

การทดสอบและการประเมินผล

ในบทนี้จะกล่าวถึงวิธีการทดสอบ และการประเมินผล เพื่อให้ทราบถึงประสิทธิภาพของระบบรับรู้เสียงพูดแบบต่างบุคคลที่พัฒนาขึ้น การประเมินผลจะเน้นถึงความถูกต้องในการรับรู้

5.1 ข้อมูลทดสอบ

คำที่ใช้ในการทดสอบได้แบ่งเป็น 2 กลุ่ม ซึ่งอาศัยเกณฑ์ในการคัดเลือกต่างกัน คือ

ก. กลุ่มที่ 1 เป็นคำที่ได้จากการอาศัยหลักเกณฑ์ทางด้านภาษาศาสตร์เป็นหลัก คำที่คัดเลือกมาทดสอบในกลุ่มนี้มี 15 คำ ได้แก่ คำว่า ปาก พาด บาง ฟาง มาก ปิด ผิด บิน ผืน มิติ ปั้น ฬั บับ ฬี่ มี รายละเอียดวิธีการคัดเลือกคำในกลุ่มนี้อยู่ในภาคผนวก ก.

ข. กลุ่มที่ 2 เป็นคำที่ตัดมาจากรายงานการวิจัยเรื่อง การประมวลผลข้อมูลภาษาไทยด้วยไมโครคอมพิวเตอร์ ของยี่น กู่วรรณและคณะ ในรายงานนี้ได้มีการจัดทำคำที่ใช้มากที่สุดในชีวิตประจำวัน คำที่คัดเลือกมาทดสอบในกลุ่มนี้ คือคำที่มีความถี่ในการใช้สูงสุด 10 อันดับแรก ได้แก่ คำว่า ที่ การ เป็น ได้ จะ ไน ไม่ ก็ ของ ให้

ได้มีการเก็บเสียงพูดเพื่อนำไปสร้างแบบอ้างอิง จากนักศึกษาชายทั้งหมด 30 คน โดยให้นักศึกษาแต่ละท่านพูดคำทั้ง 2 กลุ่มคำละ 1 ครั้ง จะทำให้ได้แบบอ้างอิงทั้งหมดของคำในกลุ่มที่ 1 จำนวน 450 แบบ และกลุ่มที่ 2 จำนวน 300 แบบ จากแบบอ้างอิงทั้งหมดจะนำมาผ่านโปรแกรมการคัดเลือกแบบอ้างอิง เพื่อให้ได้แบบอ้างอิงในแต่ละคำเพียง 10 แบบ ดังนั้นจะทำให้กลุ่มของแบบอ้างอิงลดลง โดยในกลุ่มที่ 1 จะเหลือเพียง 150 แบบ และกลุ่มที่ 2 จะเหลือ 100 แบบ ในส่วนของข้อมูลที่ใช้ในการทดสอบ จะได้จากนักศึกษาชาย 5 คน ซึ่งเป็นคนละกลุ่มกับที่นำมาสร้างแบบอ้างอิง ให้นักศึกษาทั้ง 5 ท่านทำการพูดคำทั้ง 2 กลุ่มคำละ 10 ครั้ง จากนั้นแบบทดสอบทั้งหมดจะถูกนำมาผ่านขั้นตอนต่าง ๆ เพื่อทดสอบการรับรู้ต่อไป

ในการเก็บเสียงพูดเพื่อนำไปสร้างเป็นแบบอ้างอิง หรือแบบทดสอบนั้น จำเป็นต้องมีการกำหนดความดังของเสียงที่พูด เนื่องจากความดังของเสียงจะมีผลโดยตรงต่อความครบถ้วนของข้อมูล คือถ้าเสียงพูดมีความดังเกินไป ข้อมูลที่ได้จากการสุ่มตัวอย่าง (sampling) จะมีค่าสูงเกินไป ทำให้ข้อมูลที่ควรจะมีค่าต่ำหายไป และในทางกลับกันถ้าเสียงพูดเบาเกินไป ข้อมูลที่ได้ก็

จะขาดในส่วนที่ควรจะมีค่าสูงไป การควบคุมความดังของเสียงสามารถทำได้โดยอาศัย โปรแกรม SIGNAL EDITOR ซึ่งเป็นโปรแกรมที่พัฒนาขึ้นเพื่อที่จะนำข้อมูลสัญญาณเสียง แสดงขึ้นบนจอภาพ พร้อมกับสเกล (Scale) เพื่อให้ทราบถึงความดังของเสียงได้

5.2 ผลการทดสอบ

ในการทดสอบการรับรู้จากนักศึกษาชาย 5 คน โดยนำไปผ่านขั้นตอนการรับรู้บนเครื่อง ไมโครคอมพิวเตอร์ขนาด 16 บิต เริ่มต้นด้วยการเปรียบเทียบแบบไดนามิค โปรแกรมแกรมมิ่ง และกฎการตัดสินใจเป็น K-nearest Neighbour (KNN) แบบที่ 2 โดยกำหนดค่า K เป็น 3 ได้อัตราความถูกต้องในการรับรู้เสียงพูดของคำในกลุ่มที่ 1 เป็นร้อยละ 75.7 และคำในกลุ่มที่ 2 เป็นร้อยละ 89.4 โดยมีรายละเอียดการกระจายความถูกต้องของการรับรู้ในตารางที่ 5.1 และ 5.2 ตามลำดับ


จากตารางที่ 5.1 และ 5.2 จะพบว่าอัตราความถูกต้องในการรับรู้มีค่าไม่สูงนัก ทั้งนี้เนื่องจากคำที่ใช้ในการทดสอบมีเสียงค่อนข้างใกล้เคียงกัน คือมีเสียงสระ พยัญชนะท้าย และวรรณยุกต์เสียงเดียวกัน เช่นคำว่า "ปาก" กับ "มาก" เป็นต้น จากการสังเกตค่าระยะทางของผลลัพธ์ที่ได้ซึ่งไม่ถูกต้อง ส่วนใหญ่มีค่าใกล้เคียงกับผลลัพธ์ที่ควรจะเป็น ดังนั้นจึงมีการกำหนดค่าความแตกต่างของระยะทางสำหรับผลลัพธ์ที่มีระยะทางที่น้อยที่สุดกับระยะทางของผลลัพธ์ถัดขึ้นมา ค่าความต่างนี้เรียกว่า Reject Value ผลลัพธ์ที่ถูกต้องจะต้องมีค่าความแตกต่างของระยะทางกับค่าของผลลัพธ์ถัดไปมากกว่าค่า Reject Value ที่กำหนดไว้ ถ้าค่าความแตกต่างมีค่าน้อยกว่าค่า Reject Value นี้จะถือว่าการรับรู้สำหรับคำนั้นไม่สามารถทำได้ อัตราความถูกต้องในการรับรู้หลังการปรับปรุงของคำในกลุ่มที่ 1 จะมีค่าเพิ่มเป็นร้อยละ 85.8 และความถูกต้องในกลุ่มที่ 2 เพิ่มเป็น 94.3 รายละเอียดการกระจายความถูกต้องของการรับรู้แสดงในตารางที่ 5.3 และ 5.4 ตามลำดับ

จากตารางที่ 5.3 จะพบว่า การรับรู้สำหรับคำบางคำยังมีความสับสนอยู่ เช่นคำว่า "ปาก" กับ "มาก" เป็นต้น ดังนั้นจึงได้มีการเพิ่มขั้นตอนขึ้นหลังจากที่ผ่านการตัดสินใจมาแล้ว โดยการทำ Post Processing ซึ่งในขั้นตอนนี้จะได้มาจากการทดลอง โดยทำการหาลักษณะ (Feature) เฉพาะคำที่มีความสับสนเพื่อหาความแตกต่างของคำเหล่านั้น ในที่นี้สำหรับคำว่า "ปาก" กับ "มาก" จะใช้ค่า Zero Crossing เป็นลักษณะในการจำแนกความแตกต่าง จากการทดลองพบว่าค่า Zero Crossing ของคำทั้งสองในช่วง 10 เฟรมแรกมีค่าแตกต่างกันอย่างชัดเจน โดยคำว่า "ปาก" จะมีค่าเฉลี่ยต่ำกว่า 20 และคำว่า "มาก" มีค่าเฉลี่ยสูงกว่า 25 จากข้อมูลที่ได้นำมากำหนดเงื่อนไขในการทำ Post Processing ของคำทั้งสอง ดังนี้

ก. ขั้นตอน Post Processing จะทำเมื่อ ผลลัพธ์ที่ได้จากการตัดสินใจเป็นคำว่า "ปาก" และผลลัพธ์ถัดไปเป็นคำว่า "มาก" หรือในทางกลับกันเท่านั้น

ข. จะทำการหาค่า Zero Crossing ในช่วง 10 เฟรมแรกของสัญญาณ ค่าเฉลี่ยที่ได้จะนำมาผ่านกฎ ดังนี้

1. ถ้าค่า Zero Crossing มีค่าต่ำกว่า 20 ให้แสดงผลเป็นคำว่า "ปาก"
2. ถ้าค่า Zero Crossing มีค่าสูงกว่า 25 ให้แสดงผลเป็นคำว่า "มาก"
3. ถ้าค่า Zero Crossing มีค่าอยู่ในช่วง 20 ถึง 25 ให้ปฏิเสธแบบทดสอบนี้



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

แบบอ้างอิง แบบทดสอบ	ปาก	พาด	บาง	ฟาง	มาก	ปิด	ผิด	บิน	ฝั่ง	มิด	บัน	ง	บับ	ง	ม	Rej.	%
ปาก	35			2	12	1										0	70
พาด	2	40		3	5											0	80
บาง	8		41	1												0	82
ฟาง		5	12	33												0	66
มาก			6	8	36											0	72
ปิด	3					39	8									0	78
ผิด						3	45	2								0	90
บิน						8	3	39								0	78
ฝั่ง						5			33		3				9	0	66
มิด						5	8			37						0	74
บัน						6	5	5			34					0	68
ง							7					37		6		0	74
บับ											10	4	36			0	72
ง									5					45		0	90
ม										4		8			38	0	76

เฉลี่ย 75.7

ตารางที่ 5.1 แสดงการกระจายความถูกต้องในการรับรู้ของคำในกลุ่มที่ 1



แบบอ้างอิง แบบทดสอบ	ที่	การ	เป็น	ได้	จะ	ใน	ไม่	ก็	ของ	ให้	Rej.	%
ที่	45		2	1	2						0	90
การ		44	3			3					0	88
เป็น		3	46			1					0	92
ได้	3			44				1		2	0	88
จะ	3			1	45		1				0	96
ใน	2			3		42	3				0	80
ไม่				6		2	42				0	84
ก็		2		5				43			0	86
ของ				1					48	1	0	96
ให้				3			2			45	0	90

เฉลี่ย 89.4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.2 แสดงการกระจายความถูกต้องในการรับรู้ของคำในกลุ่มที่ 2

แบบอ้างอิง แบบทดสอบ	ปาก	พาด	บาง	ฟาง	มาก	ปิด	ผิด	โบน	ฝัน	มิด	ป็น	ฟ	บับ	ฝ	ม	Rej.	%
ปาก	38				6											5	84
พาด	1	42			2	3										2	87
บาง	5		41													4	89
ฟาง		3	6	35												6	79
มาก			3	3	38											6	86
ปิด	1					39	5									5	86
ผิด						1	45	2								2	94
โบน						5	2	41								2	85
ฝัน						3			37		1			5		4	80
มิด						3	5			40						2	83
ป็น						3	2	2			38					5	84
ฟ							3					40		2		5	88
บับ										6	2		38			4	82
ฝ									2					45		3	95
ม										2		5			40	3	85

เฉลี่ย 85.8

ตารางที่ 5.3 แสดงการกระจายความถูกต้องในการรับรู้ของคำกลุ่มที่ 1
หลังจากกำหนดค่าความแตกต่างที่ยอมรับได้

แบบอ้างอิง แบบทดสอบ	ที่	การ	เป็น	ได้	จะ	ใน	ไม่	ก็	ของ	ให้	Rej.	%
ที่	45		1	1	1						2	93
การ		45	1			1					3	95
เป็น		2	46			1					1	93
ได้	1			45				1			2	93
จะ				1	45	1					3	95
ใน	1			1		43	2				3	91
ไม่				1		2	43				4	93
ก็				2				43			5	96
ของ									48	1	1	98
ให้				1			1			45	3	96

เฉลี่ย 94.3

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.4 แสดงการกระจายความถูกต้องในการรับรู้ของคำกลุ่มที่ 2
หลังจากกำหนดค่าความแตกต่างที่ยอมรับได้