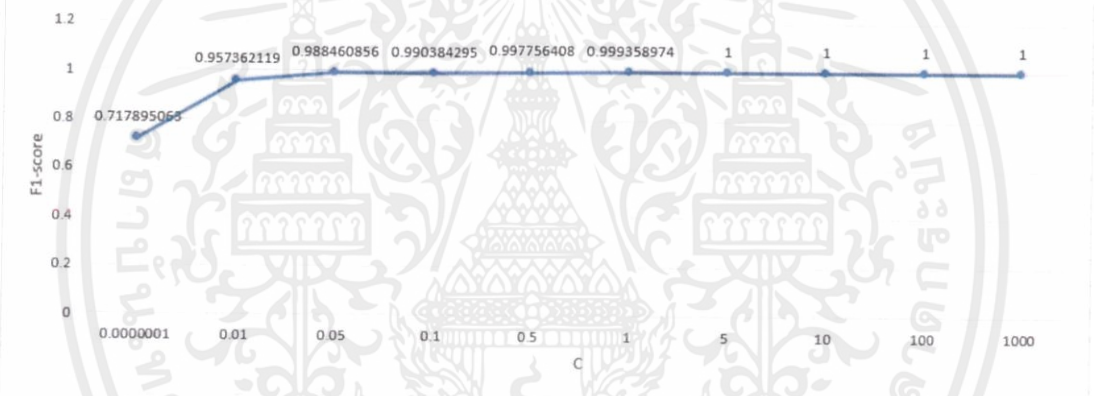
รูปที่ 4.4 จำนวนรูปภาพทั้งหมดที่ใช้ทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

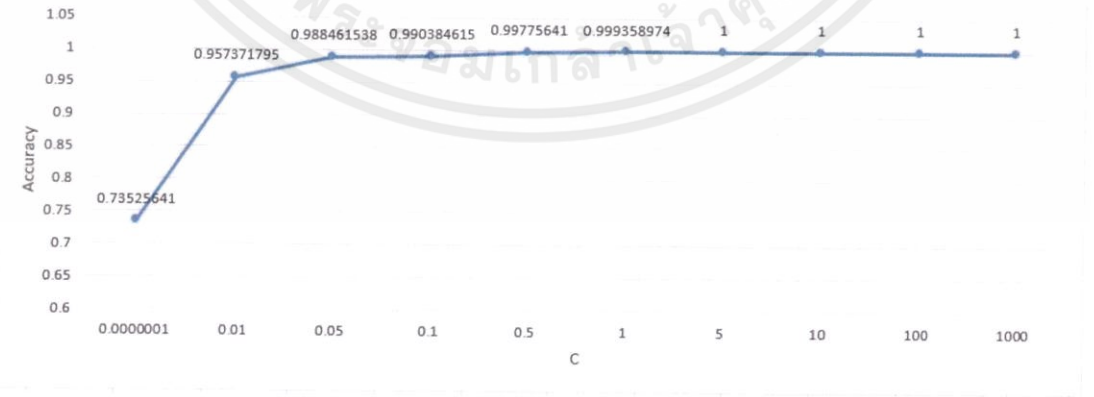
4.2 การทดสอบพารามิเตอร์สำหรับการเรียนรู้ของโปรแกรมแปลภาษามือเป็นตัวอักษร

4.2.1 ทดสอบพารามิเตอร์ของ Binary และ multi-class SVM แบบ Linear kernel

ในการทดสอบ Linear SVM แบบ Binary ได้ใช้ข้อมูลรูปภาพภาษามือตัวอักษร A และ B จำนวนตัวอักษรละ 1,560 รูปสำหรับการฝึกสอน และใช้ข้อมูลรูปภาพภาษามือตัวอักษร A และ B จำนวนตัวอักษรละ 100 รูปสำหรับการทดสอบ ทำการกำหนดพารามิเตอร์ C เพื่อใช้ในการทดสอบ และใช้วิธี SVM ในการจำแนกข้อมูลเพื่อหาค่า f1-score และ Accuracy ตามสมการที่ (3.1) และ (3.4) ตามลำดับ โดยทดสอบค่า C 10 ค่าคือ 0.0000001, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 100, 1000 กับข้อมูลที่ใช้ฝึกสอน พบว่าเมื่อค่า C มากขึ้นทำให้ f1-score มากขึ้นดังรูปที่ 4.5 และทำให้ Accuracy มากขึ้นดังรูปที่ 4.6



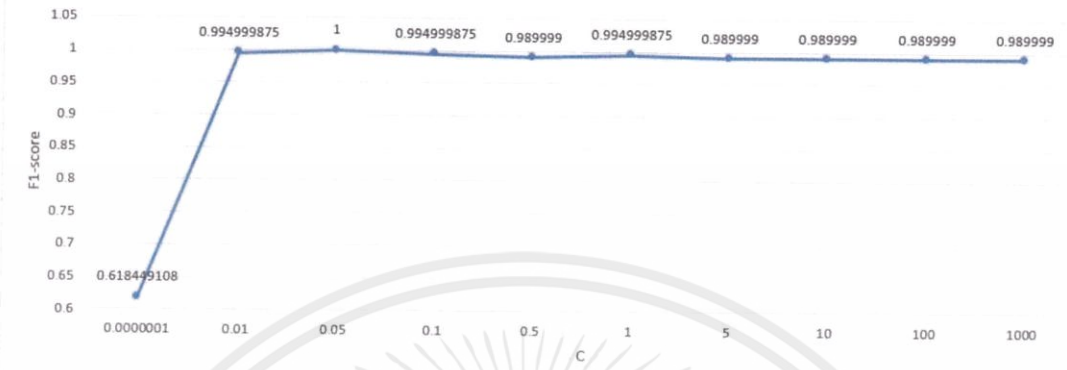
รูปที่ 4.5 กราฟความสัมพันธ์ของ C และ f1-score
ของข้อมูล A-B สำหรับข้อมูลที่ใช้ฝึกสอน



รูปที่ 4.6 กราฟความสัมพันธ์ของ C และ Accuracy
ของข้อมูล A-B สำหรับข้อมูลที่ใช้ฝึกสอน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับข้อมูลที่ใช้ทดสอบพบว่าที่ค่า C = 0.05 ได้ผล f1-score มากที่สุดเท่ากับ 1 ดังรูปที่ 4.7 และทำให้ Accuracy มากที่สุดเท่ากับ 1 ดังรูปที่ 4.8



รูปที่ 4.7 กราฟความสัมพันธ์ของ C และ f1-score
ของข้อมูล A-B สำหรับข้อมูลที่ใช้ทดสอบ

รูปที่ 4.8 กราฟความสัมพันธ์ของ C และ Accuracy
ของข้อมูล A-B สำหรับข้อมูลที่ใช้ทดสอบ

จากนั้นทำการทดสอบ Linear SVM แบบ multi-class โดยการเพิ่มข้อมูลรูปภาพภาษามือเป็นตัวอักษร A ถึง N ยกเว้นอักษร J เนื่องจากเป็นอักษรที่มีการเคลื่อนไหวของมือ โดยใช้จำนวนตัวอักษรดังตารางที่ 4.1 สำหรับการฝึกสอน และใช้ข้อมูลรูปภาพภาษามือดังตารางที่ 4.2 สำหรับการทดสอบ ทำการกำหนดพารามิเตอร์ C เพื่อใช้ในการทดสอบ และหาค่า f1-score และ Accuracy โดยทดสอบค่า C 10 ค่าคือ 0.0000001, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 100, 1000 กับข้อมูลที่ใช้ฝึกสอนพบว่าเมื่อค่า C มากขึ้นส่งผลให้ f1-score และ Accuracy มากขึ้นดังรูปที่ 4.9 และรูปที่ 4.10 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

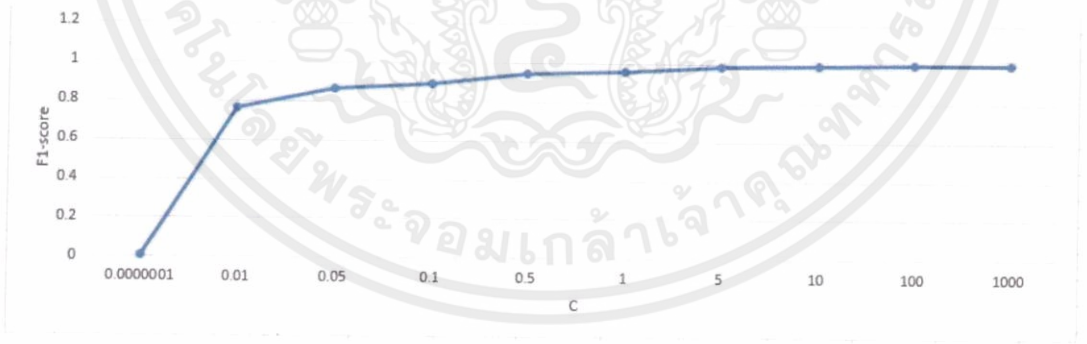
ตารางที่ 4.1 จำนวนรูปภาพที่ใช้ในการเรียนรู้อักษร A-N

ตัวอักษร	จำนวนรูปภาพที่ใช้ในการเรียนรู้
A	1560
B	1560
C	1495
D	1586
E	1597
F	1610
G	1325
H	1608
I	1617
K	1592
L	1559
M	1609
N	1592

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

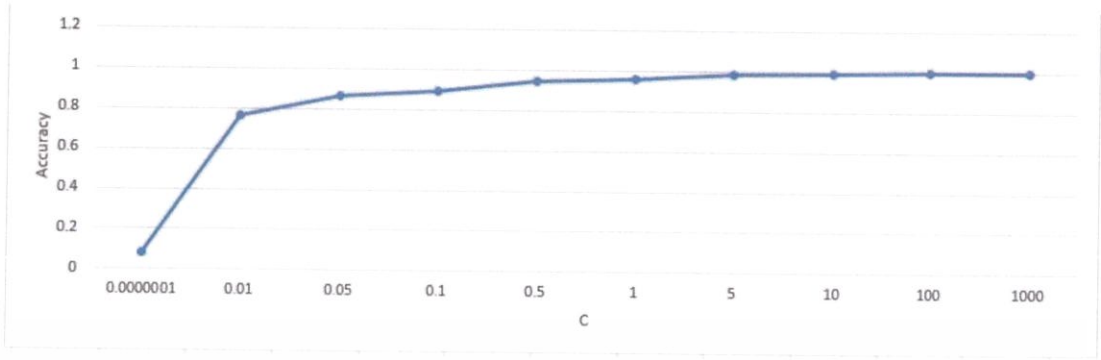
ตารางที่ 4.2 จำนวนรูปภาพที่ใช้ในการทดสอบอักษร A-N

ตัวอักษร	จำนวนรูปภาพที่ใช้ทดสอบ
A	100
B	100
C	100
D	100
E	100
F	100
G	100
H	100
I	100
K	75
L	25
M	25
N	25



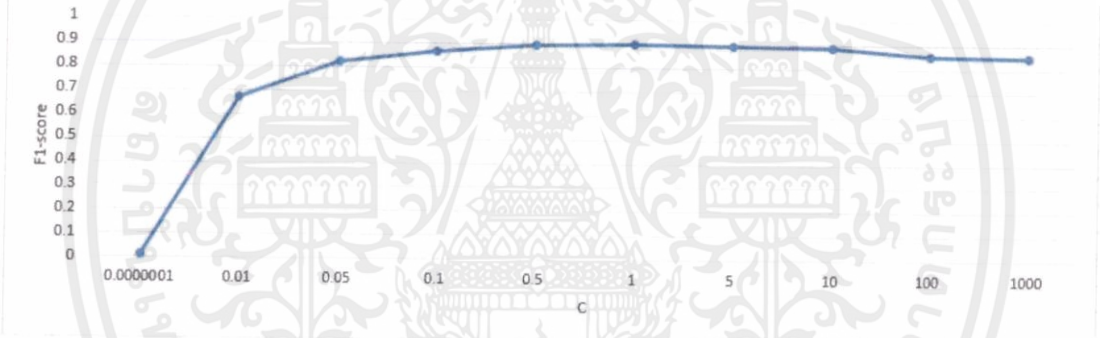
รูปที่ 4.9 กราฟความสัมพันธ์ของ C และ f1-score
ของข้อมูล A-N สำหรับข้อมูลที่ใช้ฝึกสอน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

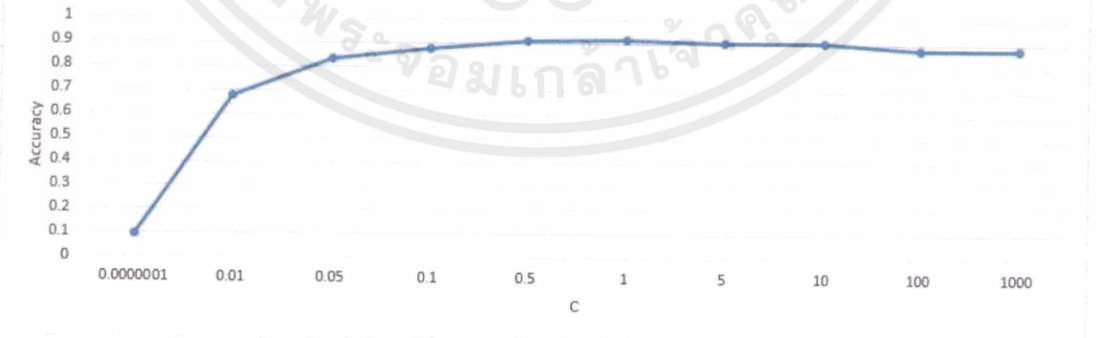


รูปที่ 4.10 กราฟความสัมพันธ์ของ C และ Accuracy
ของข้อมูล A-N สำหรับข้อมูลที่ใช้ฝึกสอน

สำหรับข้อมูลที่ใช้ทดสอบพบว่าที่ค่า C = 1 ได้ผล f1-score มากที่สุดเท่ากับ 0.901185636 ดังรูปที่ 4.11 และทำให้ Accuracy มากที่สุดเท่ากับ 0.915238095 ดังรูปที่ 4.12



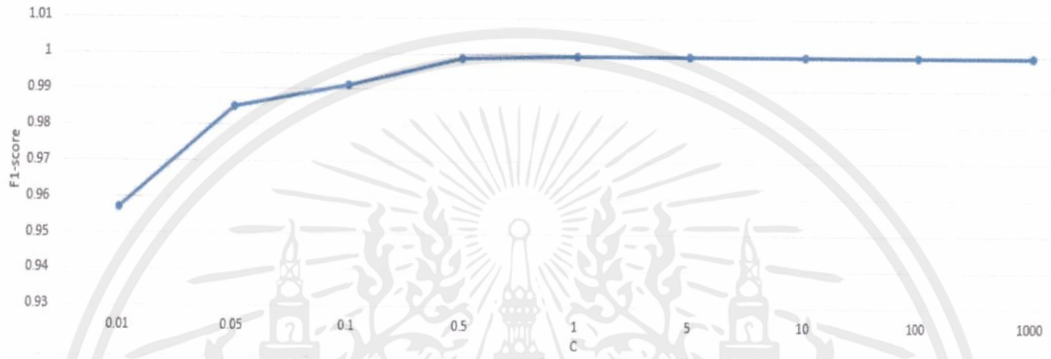
รูปที่ 4.11 กราฟความสัมพันธ์ของ C และ f1-score
ของข้อมูล A-N สำหรับข้อมูลที่ใช้ทดสอบ



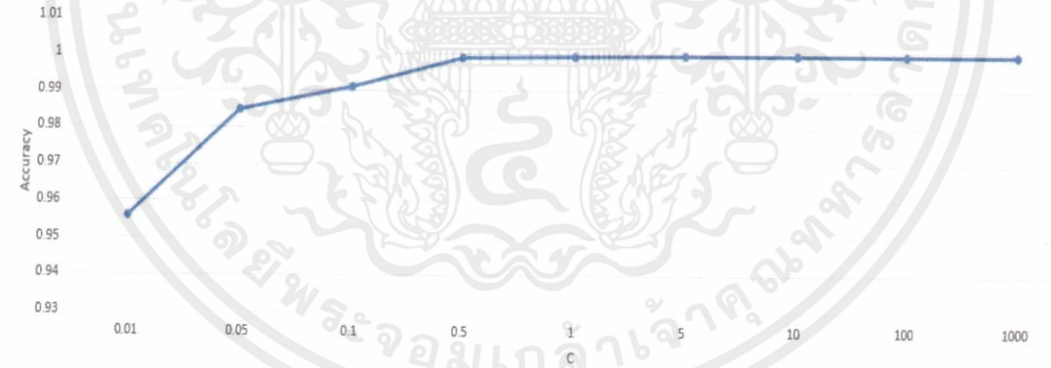
รูปที่ 4.12 กราฟความสัมพันธ์ของ C และ Accuracy
ของข้อมูล A-N สำหรับข้อมูลที่ใช้ทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดสอบถัดไปทำการเพิ่มข้อมูลรูปภาพภาษามือเป็นตัวอักษรทุกตัวอักษรคือ A ถึง Z และเครื่องหมายวรรคตอนต่างๆ มีจำนวนดังตารางที่ 3.1 สำหรับการฝึกสอน และใช้ข้อมูลรูปภาพภาษามือจำนวน 100 รูปต่อตัวอักษรสำหรับการทดสอบ ทำการกำหนดพารามิเตอร์ C เพื่อใช้ในการทดสอบ และใช้วิธี SVM แบบ multi-class ในการจำแนกข้อมูลเพื่อหาค่า f1-score และ Accuracy ทดสอบค่า C รวม 9 ค่าคือ 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 100, 1000 กับข้อมูลที่ใช้ฝึกสอนพบว่าเมื่อค่า C มากขึ้นส่งผลให้ f1-score และ Accuracy มากขึ้นดังรูปที่ 4.13 และรูปที่ 4.14 ตามลำดับ



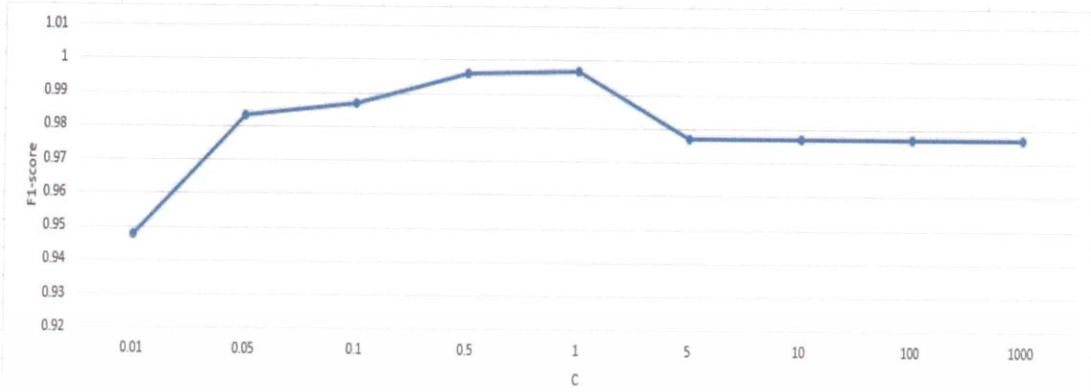
รูปที่ 4.13 กราฟความสัมพันธ์ของ C และ f1-score
ของข้อมูลทุกตัวอักษรสำหรับข้อมูลที่ใช้ฝึกสอน



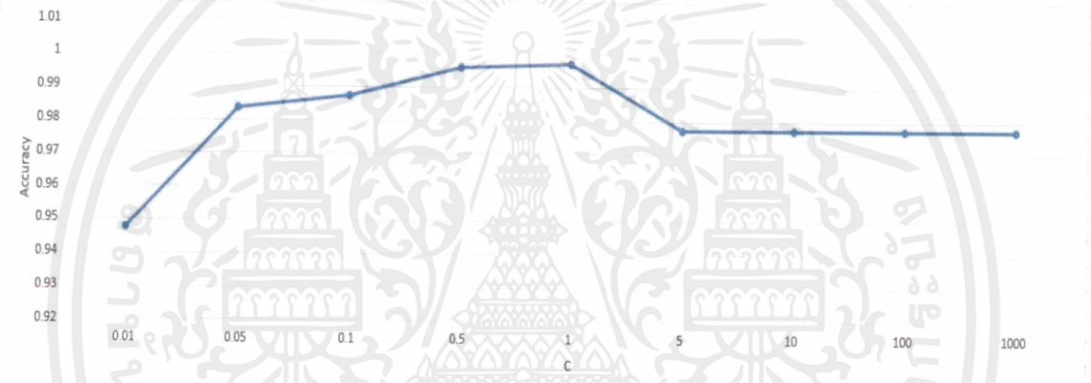
รูปที่ 4.14 กราฟความสัมพันธ์ของ C และ Accuracy
ของข้อมูลทุกตัวอักษรสำหรับข้อมูลที่ใช้ฝึกสอน

สำหรับข้อมูลที่ใช้ทดสอบพบว่าที่ค่า C = 1 ได้ผล f1-score มากที่สุดเท่ากับ 0.996874844 ดังรูปที่ 4.15 และทำให้ Accuracy มากที่สุดเท่ากับ 0.996875 ดังรูปที่ 4.16 เช่นเดียวกับการใช้ข้อมูลตัวอักษร A-N

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.15 กราฟความสัมพันธ์ของ C และ f1-score
ของข้อมูลทุกตัวอักษรสำหรับข้อมูลที่ใช้ทดสอบ

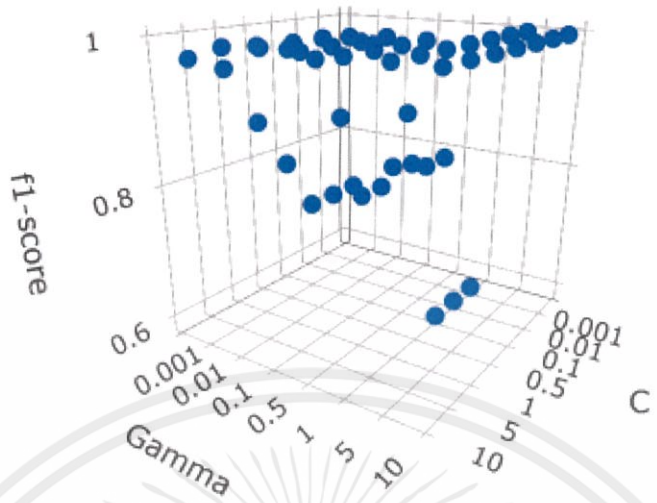


รูปที่ 4.16 กราฟความสัมพันธ์ของ C และ Accuracy
ของข้อมูลทุกตัวอักษรสำหรับข้อมูลที่ใช้ทดสอบ

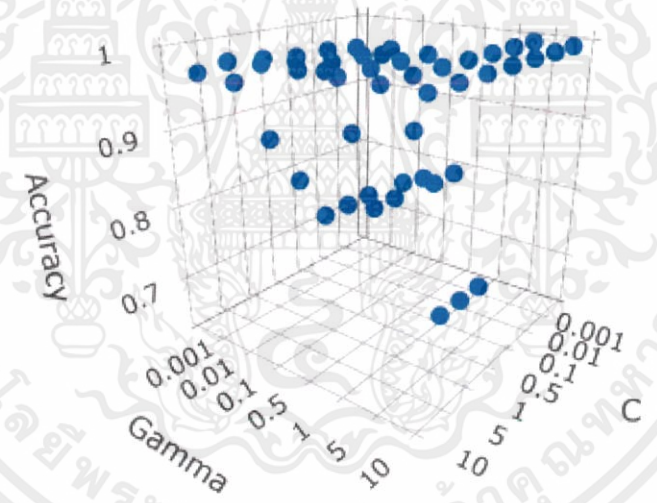
4.2.2 ทดสอบพารามิเตอร์ของ Binary SVM แบบ RBF kernel

ในการทดสอบได้ใช้ข้อมูลรูปภาพภาษามือตัวอักษร A และ B จำนวนตัวอักษรละ 1,560 รูปสำหรับการฝึกสอน และใช้ข้อมูลรูปภาพภาษามือตัวอักษร A และ B จำนวนตัวอักษรละ 100 รูป สำหรับการทดสอบ ทำการกำหนดพารามิเตอร์ C และ gamma เพื่อใช้ในการทดสอบ และใช้วิธี SVM ในการจำแนกข้อมูลเพื่อหาค่า f1-score และ Accuracy ทดสอบค่า C 7 ค่าคือ 0.001, 0.01, 0.1, 0.5, 1, 5, 10 และค่า gamma 7 ค่าคือ 0.001, 0.01, 0.1, 0.5, 1, 5, 10 กับข้อมูลที่ใช้ฝึกสอนพบว่าเมื่อค่า C และ gamma มากขึ้นทำให้ f1-score มากขึ้นดังรูปที่ 4.17 และทำให้ Accuracy มากขึ้นดังรูปที่ 4.18

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.17 กราฟความสัมพันธ์ของ C, gamma และ f1-score
ของข้อมูล A-B สำหรับข้อมูลที่ใช้ฝึกสอน



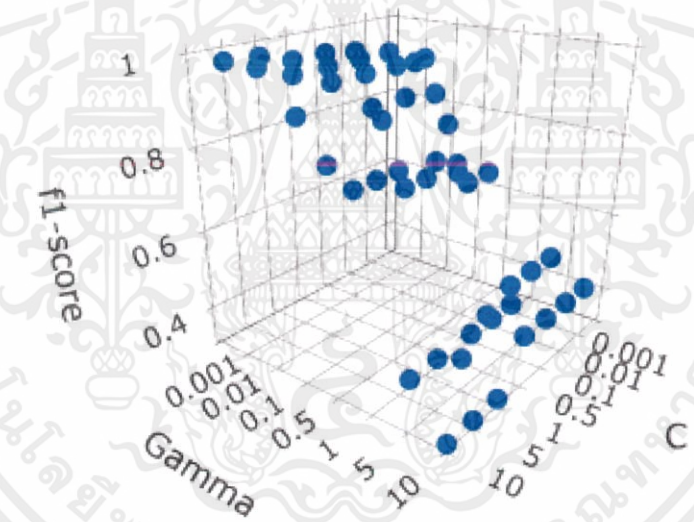
รูปที่ 4.18 กราฟความสัมพันธ์ของ C, gamma และ Accuracy
ของข้อมูล A-B สำหรับข้อมูลที่ใช้ฝึกสอน

ผลลัพธ์ของการทดสอบพารามิเตอร์ C และ gamma สำหรับข้อมูลที่ใช้ทดสอบ ค่า C และ gamma ที่ทำให้ f1-score มากที่สุดเท่ากับ 1 เป็นดังรูปที่ 4.19 และทำให้ Accuracy มากที่สุดเท่ากับ 1 เป็นดังรูปที่ 4.20 โดยค่า C และ gamma ที่ทำให้ f1-score และ Accuracy มากที่สุดมีดังนี้

- C = 0.5, gamma = 0.1
- C = 0.5, gamma = 0.5
- C = 1, gamma = 0.001

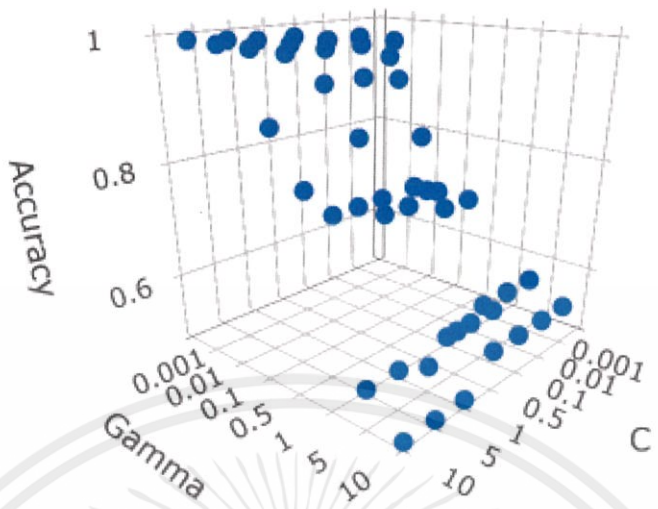
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- C = 1, gamma = 0.1
- C = 1, gamma = 0.5
- C = 5, gamma = 0.01
- C = 5, gamma = 0.1
- C = 5, gamma = 0.5
- C = 10, gamma = 0.001
- C = 10, gamma = 0.01
- C = 10, gamma = 0.1
- C = 10, gamma = 0.5



รูปที่ 4.19 กราฟความสัมพันธ์ของ C, gamma และ f1-score ของข้อมูล A-B สำหรับข้อมูลที่ใช้ทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.20 กราฟความสัมพันธ์ของ C, gamma และ Accuracy
ของข้อมูล A-B สำหรับข้อมูลที่ใช้ทดสอบ

จากผลการทดสอบพารามิเตอร์ของ SVM แบบ Linear kernel และแบบ RBF kernel ในปริญญานิพนธ์นี้ผู้จัดทำได้เลือกใช้ SVM แบบ Linear kernel ที่ C เท่ากับ 1 เนื่องจากทำให้ค่า f1-score และ Accuracy มีค่าสูงที่สุดเมื่อเพิ่มข้อมูลเป็นทุกตัวอักษร คือตัวอักษร A ถึง Z และเครื่องหมายวรรคตอนต่างๆ และมีการคำนวณที่รวดเร็วกว่า

4.3 ความแม่นยำในการจำแนกตัวอักษรที่ของ Linear SVM แบบ multi-class

ค่า Accuracy ของแต่ละตัวอักษร คือตัวอักษร A ถึง Z ที่ถูกจำแนกด้วย Linear SVM แสดงดังตารางที่ 4.3 และเครื่องหมายวรรคตอนต่างๆ ดังตารางที่ 4.4 โดยใช้ข้อมูลที่ฝึกสอนด้วยวิธี SVM แบบ Linear kernel ที่ C เท่ากับ 1 และรูปภาพที่ใช้ทดสอบมีจำนวน 100 ภาพต่อตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 Accuracy ของแต่ละตัวอักษร

ตัวอักษร	Accuracy
a	1
b	0.99
c	1
d	1
e	1
f	1
g	1
h	1
i	1
j	1
k	1
l	1
m	1
n	1
o	1
p	1
q	1
r	0.98
s	0.99
t	0.99
u	1
v	0.99
w	0.99
x	1
y	0.99
z	1

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

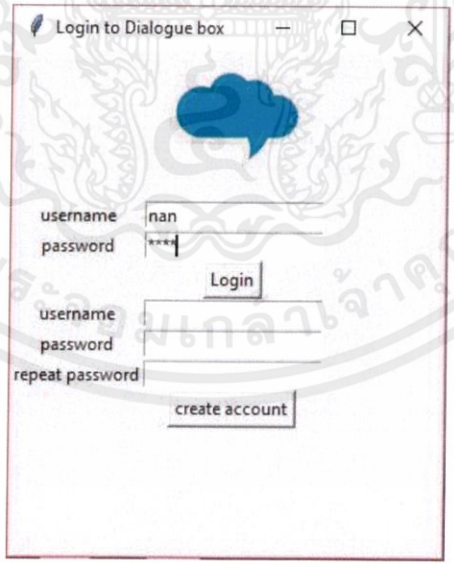
ตารางที่ 4.4 Accuracy ของเครื่องหมายวรรคตอนต่างๆ

เครื่องหมายวรรคตอน	Accuracy
เครื่องหมายจุลภาค (,)	1
เครื่องหมายมหัพภาค (.)	1
เครื่องหมายปรัศนี (?)	0.98
เครื่องหมายเว้นวรรค	1
สัญลักษณ์สำหรับลบตัวอักษร	1
สัญลักษณ์แสดงการส่งข้อความ	1

4.4 การทดสอบส่วนต่อประสานกราฟิกกับผู้ใช้และฐานข้อมูล

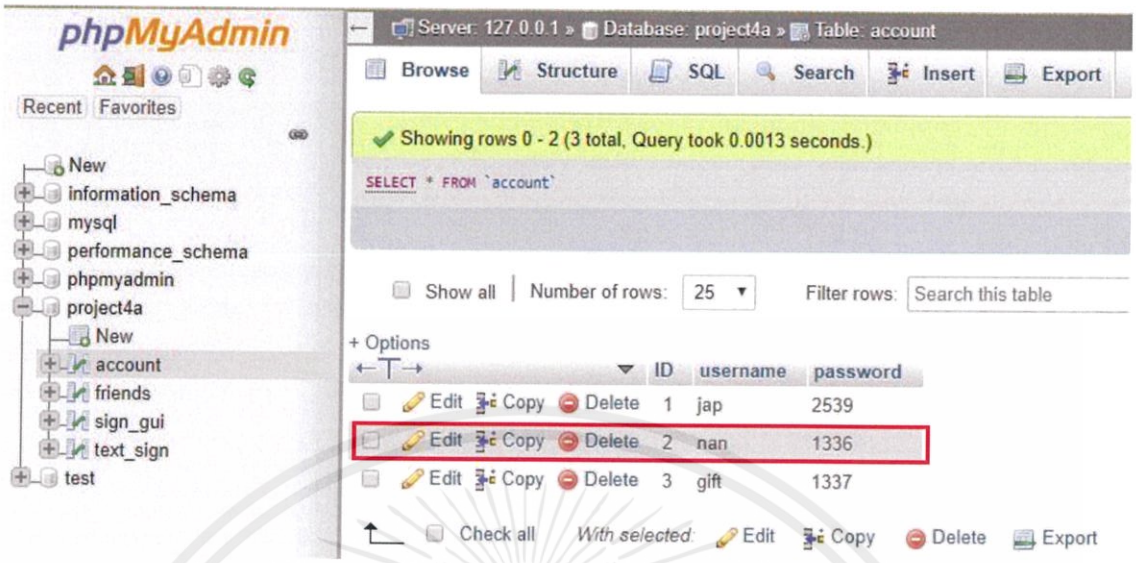
4.4.1 การทดสอบหน้าต่าง Login และลงทะเบียนเข้าใช้งานกับฐานข้อมูล

หน้าต่าง Login ดังรูปที่ 4.21 ใช้สำหรับเข้าสู่หน้าต่างการสนทนา เมื่อกรอกข้อมูล username และ password ถูกต้องตรงตามฐานข้อมูลดังรูปที่ 4.22 จะสามารถเข้าสู่หน้าต่างการสนทนาได้ หากกรอกข้อมูล username และ password ไม่ถูกต้องตามฐานข้อมูลจะมีการแจ้งเตือนดังรูปที่ 4.23 และไม่สามารถเข้าสู่หน้าต่างการสนทนาได้

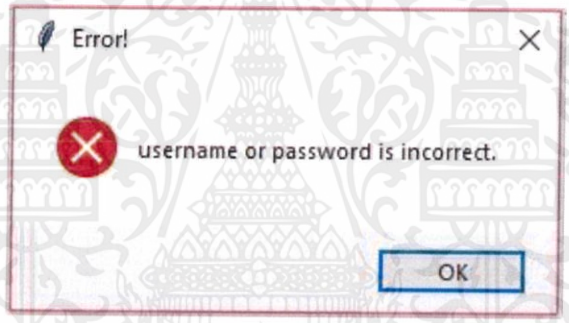


รูปที่ 4.21 หน้าต่าง Login สำหรับเข้าสู่หน้าต่างการสนทนา

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



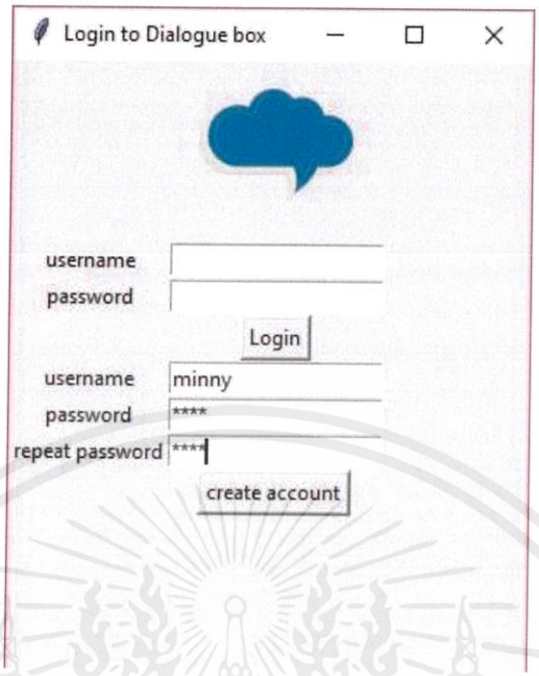
รูปที่ 4.22 ฐานข้อมูลสำหรับเก็บข้อมูล Username และ Password



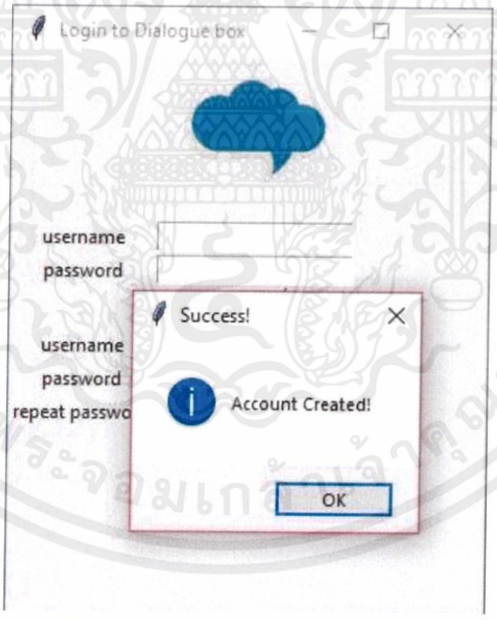
รูปที่ 4.23 การแจ้งเตือนหากกรอก Username และ Password ไม่ถูกต้อง

การลงทะเบียนเข้าใช้งานสามารถกรอกข้อมูล username และ password และกดปุ่ม create account ดังรูปที่ 4.24 เมื่อทำการลงทะเบียนสำเร็จจะมีข้อความยืนยันว่า Account Created! ดังรูปที่ 4.25 แสดงว่าข้อมูลที่ได้ลงทะเบียนถูกนำไปเก็บในฐานข้อมูลแล้วดังรูปที่ 4.26

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

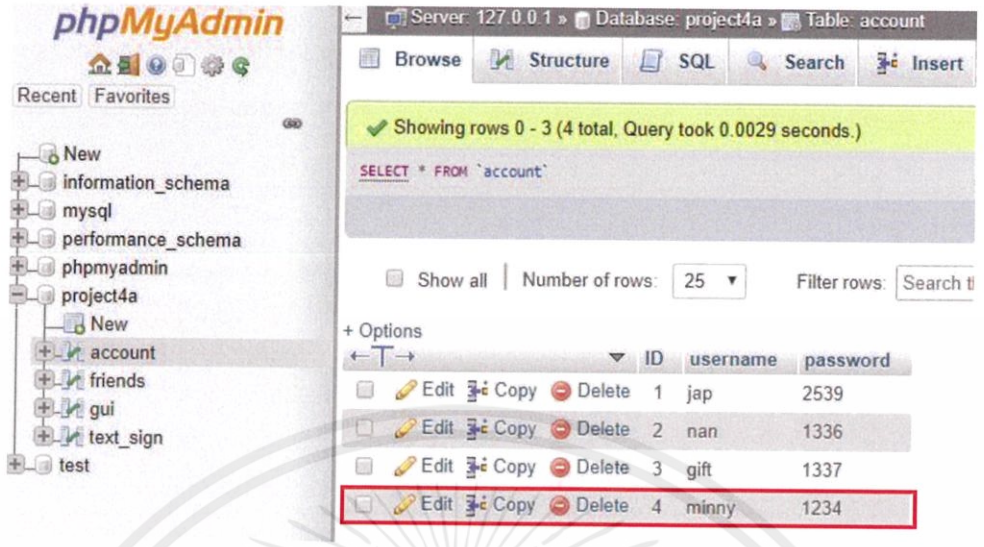


รูปที่ 4.24 หน้าต่างลงทะเบียนผู้ใช้งาน



รูปที่ 4.25 ข้อความยืนยันการลงทะเบียนสำเร็จ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



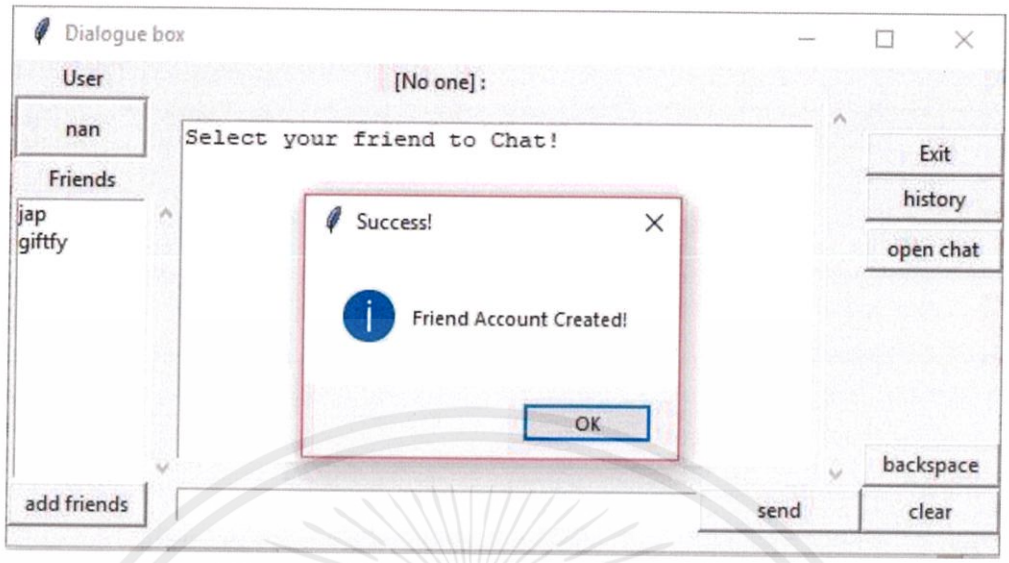
รูปที่ 4.26 การบันทึกข้อมูลการลงทะเบียนลงในฐานข้อมูล

4.4.2 การทดสอบหน้าต่าง Add Friend กับฐานข้อมูล และการแสดงรายชื่อผู้ติดต่อในหน้าต่าง Dialogue box

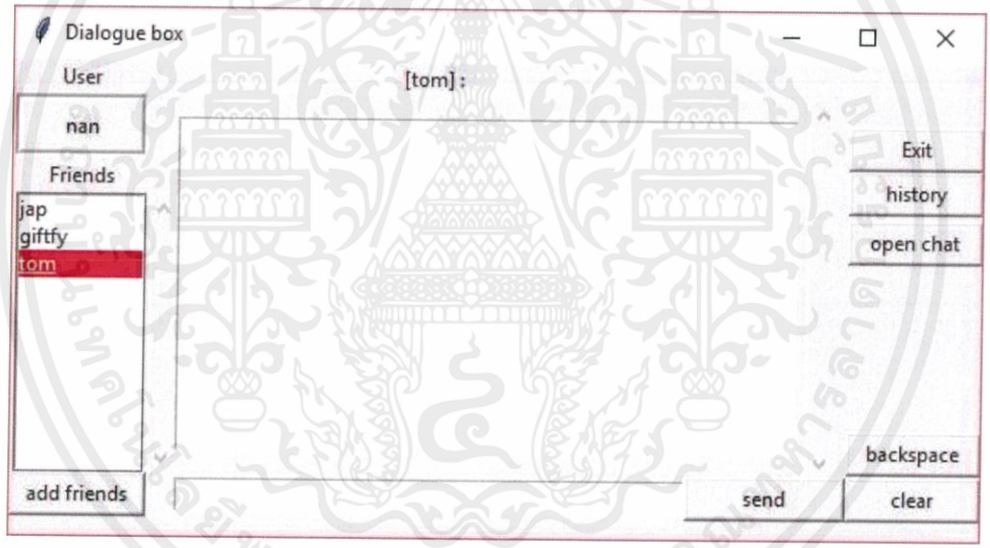
จากหน้าต่าง Add friend ดังรูปที่ 4.27 เมื่อกรอกรายชื่อผู้ติดต่อใหม่และกดปุ่ม Add Friend จะมีข้อความยืนยันว่า Friend Account Created! ดังรูปที่ 4.28 และรายชื่อผู้ติดต่อจะแสดงบนหน้าต่าง Dialogue box อัตโนมัติดังรูปที่ 4.29 โดยข้อมูลรายชื่อผู้ติดต่อใหม่จะถูกบันทึกลงในฐานข้อมูลดังรูปที่ 4.30

รูปที่ 4.27 การกรอกรายชื่อผู้ติดต่อใหม่ในหน้าต่าง Add Friend

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

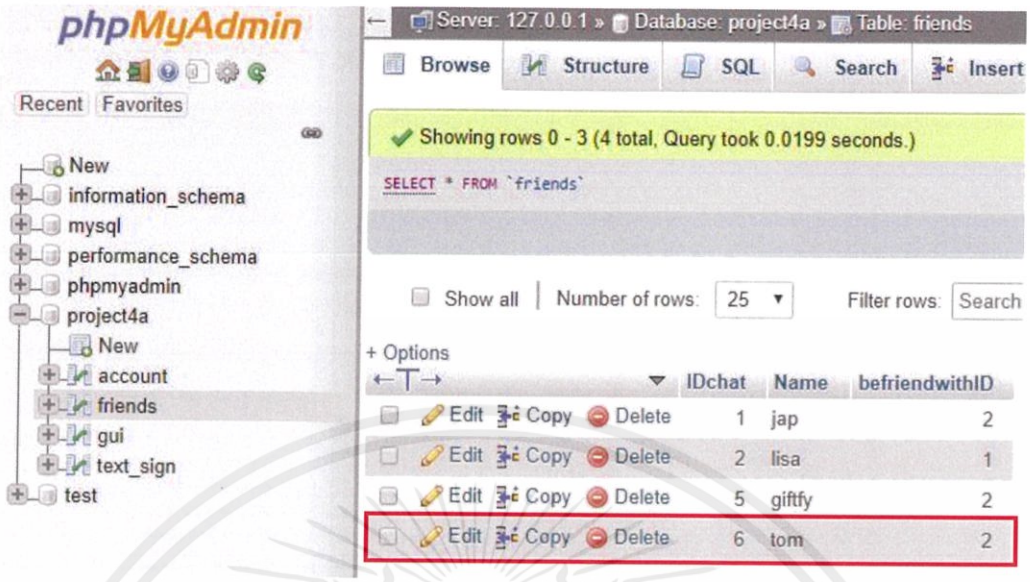


รูปที่ 4.28 ข้อความยืนยันการเพิ่มรายชื่อผู้ติดต่อสำเร็จ



รูปที่ 4.29 รายชื่อผู้ติดต่อใหม่บนหน้าต่าง Dialogue box

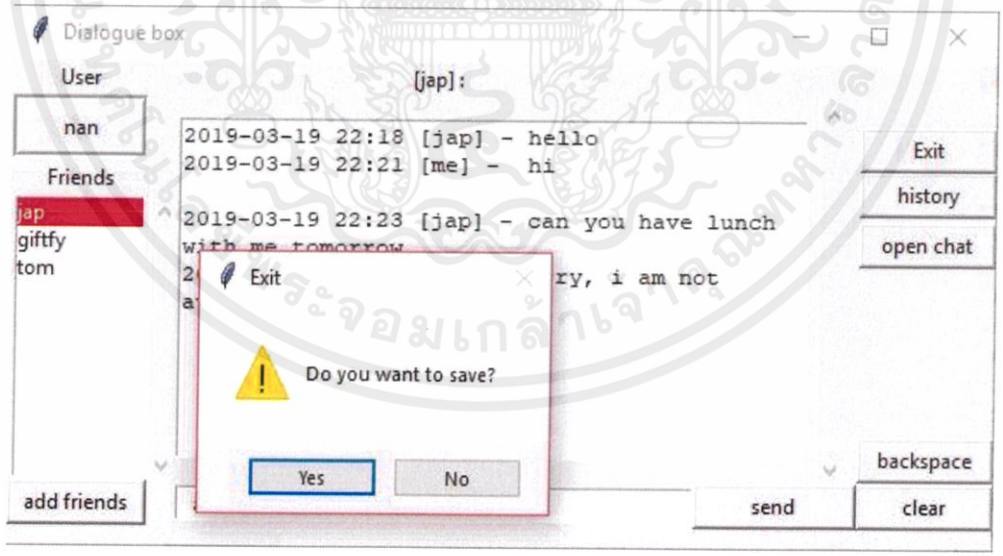
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.30 ฐานข้อมูลสำหรับเก็บข้อมูลรายชื่อผู้ติดต่อ

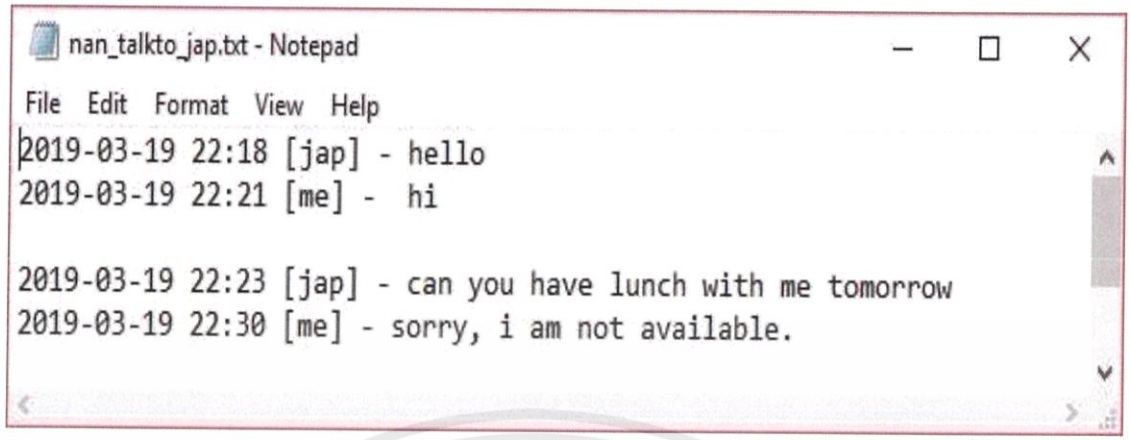
4.4.3 การทดสอบหน้าต่าง Exit และการบันทึกข้อมูลการสนทนาเป็นไฟล์ .txt

การบันทึกข้อมูลการสนทนาเป็นดังรูปที่ 4.31 กดปุ่ม Exit เมื่อต้องการออกจากการสนทนา ผู้ใช้สามารถเลือกบันทึกข้อมูลการสนทนาได้ โดยระบบจะทำการบันทึกข้อมูลการสนทนาลงในไฟล์ .txt ดังรูปที่ 4.32



รูปที่ 4.31 การบันทึกข้อมูลการสนทนา

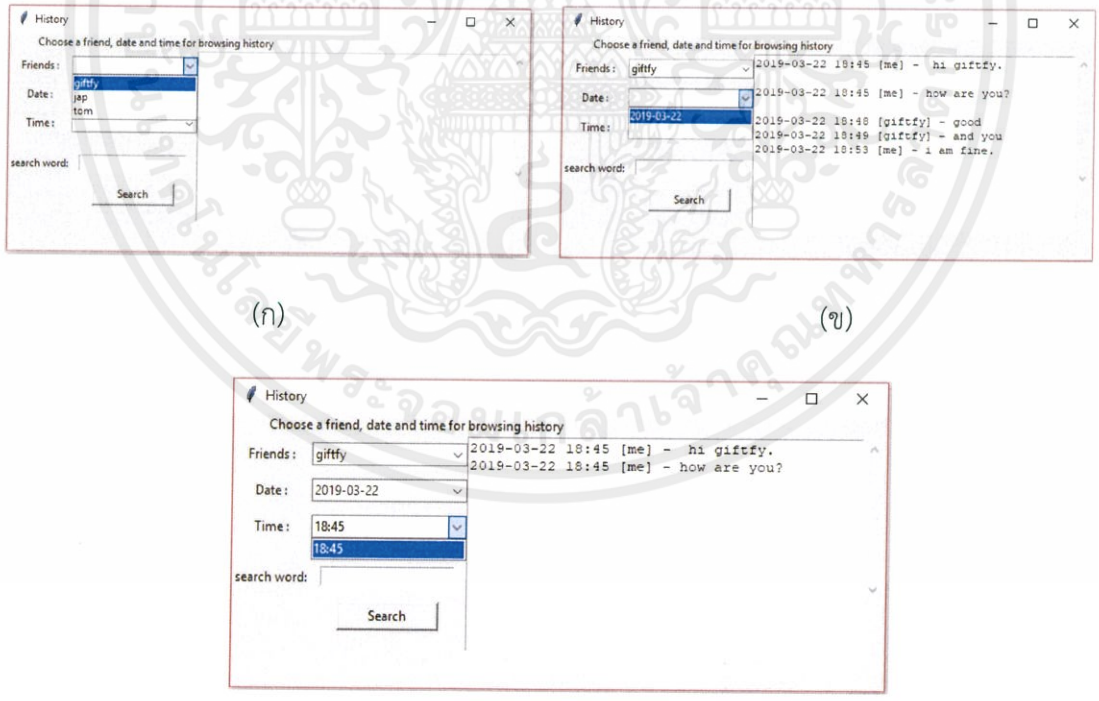
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.32 ข้อมูลการสนทนาที่ถูกบันทึกในไฟล์ .txt

4.4.4 การทดสอบหน้าต่าง History กับฐานข้อมูล

การเรียกดูประวัติการสนทนาสามารถทำได้โดยการเลือกชื่อคู่สนทนา วันที่ และเวลาที่ต้องการดู ประวัติการสนทนาจะแสดงที่หน้าต่างแสดงประวัติการสนทนาดังรูปที่ 4.33 โดยประวัติการสนทนาจะตรงกับฐานข้อมูลที่ถูกบันทึกไว้ดังรูปที่ 4.34 การค้นหาคำทำได้โดยพิมพ์ที่ช่อง search word และกดปุ่ม Search คำที่ค้นหาจะแสดงที่หน้าต่างแสดงประวัติการสนทนาดังรูปที่ 4.35



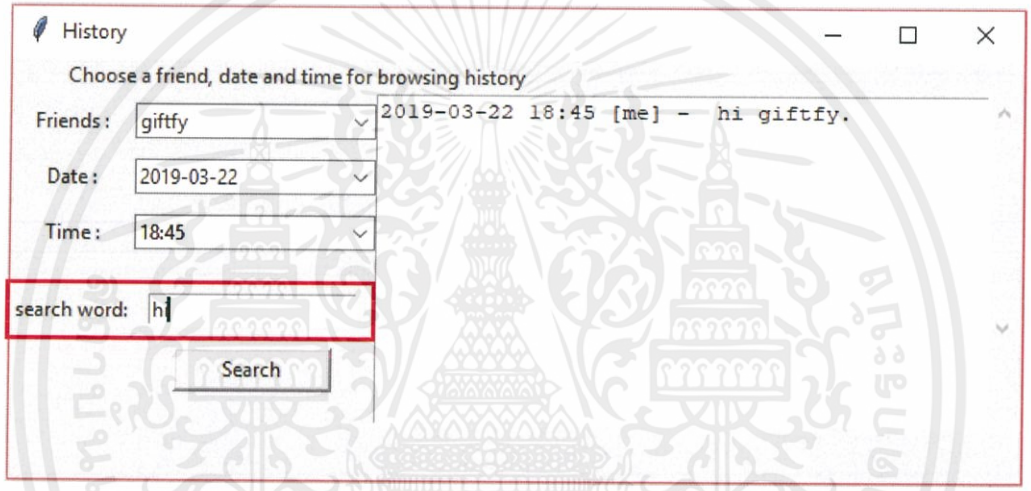
(ก) (ข)

(ค)

รูปที่ 4.33 การเรียกดูประวัติการสนทนา (ก) การเลือกชื่อคู่สนทนา
(ข) การเลือกวันที่ (ค) การเลือกเวลา

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.34 ฐานข้อมูลประวัติการสนทนา



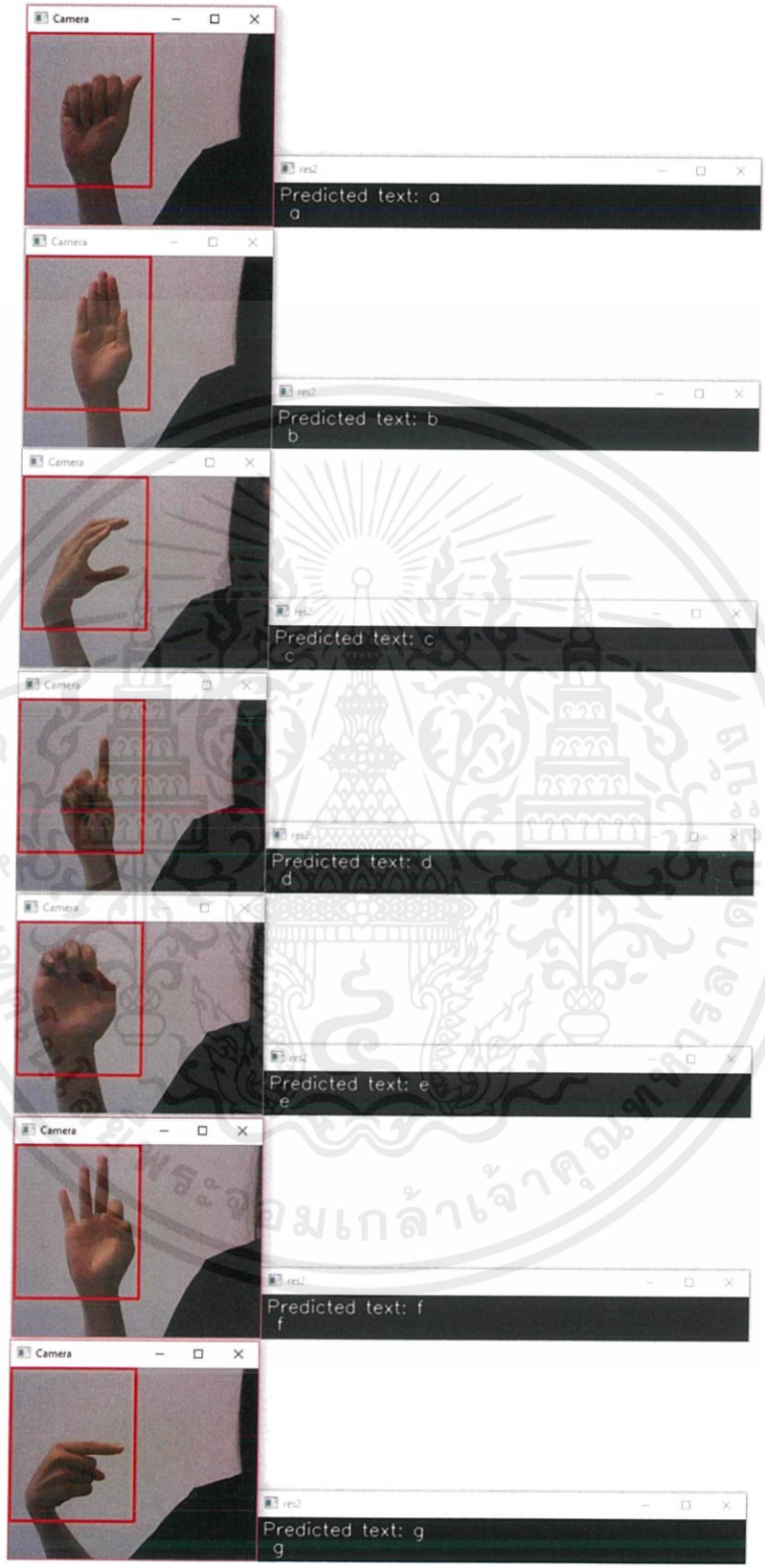
รูปที่ 4.35 ประวัติการสนทนาจากการค้นหาคำ

4.5 การทดสอบผลลัพธ์โปรแกรมแปลภาษามือเป็นตัวอักษร

4.5.1 ทดสอบผลลัพธ์การแปลภาษามือของตัวอักษร A-Z และเครื่องหมายวรรคตอนต่างๆ

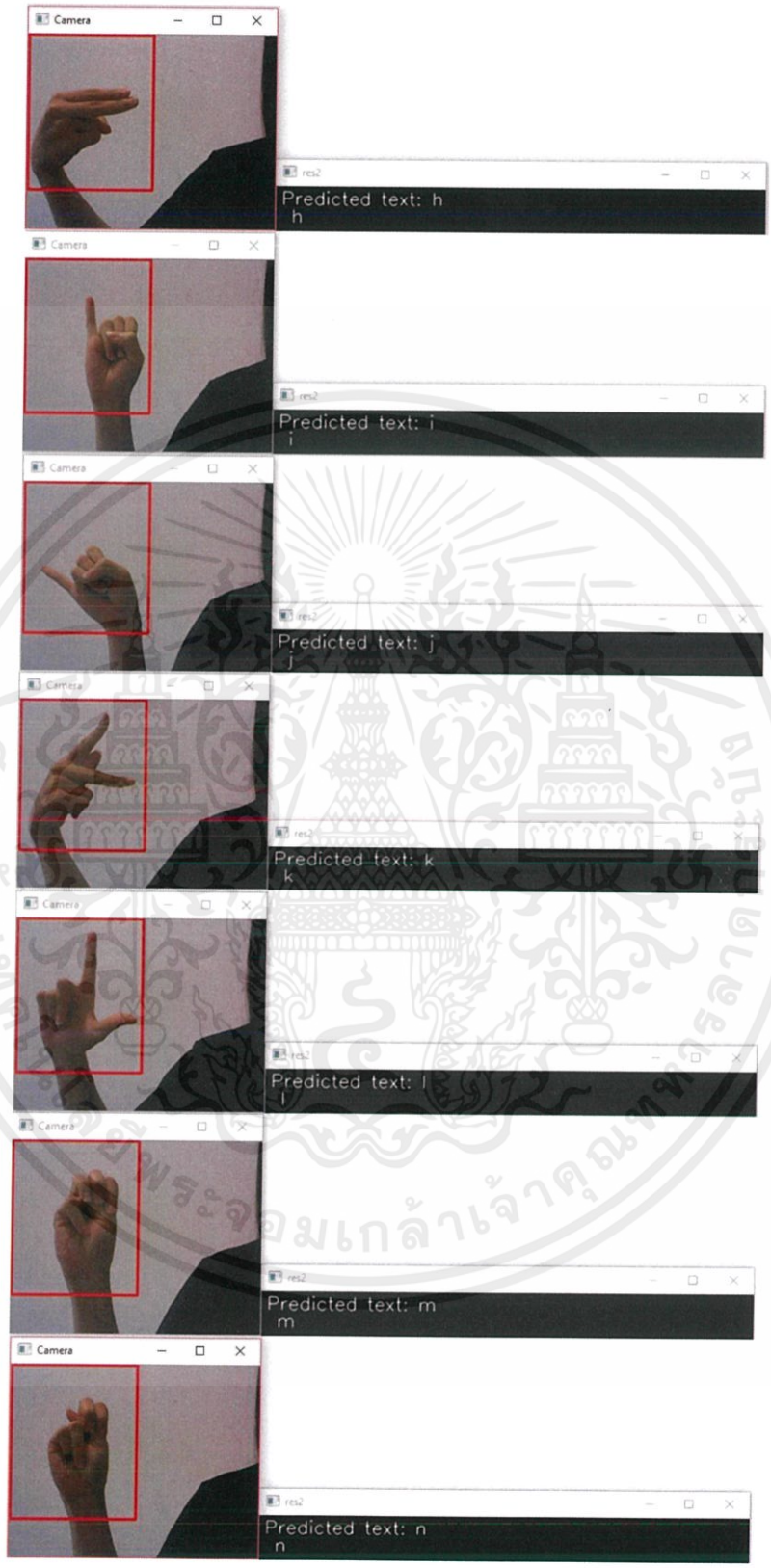
ผลลัพธ์ในการแปลภาษามือตัวอักษร a ถึง g เป็นดังรูปที่ 4.36 ผลลัพธ์ในการแปลภาษามือตัวอักษร h ถึง n เป็นดังรูปที่ 4.37 ผลลัพธ์ในการแปลภาษามือตัวอักษร o ถึง u เป็นดังรูปที่ 4.38 ผลลัพธ์ในการแปลภาษามือตัวอักษร v ถึง z เป็นดังรูปที่ 4.39 และผลลัพธ์ในการแปลเครื่องหมายวรรคตอนต่างๆ คือเครื่องหมายจุลภาค (,), ภาษามือสำหรับลบตัวอักษร, เครื่องหมายมหัพภาค (.), ภาษามือแสดงการส่งข้อความ, เครื่องหมายปรัศนี (?) และเครื่องหมายเว้นวรรคดังรูปที่ 4.40

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



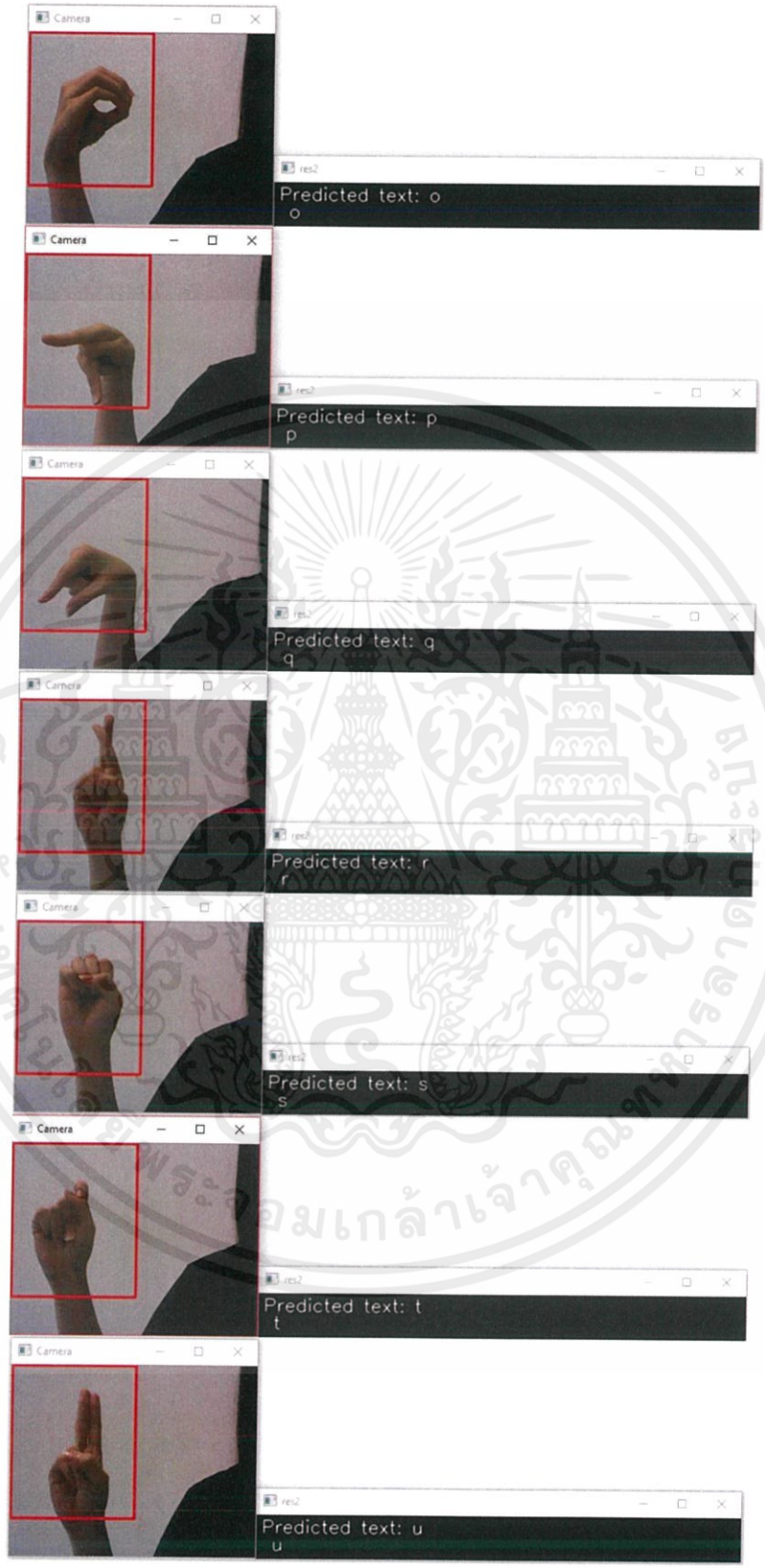
รูปที่ 4.36 ผลลัพธ์ในการแปลภาษามือตัวอักษร a ถึง g

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



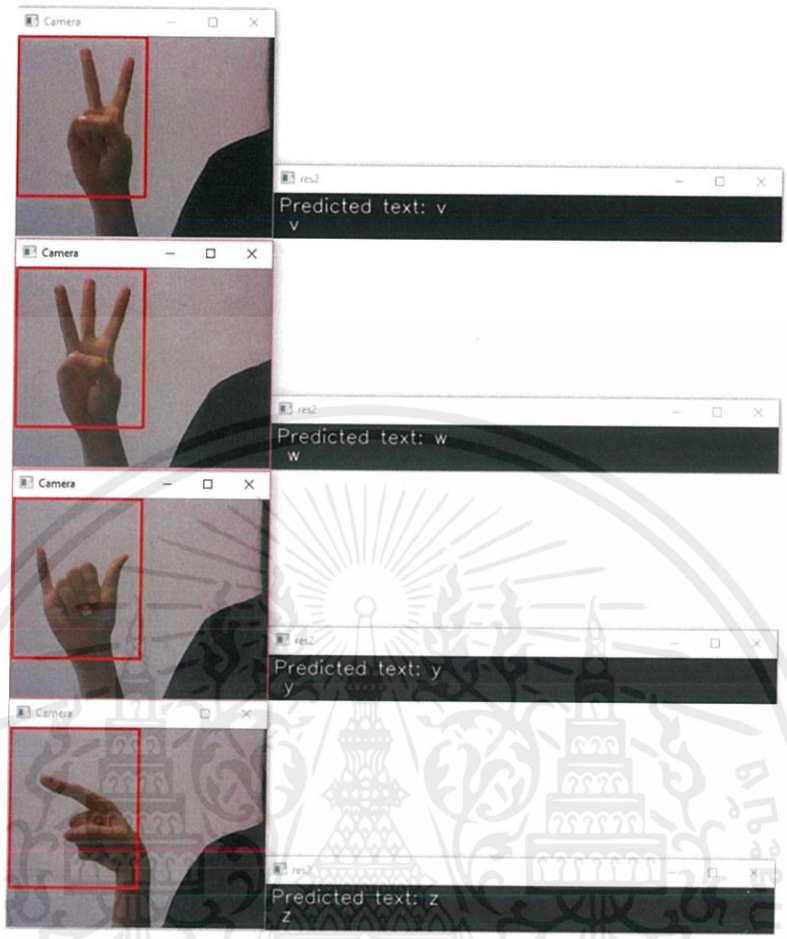
รูปที่ 4.37 ผลลัพธ์ในการแปลภาษามือตัวอักษร h ถึง n

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



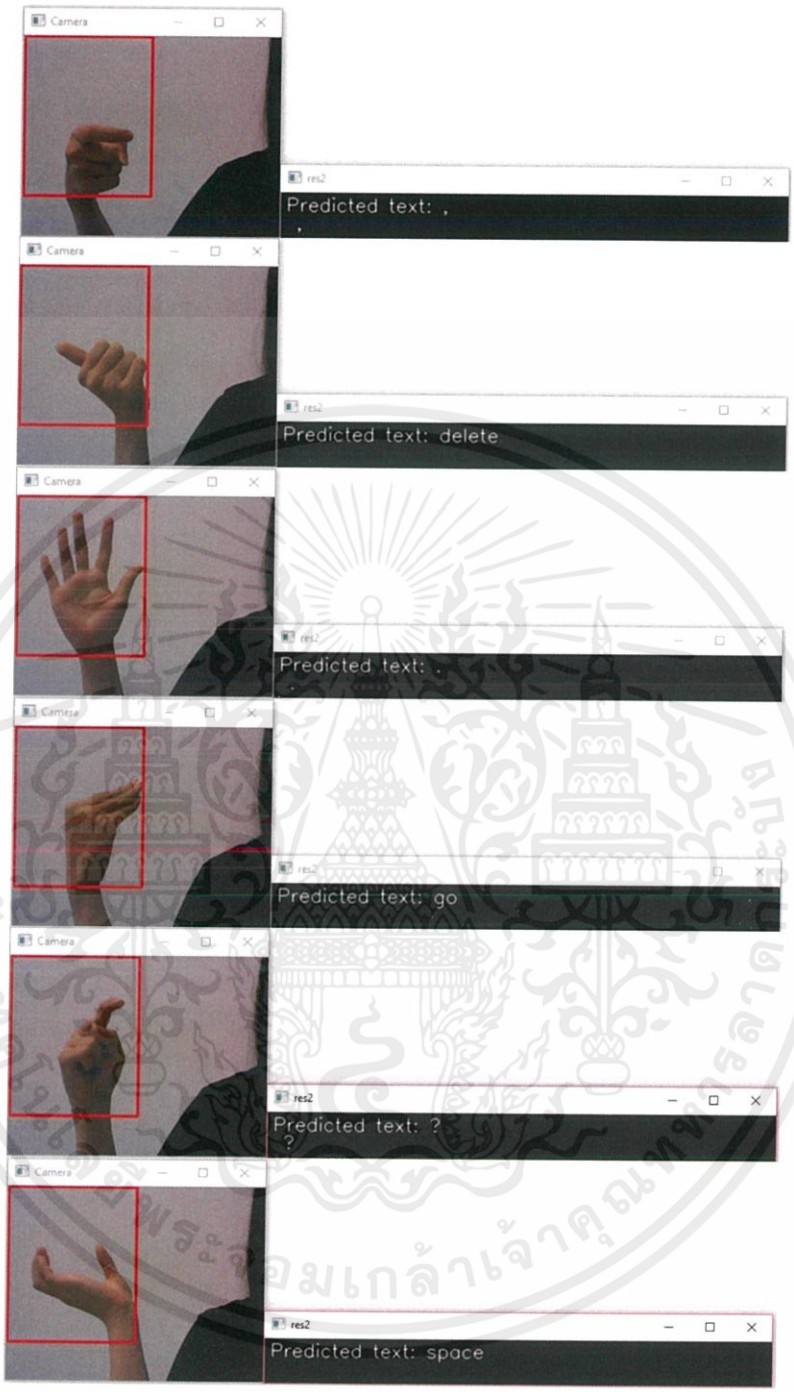
รูปที่ 4.38 ผลลัพธ์ในการแปลภาษามือตัวอักษร o ถึง u

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.39 ผลลัพธ์ในการแปลภาษามือตัวอักษร v ถึง z

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

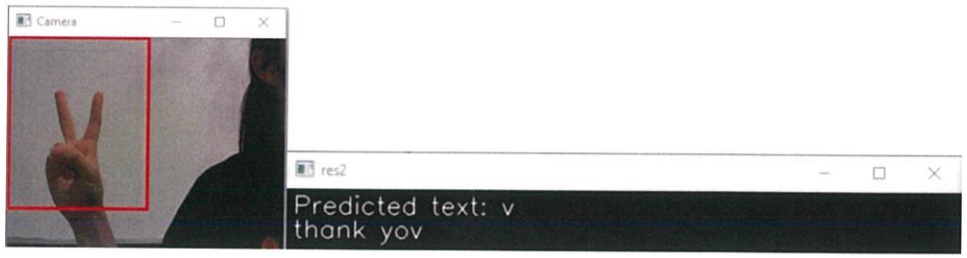


รูปที่ 4.40 ผลลัพธ์ในการแปลภาษามือเครื่องหมายวรรคตอนต่างๆ

4.5.2 ทดสอบการลบตัวอักษรด้วยภาษามือ

ทดสอบการลบตัวอักษรในระบบ ตัวอย่างใช้คำว่า thank yov โดยทำการลบตัวอักษรที่ผิด คือตัวอักษร v และทำการเพิ่มตัวอักษร u ที่ถูกต้องแทน ดังรูปที่ 4.41

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)

(ข)

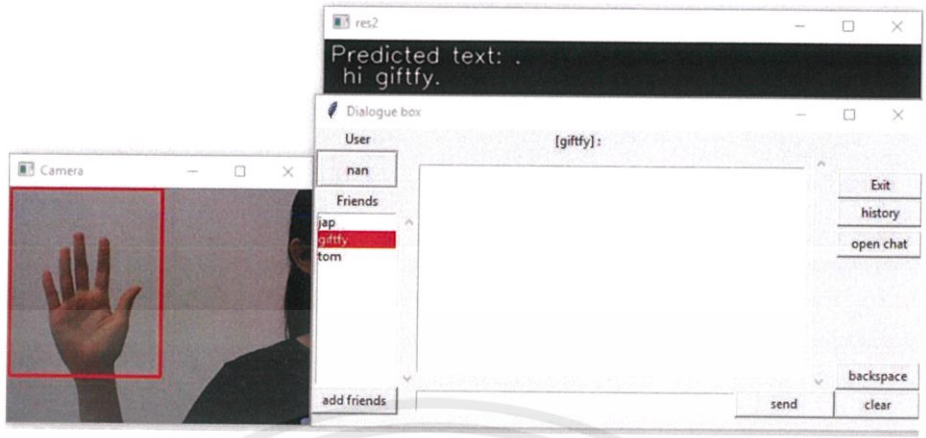
(ค)

รูปที่ 4.41 ตัวอย่างการลบตัวอักษรที่ผิด
(ก) การแปลภาษามือคำว่า thank yov (ข) การแปลภาษามือเพื่อลบตัวอักษร v
(ค) ผลลัพธ์ในการเพิ่มตัวอักษร u ที่ถูกต้อง

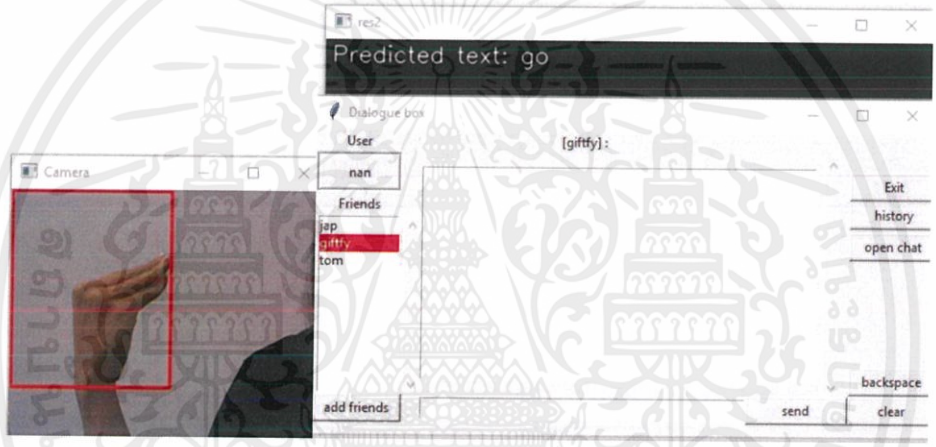
4.5.3 ทดสอบการส่งตัวอักษรด้วยภาษามือกับฐานข้อมูล

ทดสอบการส่งคำว่า hi giftfy. โดยเมื่อทำภาษามือคำว่า hi giftfy. ดังรูปที่ 4.42 (ก) จากนั้นทำสัญลักษณ์แสดงการส่งข้อความดังรูปที่ 4.42 (ข) ข้อความจะถูกบันทึกลงในฐานข้อมูลดังรูปที่ 4.43

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



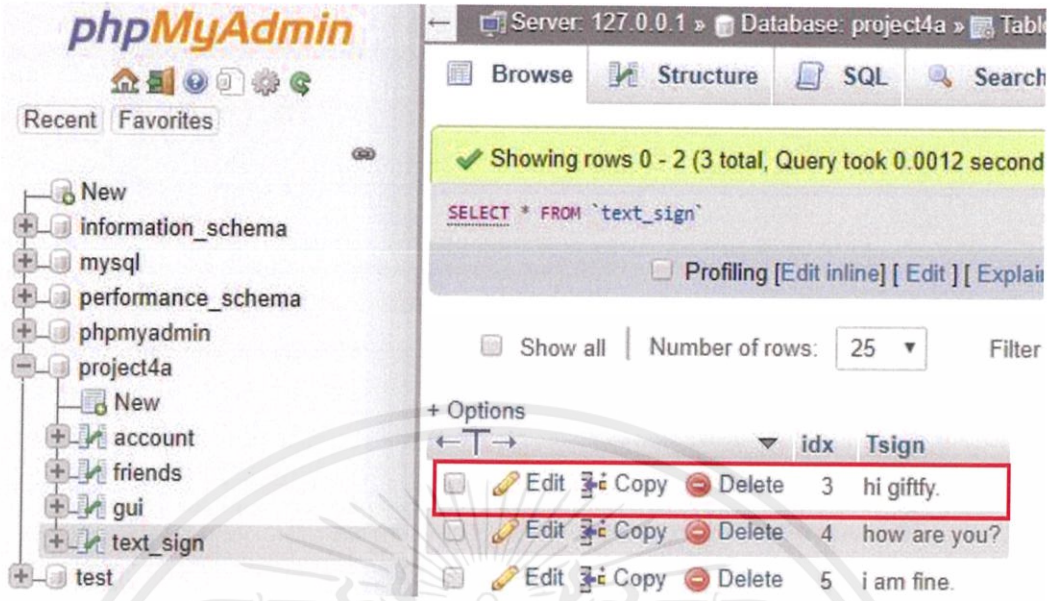
(ก)



(ข)

รูปที่ 4.42 ตัวอย่างการส่งข้อความ (ก) ผลลัพธ์ในการแปลภาษามือคำว่า hi giftfy. (ข) ผลลัพธ์ในการแปลภาษามือแสดงการส่งข้อความ hi giftfy.

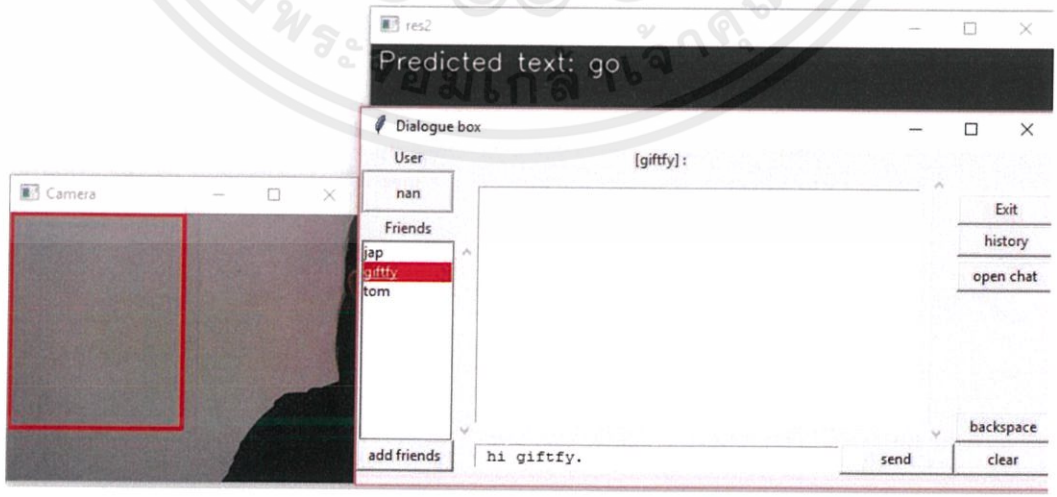
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.43 ข้อความ hi giftfy. ถูกบันทึกลงในฐานข้อมูล

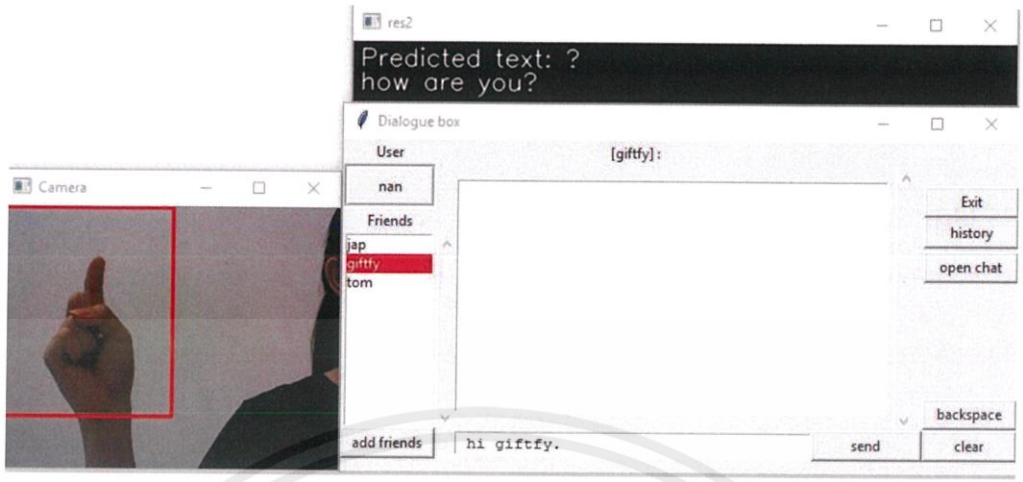
4.5.4 ทดสอบการแสดงผลลัพธ์โปรแกรมแปลภาษามือเป็นตัวอักษรบนส่วนต่อประสานกราฟิกกับผู้ใช้

จากรูปที่ 4.42 (ก) ข้อความ hi giftfy. ในฐานข้อมูลจะแสดงที่หน้าต่าง GUI ดังรูปที่ 4.44 โดยหากมีข้อความใหม่เช่นคำว่า How are you? จากภาษามือดังรูปที่ 4.45 เมื่อกดปุ่ม send ข้อความใหม่ซึ่งคือคำว่า How are you? จะขึ้นทันทีดังรูปที่ 4.46 และข้อความจะถูกบันทึกลงฐานข้อมูลดังรูปที่ 4.47 ถ้าข้อความที่ถูกส่งไปแล้วยังค้างที่หน้าต่าง GUI ให้กดปุ่ม clear ข้อความใหม่จากภาษามือซึ่งคือ I am fine. จะขึ้นแทนที่ดังรูปที่ 4.48 กดปุ่ม Exit เมื่อต้องการบันทึกบทสนทนา และบทสนทนาจะถูกบันทึกลงไฟล์ .txt ดังรูปที่ 4.49

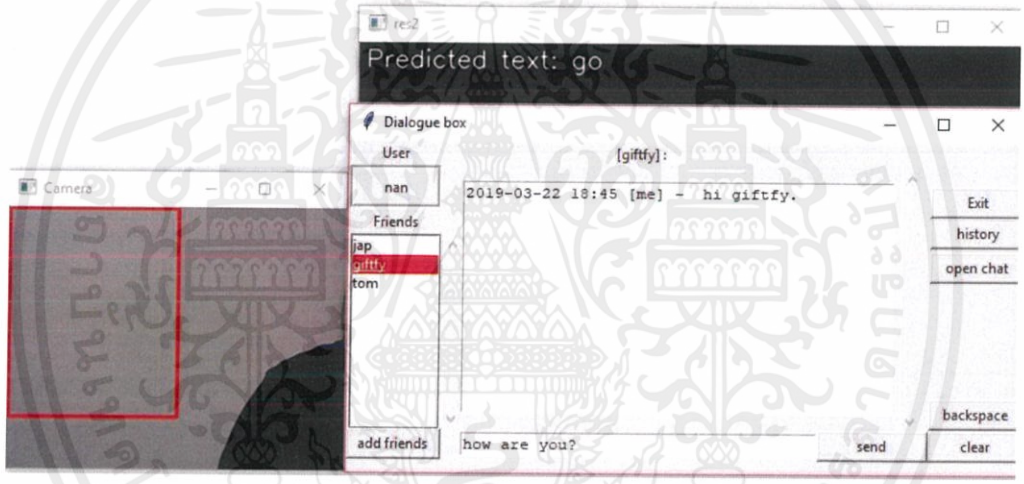


รูปที่ 4.44 ภาษามือคำว่า “hi giftfy.” ที่แสดงผ่านหน้าจอ GUI

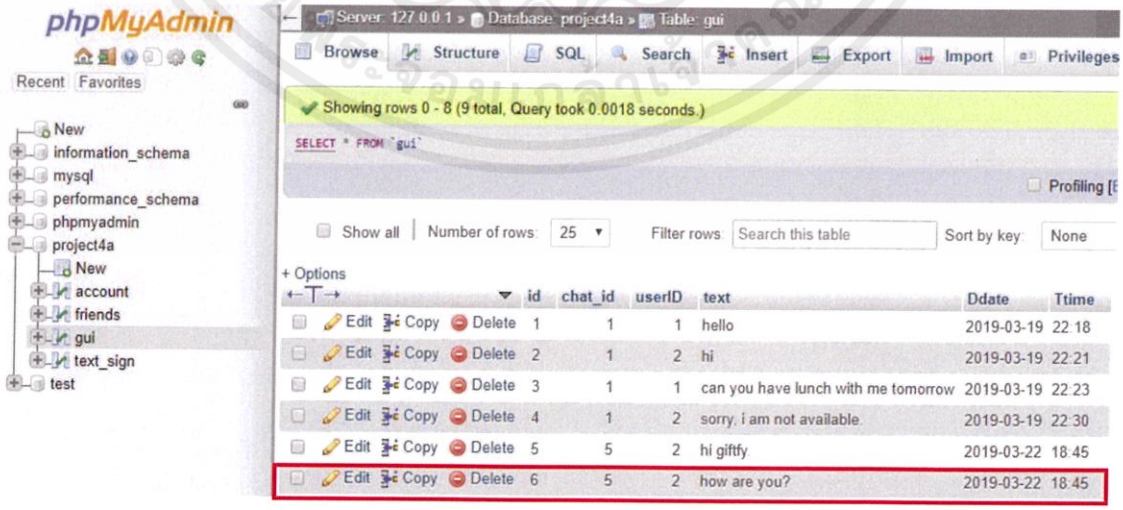
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.45 ผลลัพธ์ในการแปลภาษามือคำว่า How are you?

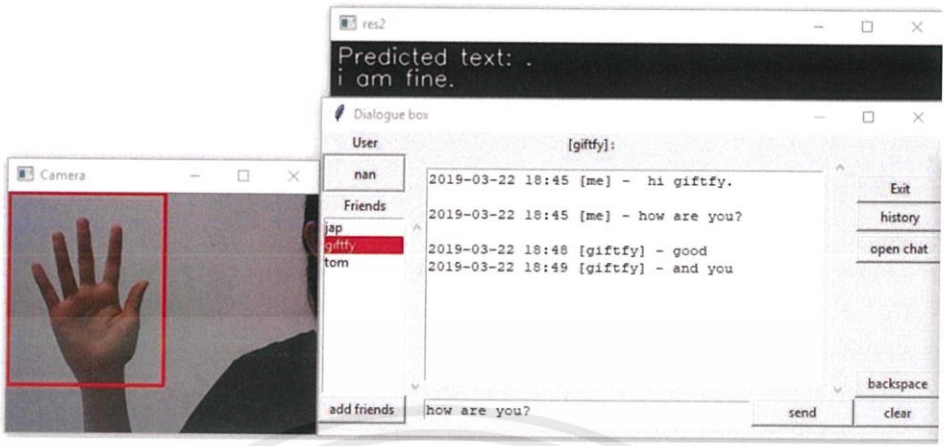


รูปที่ 4.46 ข้อความใหม่ว่า “How are you?” บน Textbox ของ GUI เมื่อส่งคำว่า go

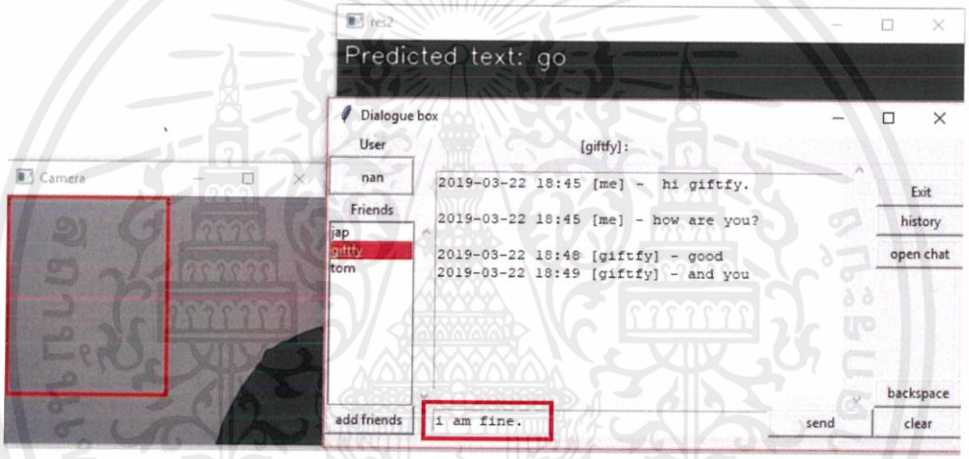


รูปที่ 4.47 ข้อความที่ถูกบันทึกลงฐานข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



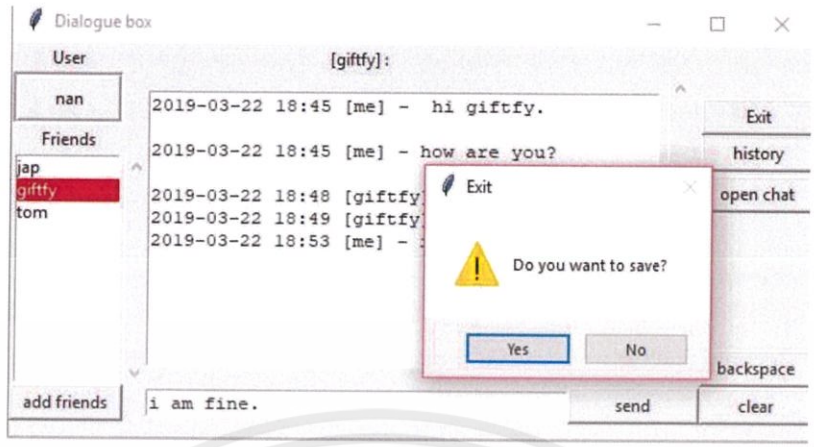
(ก)



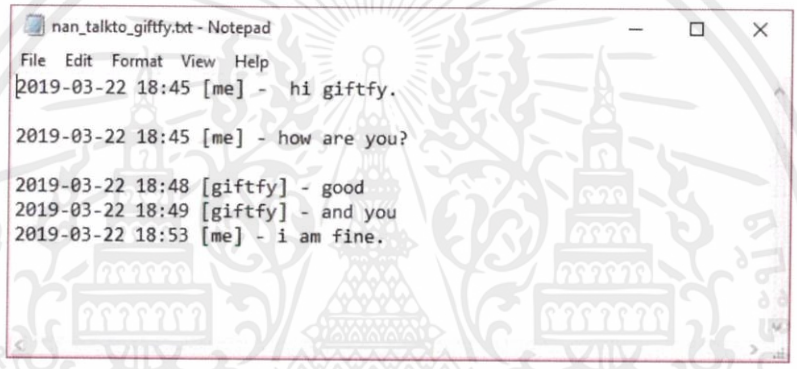
(ข)

รูปที่ 4.48 ตัวอย่างการอัพเดตข้อความใหม่ (ก) ข้อความ “how are you?” ที่ค้างบน Textbox ก่อนกดปุ่ม clear (ข) ข้อความใหม่จากภาษามือหลังกดปุ่ม clear

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)

รูปที่ 4.49 ตัวอย่างการบันทึกบทสนทนา (ก) การกดปุ่ม Exit เมื่อต้องการบันทึกบทสนทนา (ข) บทสนทนาที่ถูกบันทึกลงไฟล์ .txt

4.6 การทดสอบผลลัพธ์โปรแกรมแปลเสียงเป็นตัวอักษร

หน้าต่างสำหรับแปลเสียงเป็นตัวอักษรแสดงดังรูปที่ 4.50 เป็นหน้าต่างสำหรับคนปกติใช้ในการสื่อสารโต้ตอบกับผู้พิการทางการได้ยิน โดยใช้ระบบการแปลเสียงพูดเป็นตัวอักษรจาก Google cloud speech API

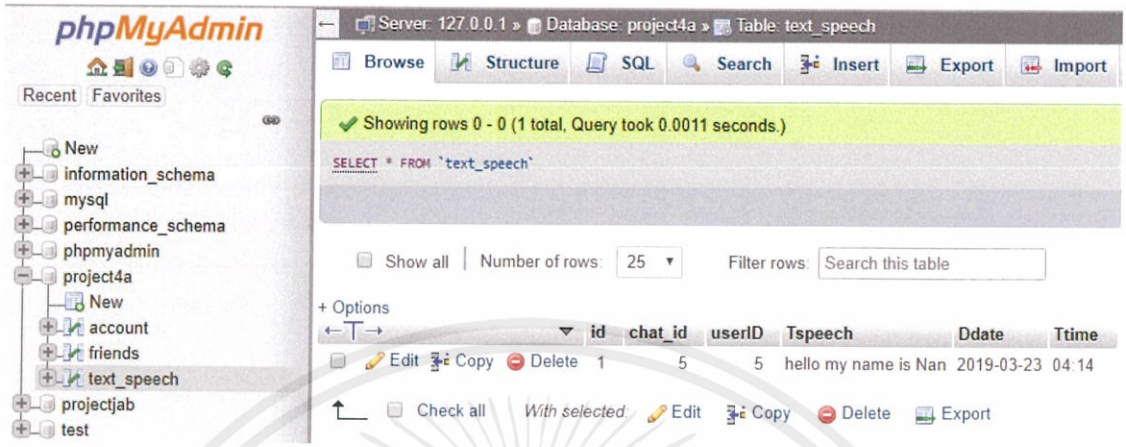
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.50 หน้าต่างสำหรับแปลเสียงพูดเป็นตัวอักษร

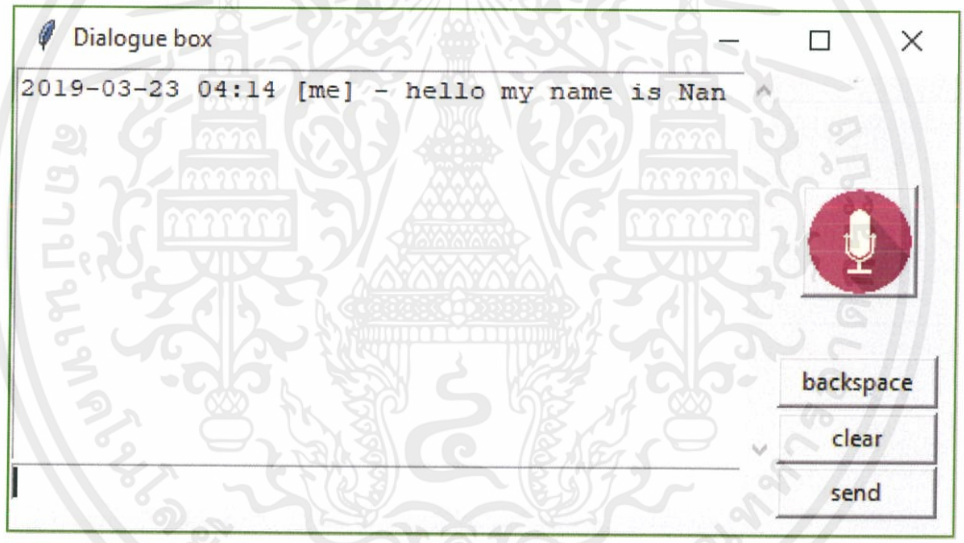
เมื่อต้องการสนทนา ให้กดปุ่มรูปไมโครโฟนในรูปที่ 4.50 แล้วพูดข้อความที่ต้องการสนทนา เช่น hello my name is Nan ระบบจะแปลเสียงพูดเป็นข้อความตัวอักษรแสดงบนพื้นที่แก้ไขข้อความ (edit text) ในส่วนต่อประสานกราฟิกกับผู้ใช้ดังรูปที่ 4.51 ผลลัพธ์ของข้อความที่แปลได้ คือ hello my name is Nan มีความถูกต้องตรงตามเสียงพูด จากนั้นกดปุ่ม send เพื่อส่งข้อความแสดงบนพื้นที่การสนทนา (chat space) ข้อความจะถูกบันทึกลงในฐานข้อมูลดังรูปที่ 4.52 และแสดงบนพื้นที่การสนทนาพร้อมระบุวันที่และเวลาดังรูปที่ 4.53

รูปที่ 4.51 ข้อความที่แปลได้จากเสียงพูด

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.52 ข้อความที่ถูกบันทึกลงฐานข้อมูล



รูปที่ 4.53 ข้อความถูกแสดงบนพื้นที่การสนทนา

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.7 การทดสอบการทำงานของระบบรวม

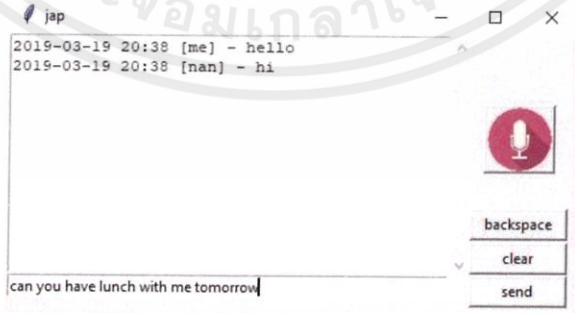
4.7.1 การทดสอบการทำงานของระบบรวมบนคอมพิวเตอร์

หน้าต่างสนทนาของผู้พิการทางการได้ยินเป็นดังรูปที่ 4.54 (ก) เมื่อต้องการสนทนากับบุคคลปกติให้กดปุ่ม open chat หน้าต่างสนทนาของคู่สนทนาขึ้นมาแสดงดังรูปที่ 4.54 (ข)

(ก) (ข)

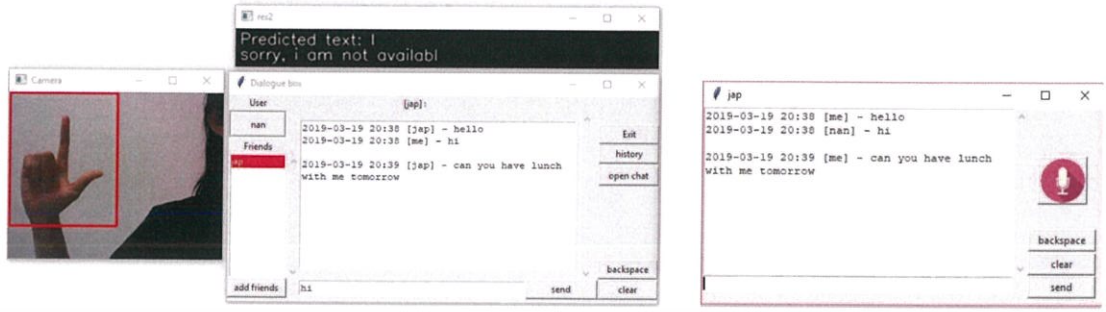
รูปที่ 4.54 หน้าต่างการสนทนา (ก) หน้าต่างสนทนาของผู้พิการทางการได้ยิน (ข) หน้าต่างสนทนาของคู่สนทนา

เมื่อคู่สนทนาต้องการสนทนากับผู้พิการทางการได้ยินให้กดปุ่มรูปไมโครโฟนเพื่อพูด ในตัวอย่างพูดคำว่า Can you have lunch with me tomorrow? ผลลัพธ์การแปลเสียงเป็นตัวอักษรแสดงบนส่วนต่อประสานกราฟิกกับผู้ใช้แสดงดังรูปที่ 4.55 เมื่อกดปุ่ม send คำว่า can you have lunch with me tomorrow จะแสดงที่หน้าต่างสนทนาทั้งของผู้พิการทางการได้ยินและคู่สนทนา ผู้พิการทางการได้ยินตอบกลับโดยใช้ภาษามือคำว่า Sorry, I am not available. ดังรูปที่ 4.56 และทำท่าทางภาษามือเพื่อส่งข้อความ ประโยคว่า Sorry, I am not available. จะถูกบันทึกลงในฐานข้อมูลดังรูปที่ 4.57 และจะแสดงที่ส่วนต่อประสานกราฟิกกับผู้ใช้ดังรูปที่ 4.58



รูปที่ 4.55 ผลลัพธ์การแปลเสียงเป็นตัวอักษร

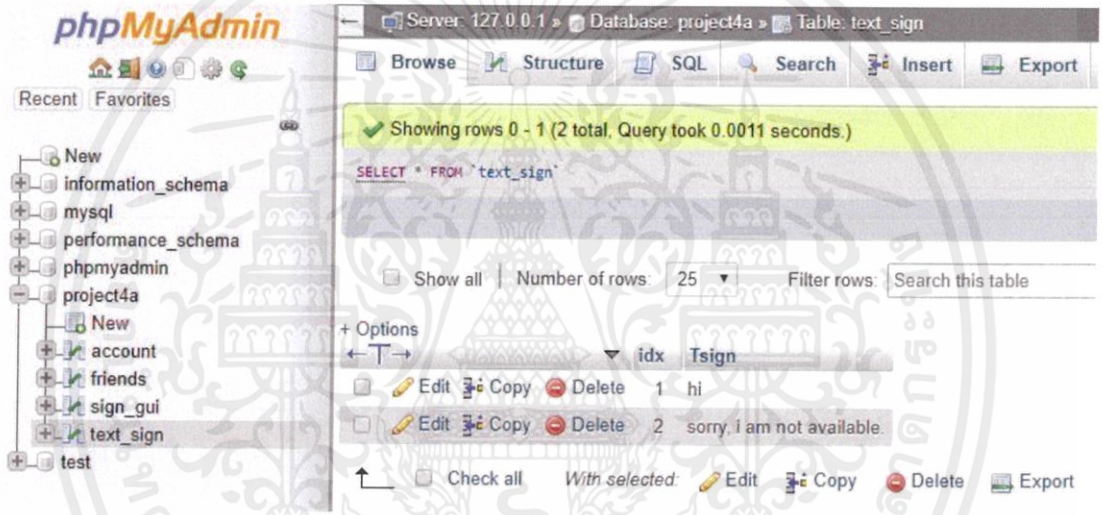
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



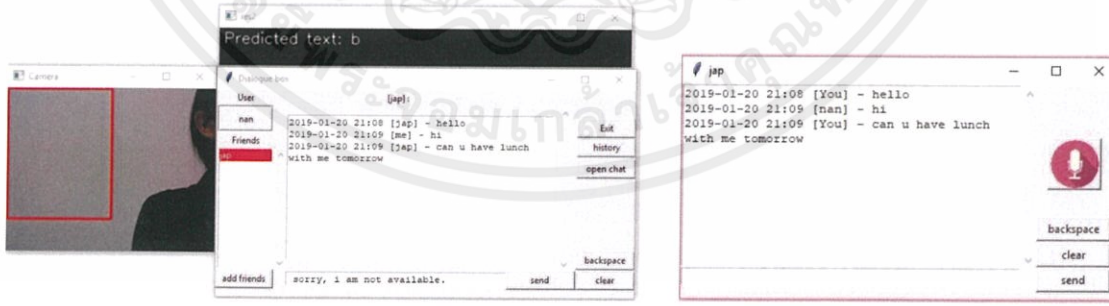
(ก) (ข)

รูปที่ 4.56 ผลลัพธ์การทำท่าทางภาษามือคำว่า “Sorry, I am not available.”

(ก) หน้าต่างของผู้พิการทางการได้ยิน (ข) หน้าต่างของคู่สนทนา



รูปที่ 4.57 ภาษามือที่ถูกบันทึกลงในฐานข้อมูล



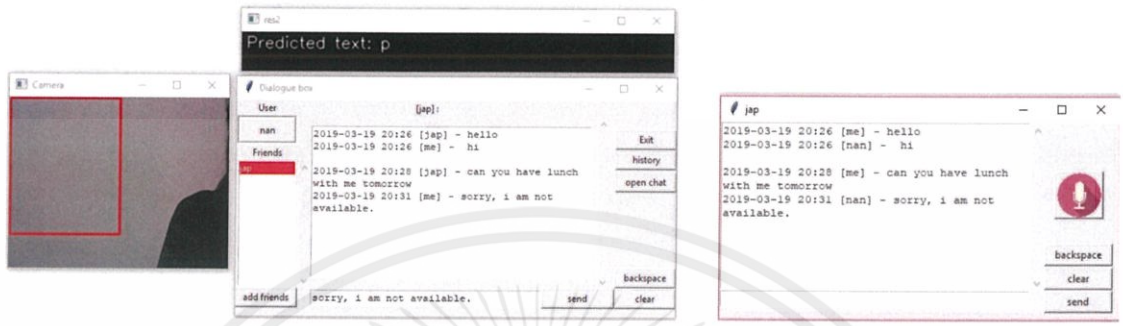
(ก) (ข)

รูปที่ 4.58 ประโยค “Sorry, I am not available.” บนส่วนต่อประสานกราฟิกกับผู้ใช้

(ก) หน้าต่างของผู้พิการทางการได้ยิน (ข) หน้าต่างของคู่สนทนา

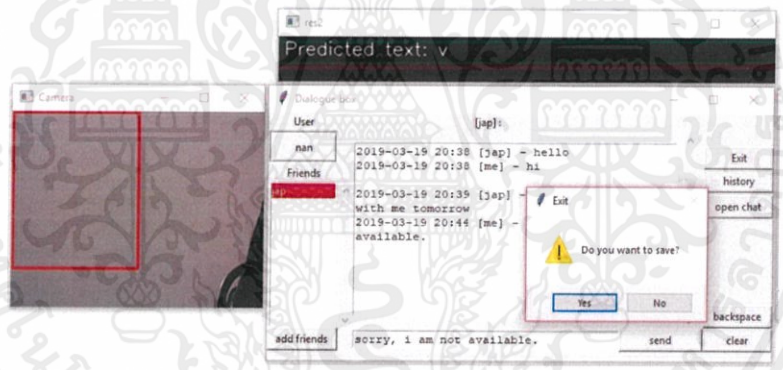
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กดปุ่ม send เพื่อส่งคำว่า “Sorry, I am not available.” บน textbox ให้คู่สนทนาดังรูปที่ 4.59 กดปุ่ม Exit เมื่อต้องการจบการสนทนา บันทึกบทสนทนาโดยเลือก save บทความสนทนาถูกบันทึกเป็นไฟล์ .txt ดังรูปที่ 4.60

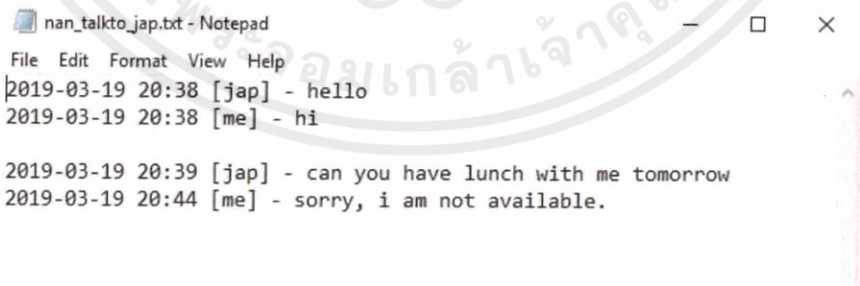


(ก) (ข)

รูปที่ 4.59 การส่งคำว่า “Sorry, I am not available.” บนส่วนต่อประสานกราฟิกกับผู้ใช้ (ก) หน้าต่างของผู้พิการทางการได้ยิน (ข) หน้าต่างของคู่สนทนา



(ก)

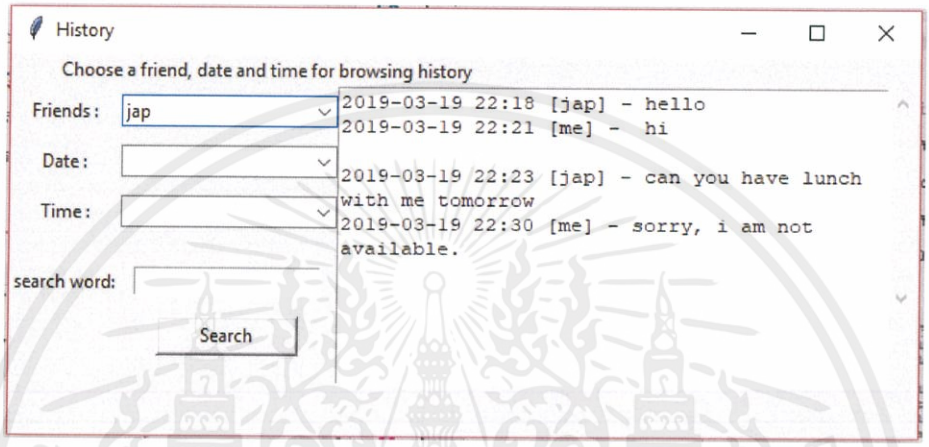


(ข)

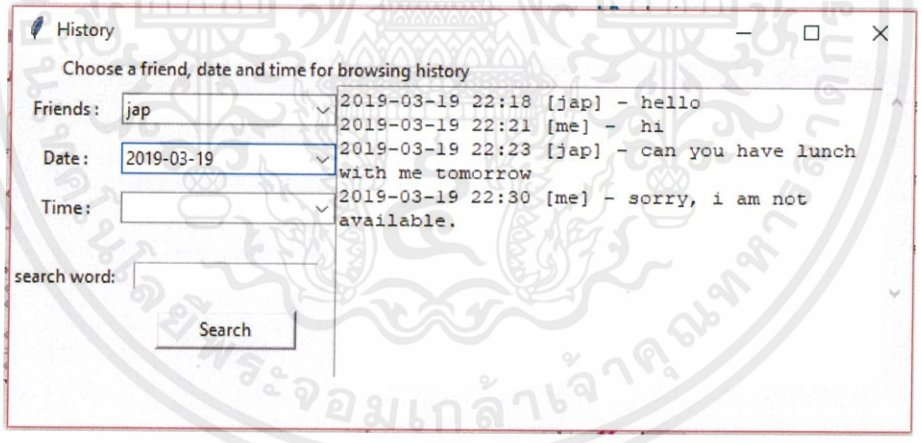
รูปที่ 4.60 การบันทึกข้อความสนทนา (ก) การกดปุ่ม Exit เพื่อจบการสนทนา และการกดปุ่ม save เพื่อบันทึก (ข) ไฟล์ .txt ที่บันทึกข้อความสนทนา

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เรียกดูประวัติการสนทนาโดยการกดปุ่ม history ที่หน้าต่าง Dialogue box โดยเมื่อเลือกชื่อคู่สนทนา ประวัติการสนทนาทั้งหมดจะขึ้นมาที่ช่องประวัติการสนทนาดังรูปที่ 4.61 เมื่อเลือกวันที่จะมีประวัติการสนทนาแค่วันที่ที่เลือกแสดงดังรูปที่ 4.62 และเมื่อเลือกเวลาจะมีประวัติการสนทนาแค่เวลาที่เลือกแสดงดังรูปที่ 4.63 เมื่อต้องการค้นหาคำในตัวอย่างต้องการค้นหาคำว่า can และเมื่อกดปุ่ม search ประวัติการสนทนาที่มีคำว่า can จะแสดงดังรูปที่ 4.64

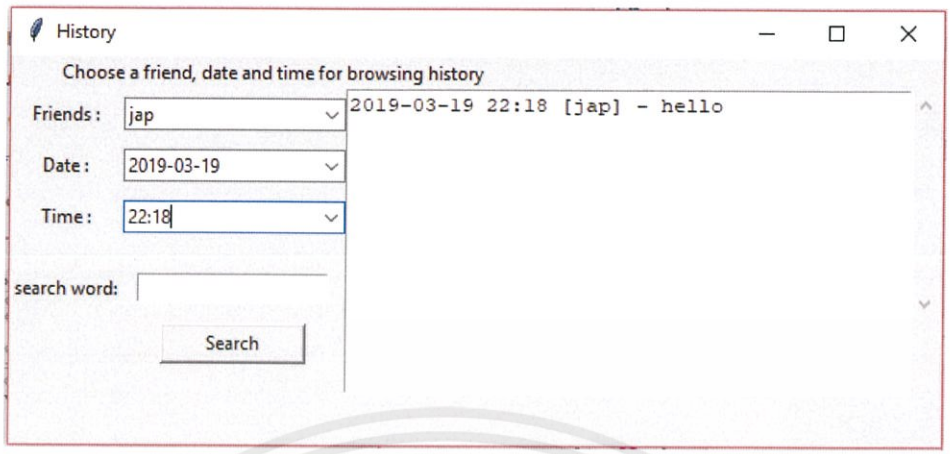


รูปที่ 4.61 ประวัติการสนทนาเมื่อเลือกคู่สนทนา

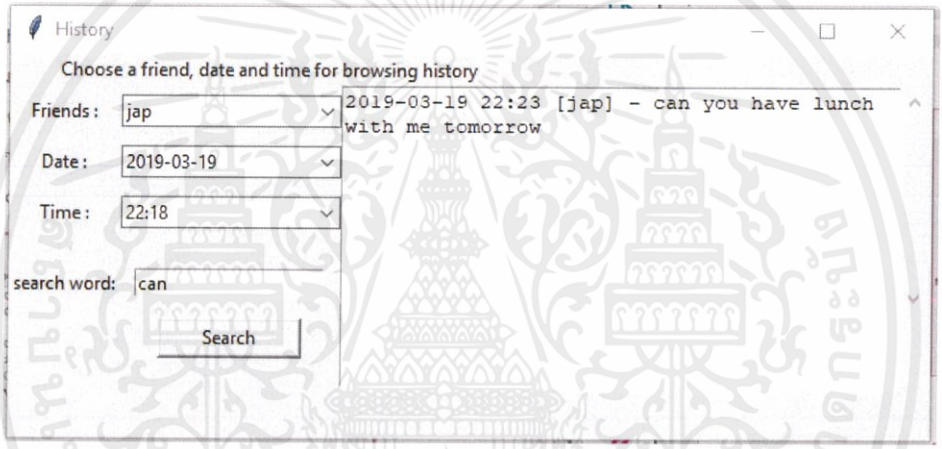


รูปที่ 4.62 ประวัติการสนทนาเมื่อเลือกวันที่

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.63 ประวัติการสนทนาเมื่อเลือกเวลา

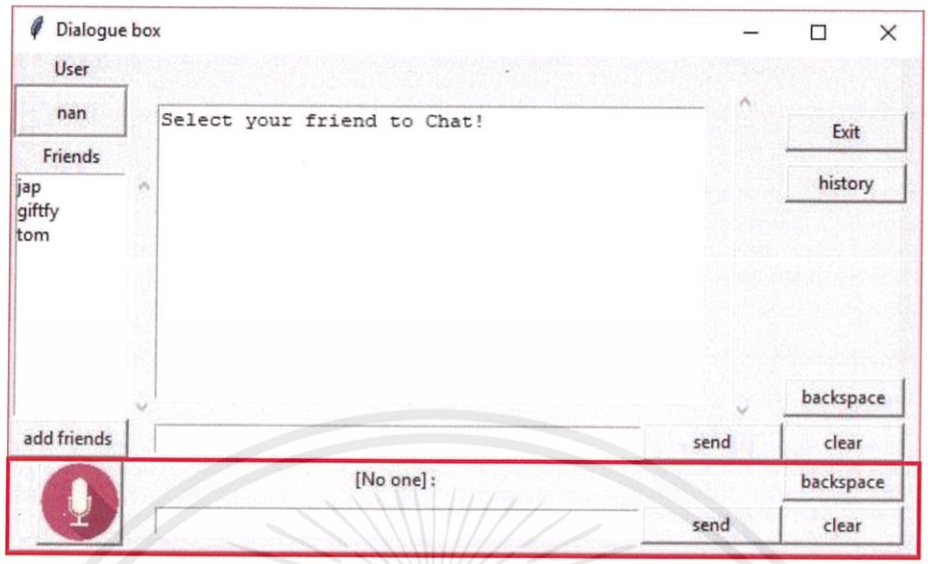


รูปที่ 4.64 ประวัติการสนทนาที่มีคำว่า can

4.7.2 การทดสอบการทำงานของระบบรวมบน Raspberry Pi

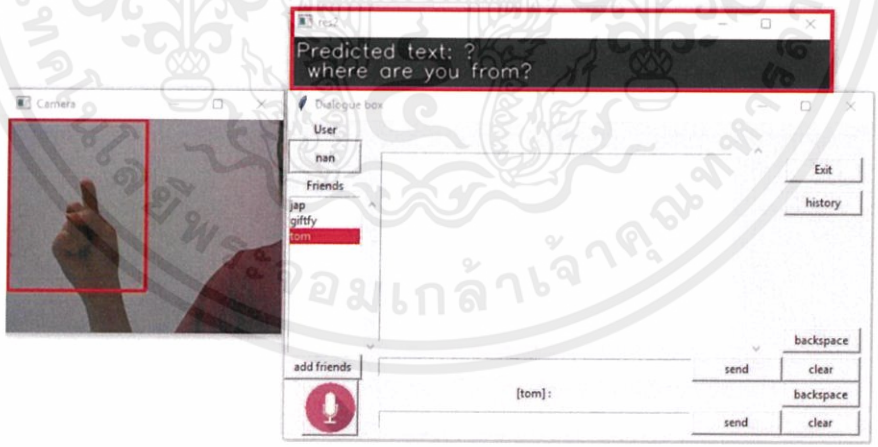
เมื่อทดสอบระบบบน Raspberry Pi แล้วพบว่า Raspberry Pi ไม่สามารถรองรับการตั้งค่าจอ LCD HDMI สองจอแบบขยาย (Extend) จึงใช้การต่อจอ LCD HDMI แบบภาพสะท้อน (Mirror) ซึ่งทั้งสองจอจะแสดงภาพเหมือนกัน และปรับหน้าต่าง Dialogue box โดยเพิ่มช่อง chat window สำหรับคู่สนทนาที่ด้านล่างในกรอบสีแดงดังรูปที่ 4.65 เพื่อไม่ให้หน้าต่าง Dialogue box ที่แสดงบนจอเล็กเกินไป โดยมีปุ่มไมโครโฟนสำหรับพูด ปุ่ม send สำหรับการส่งข้อความ ปุ่ม backspace สำหรับการลบทีละตัวอักษร และปุ่ม clear สำหรับการลบทั้งหมด โดยหน้าต่าง Login หน้าต่าง Add Friend และ History ยังคงเป็นแบบเดิม

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



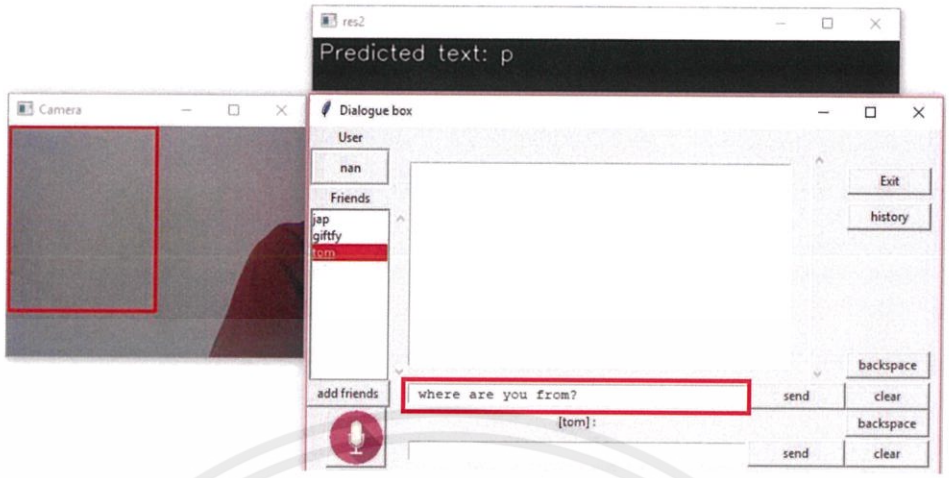
รูปที่ 4.65 หน้าต่าง Dialogue box ที่ได้ปรับปรุง

ผู้พิการทางการได้ยินใช้ภาษามือในการสนทนา ตัวอย่างใช้คำว่า “Where are you from?” ดังรูปที่ 4.66 และแสดงบน textbox ดังรูปที่ 4.67 เมื่อกด send ข้อความจะแสดงที่หน้าต่างการสนทนาดังรูปที่ 4.68 คู่สนทนาตอบกลับโดยการกดปุ่มไมโครโฟน ในตัวอย่างพูดคำว่า “I am from Thailand” แสดงบน textbox ดังรูปที่ 4.69 และเมื่อกด send ข้อความจะแสดงที่หน้าต่างการสนทนาเช่นเดียวกันดังรูปที่ 4.70

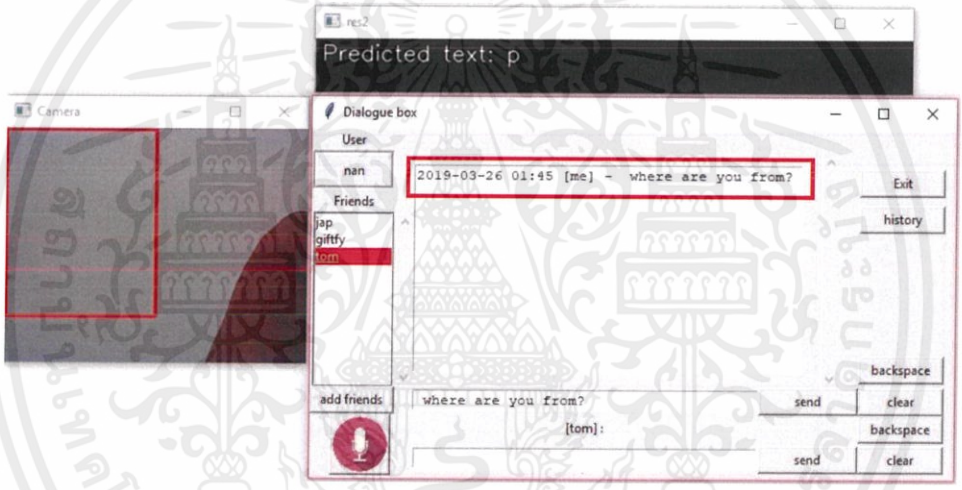


รูปที่ 4.66 ผลลัพธ์ในการแปลภาษามือคำว่า Where are you from?

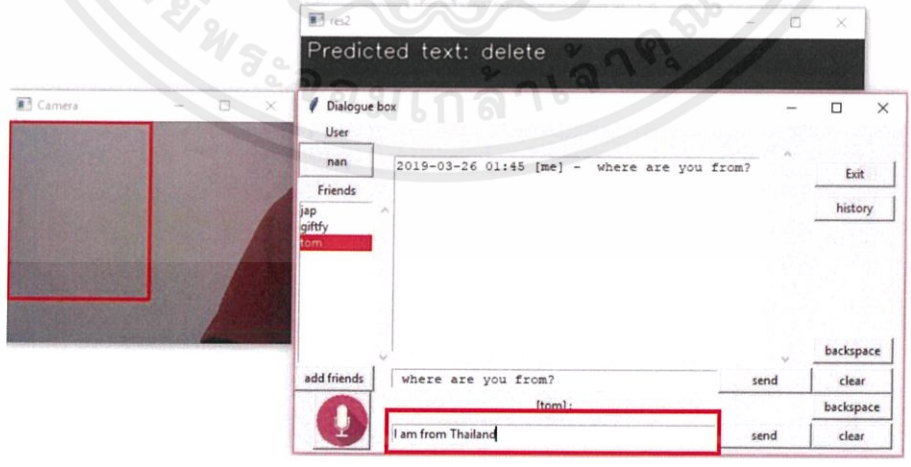
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.67 ข้อความ “Where are you from?” บน Textbox

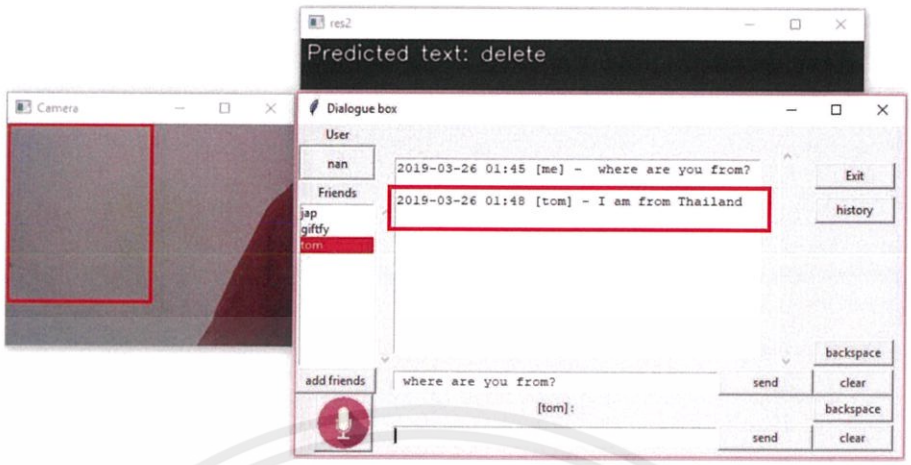


รูปที่ 4.68 ข้อความ “Where are you from?” บนส่วนต่อประสานกราฟิกกับผู้ใช้



รูปที่ 4.69 ข้อความ “I am from Thailand” จากเสียงพูดบน Textbox

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.70 ข้อความ “I am from Thailand” บนส่วนต่อประสานกราฟิกกับผู้ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1 สรุปผล

ปริญญานิพนธ์นี้มีเป้าหมายเพื่อจัดทำอุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยินเพื่อลดความไม่เท่าเทียมของผู้พิการในการสนทนากับบุคคลทั่วไป โดยระบบที่ออกแบบและพัฒนาขึ้นใช้การรับภาษามือจากผู้พิการทางการได้ยินผ่านกล้อง และตอบกลับโดยรับเสียงจากคู่สนทนาด้วยไมโครโฟน ซึ่งทั้งสองส่วนประมวลผลโดยใช้ Raspberry Pi โดยมีการเก็บบันทึกการสนทนาในฐานข้อมูล และแสดงผลการสนทนาผ่านจอแสดงผล โดยผู้พิการทางการได้ยินสามารถทำท่าทางภาษามือทีละตัวอักษรเรียงต่อกันเป็นประโยคเพื่อสนทนา โดยภาษามือที่ระบบเข้าใจ คือภาษามือแบบอเมริกัน โดยผู้พิการทางการได้ยินสามารถทำท่าทางภาษามือเพื่อสนทนาผ่านกล้อง เมื่อผ่านการประมวลผลภาพ การหาคุณลักษณะของภาพด้วยวิธี Histogram of gradients (HOG) และใช้อัลกอริทึม SVM แบบ multi-class ในการจำแนกตัวอักษรภาษามือ ข้อความที่แปลได้จากภาษามือจะแสดงผ่านหน้าต่างการสนทนา ในส่วนการตอบกลับจากคู่สนทนาสามารถทำได้โดยการกดปุ่มไมโครโฟนที่ส่วนต่อประสานกราฟิกกับผู้ใช้ ซึ่งใช้ Google Cloud Speech API ในการแปลเสียงเป็นข้อความ ซึ่งข้อความที่แปลได้จากเสียงจะแสดงที่หน้าต่างการสนทนาเช่นกัน จากการทดสอบระบบสามารถทำงานได้ดังที่ออกแบบไว้

5.2 ข้อเสนอแนะ

อุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยินที่ได้ออกแบบและจัดทำขึ้นในปริญญานิพนธ์นี้ ในส่วนการรับภาษามือจากผู้พิการทางการได้ยินผ่านกล้องนั้น ต้องกดปุ่มเพื่อลบพื้นหลัง ซึ่งต้องกดทุกครั้งเมื่อย้ายสถานที่ การแปลภาษามือเป็นตัวอักษรอาจมีบางตัวอักษรที่ผิดพลาด รวมถึงภาษามือที่ใช้สนทนาไม่รองรับภาษามือทุกประเทศ ซึ่งอาจมีผู้พิการทางการได้ยินที่ไม่สามารถสนทนาเป็นภาษาอังกฤษได้ และหากใช้จอต่อ Raspberry Pi ที่สามารถรองรับการตั้งค่าจอแบบขยาย (Extend) ผู้พิการทางการได้ยินและคู่สนทนาจะเห็นหน้าต่างสนทนาต่างกัน ซึ่งจะลดความสับสนได้ ข้อด้อยดังกล่าวของอุปกรณ์ช่วยการสนทนาสำหรับผู้พิการทางการได้ยินนี้สามารถนำไปปรับปรุงและพัฒนาต่อยอดให้มีความเหมาะสมและสะดวกต่อการใช้งานมากยิ่งขึ้นในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] Signing Savvy. "Sign Language." https://www.signingsavvy.com/sign.

[2] RASPBERRY PI FOUNDATION. "RASPBERRY PI 3 MODEL B+." https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus.

[3] ทวีรัตน์ นวลช่วย. "Python programming." https://sites.google.com/site/dotpython/installation.

[4] Softonic International S.A. "PyCharm Varies with device." https://pycharm.th.softonic.com/#app-softonic-review.

[5] Kan Ouivirach. "ดู Data Structure ของรูปภาพโดยใช้ OpenCV for Python." https://medium.com/prontotools/%E0%B8%94%E0%B8%B9-data-structure.

[6] Dan Aharon. "Cloud Speech API is now generally available." https://cloudplatform.googleblog.com/2017/04/Cloud-Speech-API-is-now-generally-available.html.

[7] Y. Jaruwan. "ความรู้ทั่วไปเกี่ยวกับระบบฐานข้อมูล." http://www3.chandra.ac.th/office/ict/document/it/it04/page01.html.

[8] บริษัทเอโอซอฟต์จำกัด. "phpMyAdmin คืออะไร." https://www.aosoft.co.th/article/310/phpMyAdminA3.html.

[9] นางสาวอรวรรณ เลียบศิริ. "การจำแนกอารมณ์ตัวการ์ตูนด้วยฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์." http://ethesisarchive.library.tu.ac.th/thesis/.

[10] ผศ.ดร.มหศักดิ์ เกตุฉ่ำ. "การสกัดคุณลักษณะเฉพาะของภาพ." https://bit.ly/2TBL27T.

[11] sit. "เทคนิค Support Vector Machine (SVM) และซอฟต์แวร์ HR-SVM." http://dataminingtrend.com/2014/support-vector-machine-svm.

[12] inform. Support Vector Machines. https://informy.tumblr.com/post/17937920903/support-vector-machines.

[13] Pushkar Mandot. "What is the Significance of C value in Support Vector Machine." https://medium.com/@pushkarmandot/what-is-the-significance-of-c-value-in-support-vector-machine.

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[14] Chris Albon. "SVC Parameters When Using RBF Kernel." https://chrisalbon.com/machine_learning/support_vector_machines/svc_parameters_using_rbf_kernel/.

[15] Rimah Amami, Dorra Ben Ayed, Noureddine Ellouze. "Practical Selection of SVM Supervised Parameters with Different Feature Representations for Vowel Recognition." https://bit.ly/2Z1XVHM.

[16] Ruslan K.,Ho P., and Ng J. "Sign Language Classification Using Webcam Images." https://stanford.io/2VBLd0g.

[17] bhedavivek. "A Deep Learning Model to Classify ASL Alphabets." https://github.com/bhedavivek/deep_asl/tree/master/dataset.

[18] Shubham Gupta. "ASL-Translator." https://github.com/mjk188/ASL-Translator/tree/master/TrainData.

[19] Renuka Joshi. "Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures." https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures.

[20] Google Cloud (2019). "Cloud Speech-to-Text basics/Cloud Speech API Documentation/Google Cloud." https://cloud.google.com/speech-to-text/docs/basics.

[21] Arduitronics. "USB Mini Microphone for Raspberry Pi (also PC และ Laptop)." https://www.arduitronics.com/product/1723/usb-mini-microphone-for-raspberry-pi-also-pc.

[22] mosfex. "จอ 5 นิ้ว Touch Screen." https://www.mosfex.com/ product/26/

[23] Logitech. "C170 WEBCAM." https://www.logitech.com/th-th/product /webcam-c170.

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้