

## บทที่ 4

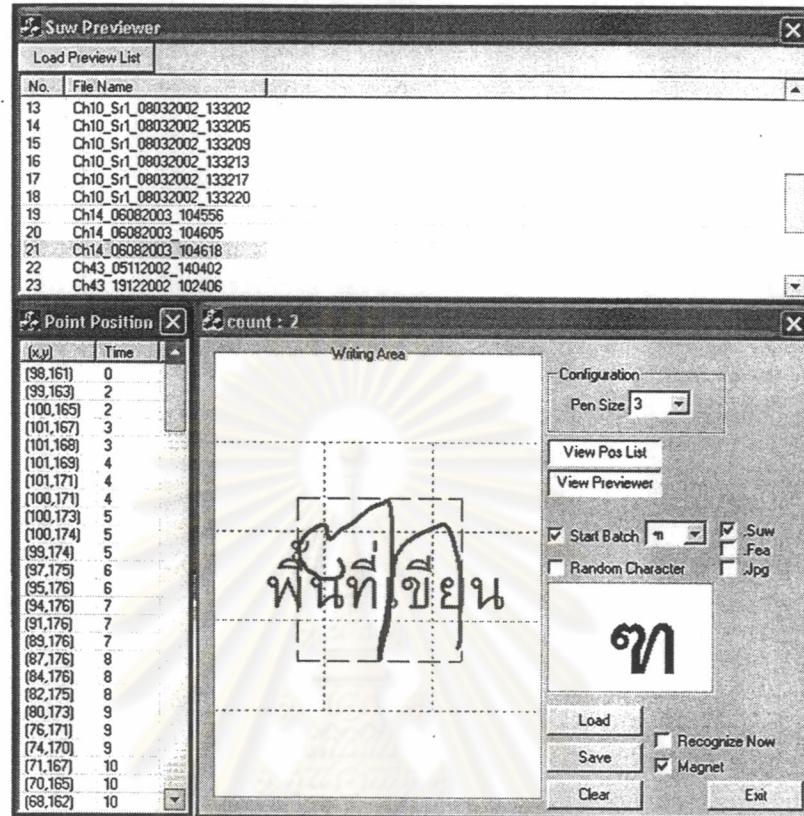
### ขั้นตอนการทดลอง ผลการทดลองและวิเคราะห์ผลการทดลอง

ในบทนี้จะกล่าวถึงการประเมินประสิทธิภาพของระบบการรู้จำตัวอักษรลายมือเขียนภาษาไทยแบบออนไลน์ในงานวิจัยนี้ ทั้งในส่วนผลการทดลองการรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟ และผลการทดลองการรู้จำด้วยวิธีที่เสนอ รวมถึงการวิเคราะห์ผลที่ได้จากการทดลอง

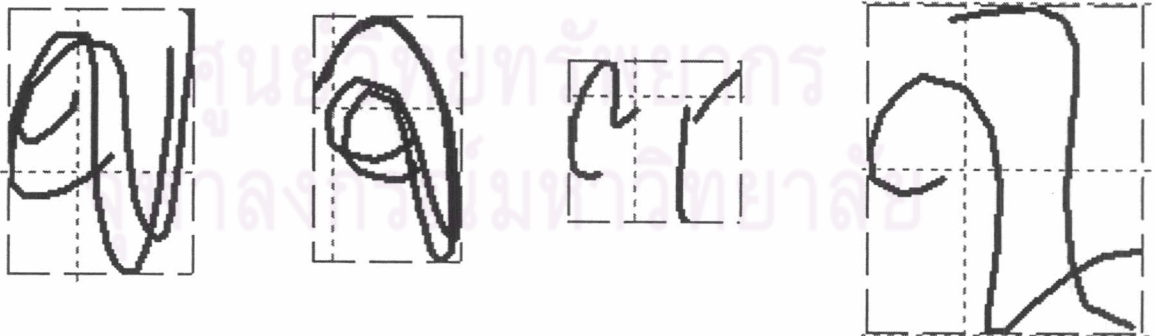
#### 4.1 เครื่องมือที่ใช้ในการทดสอบ

1. เครื่องคอมพิวเตอร์ส่วนบุคคล ใช้หน่วยประมวลผลกลางรุ่น Pentium III ความเร็ว 733 เมกะเฮิร์ตซ์ (MHz) หน่วยความจำ 256 เมกะไบต์ (MBytes)
2. กระดานอิเล็กทรอนิกส์ ซึ่งเป็นอุปกรณ์เชื่อมต่อในการรับข้อมูล
3. โปรแกรมการรู้จำตัวอักษรพัฒนาขึ้นบนโปรแกรม วิซวลซี++ ของบริษัทไมโครซอฟท์ ที่ทำงานภายใต้ระบบปฏิบัติการวินโดวส์ XP

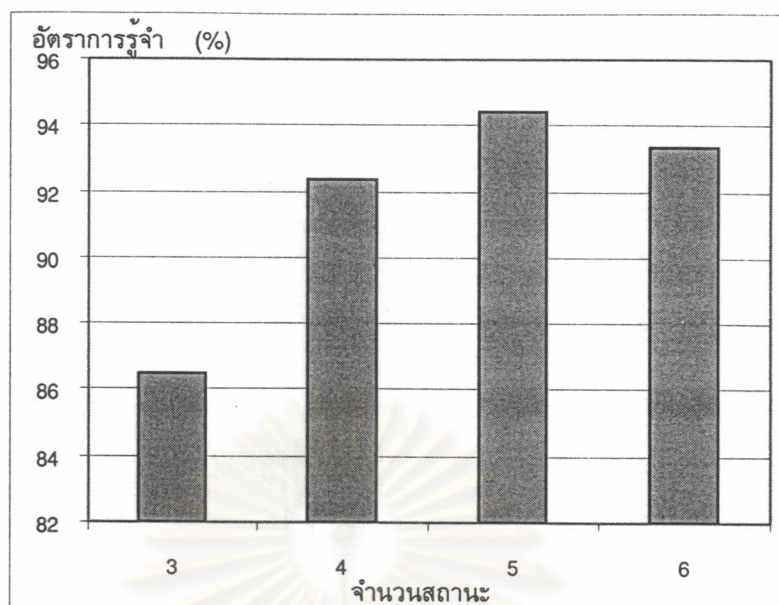
ข้อมูลที่ใช้ฝึกฝน และทดสอบเป็นตัวอักษรภาษาไทยดังแสดงในตารางที่ 1.2 เก็บตัวอักษรผ่านทางกระดานอิเล็กทรอนิกส์ที่มีอัตราการสุ่มข้อมูล 100 จุดต่อวินาที และตั้งกระดานอิเล็กทรอนิกส์ให้ตรงไม่หมุนหรือเอียง ด้วยโปรแกรมเก็บข้อมูลดังแสดงในรูปที่ 4.1 จากผู้เขียนจำนวน 59 คน โดยแบ่งข้อมูลตัวอักษรออกเป็น 2 กลุ่ม คือกลุ่มที่ใช้ฝึกฝน และกลุ่มที่ใช้ทดสอบ การเก็บข้อมูลในกลุ่มแรกมี 40 คน เก็บข้อมูลคนละ 17 ครั้งต่อหนึ่งตัวอักษร และคัดแยกที่เขียนเสียออก ตัวอักษรที่เขียนเสียนั้นคือตัวอักษรที่เขียนผิดเนื่องจากความเผอเรอ หรือใช้กระดานอิเล็กทรอนิกส์ไม่ชำนาญ เช่น การเขียนตัวอักษรทับกัน เนื่องจากยังไม่ได้เก็บข้อมูลตัวเก่าก่อนเขียนตัวใหม่ การกดปากกาในกระดานอิเล็กทรอนิกส์เบาเกินไปทำให้บางเส้นนั้นหายไป เป็นต้น ในรูปที่ 4.2 เป็นรูปตัวอย่างของตัวอักษรที่เขียนผิด นำเอาข้อมูล 10 ตัวต่อหนึ่งตัวอักษรต่อหนึ่งผู้เขียนเพื่อมาฝึกฝนระบบ และนำข้อมูลที่เหลือไปใช้ทดสอบระบบด้วยตัวอักษรของกลุ่มผู้เขียนที่ใช้ในการฝึกฝนระบบ ส่วนอีกกลุ่มหนึ่งเป็นกลุ่มผู้เขียนที่อิสระจากกลุ่มแรกมี 19 คน เก็บข้อมูลคนละ 13 ครั้งต่อหนึ่งตัวอักษรและคัดแยกที่เขียนเสียออก นำเอาข้อมูล 10 ตัวต่อหนึ่งตัวอักษรต่อหนึ่งผู้เขียนมาทดสอบระบบ



รูปที่ 4.1 โปรแกรมในการจัดเก็บข้อมูล



รูปที่ 4.2 ตัวอย่างตัวอักษรที่เขียนเสีย ในการเก็บข้อมูล



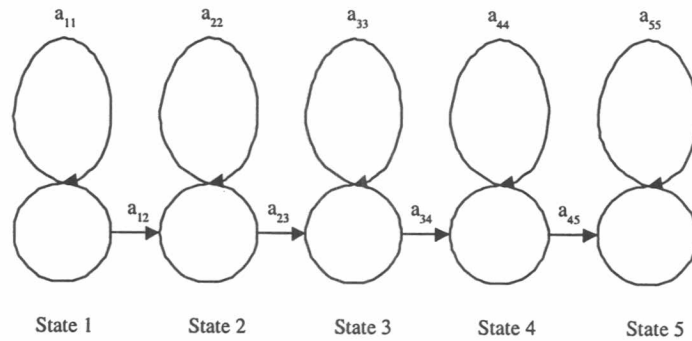
รูปที่ 4.3 อัตราการเรียนรู้จำตัวอักษร 'ก' ถึง 'ฮ' ของกลุ่มผู้เขียนที่ฝึกฝนระบบเมื่อใช้จำนวนสถานะต่างๆ

#### 4.2 แหล่งที่มาของข้อมูลและการเก็บข้อมูล

1. กลุ่มที่ใช้ฝึกฝนระบบประกอบไปด้วย ตัวอักษรและสระระดับกลางจำนวน 19200 ตัว สระและวรรณยุกต์ระดับบน 4800 ตัว และ สระระดับล่างจำนวน 800 ตัว จากผู้เขียน 40 คน
2. กลุ่มที่ใช้ทดสอบแบ่งออกเป็น 2 การทดสอบ คือการทดสอบด้วยตัวอักษรของกลุ่มผู้เขียนที่ใช้ในการฝึกฝนระบบ ประกอบไปด้วย ตัวอักษรและสระระดับกลางจำนวน 12275 ตัว สระและวรรณยุกต์ระดับบน 2504 ตัว และ สระระดับล่างจำนวน 452 ตัว จากผู้เขียน 40 คน การทดสอบด้วยตัวอักษรของกลุ่มผู้เขียนที่ไม่ได้ใช้ในการฝึกฝนระบบ ประกอบไปด้วย ตัวอักษรและสระระดับกลางจำนวน 9120 ตัว สระและวรรณยุกต์ระดับบน 2090 ตัว และ สระระดับล่างจำนวน 380 ตัว จากผู้เขียน 19 คน

#### 4.3 การทดสอบประสิทธิภาพของการรู้จำด้วยแบบจำลองฮิดเดนมาร์คคอฟ

การทดสอบการรู้จำด้วยแบบจำลองฮิดเดนมาร์คคอฟในวิทยานิพนธ์ฉบับนี้ใช้แบบจำลองฮิดเดนมาร์คคอฟแบบต่อเนื่อง (CHMM) มีจำนวนสถานะ 5 สถานะ เนื่องจากเป็นจำนวนสถานะที่มีอัตราการเรียนรู้สูงสุดเมื่อทดสอบด้วยข้อมูลตัวอักษร 'ก' ถึง 'ฮ' ของกลุ่มผู้เขียนที่ฝึกฝนระบบ มีอัตราการเรียนรู้ดังรูป 4.3 และมีโครงสร้างของแบบจำลองดังรูปที่ 4.4 โดยผลการทดสอบประสิทธิภาพสามารถแสดง



รูปที่ 4.4 แบบจำลองฮิตเดนมาร์คอฟที่ใช้ในวิทยานิพนธ์นี้

ตารางที่ 4.1 อัตราการรู้จำเฉลี่ยของระบบรู้จำตัวอักษรภาษาไทยด้วยแบบจำลองฮิตเดนมาร์คอฟ

	อัตราการรู้จำตัวอักษรระดับต่างๆ(%)		
	ระดับบน	ระดับกลาง	ระดับล่าง
ชุดข้อมูลของกลุ่มผู้เขียนที่ใช้ในการฝึกฝนระบบ	93.29	93.53	100.00
ชุดข้อมูลของกลุ่มผู้เขียนอิสระ	84.31	89.08	96.84

ได้ตั้งตารางที่ 4.2 ถึงตารางที่ 4.7 จากผลการทดลองพบว่ามีอัตราการรู้จำเฉลี่ยดังตารางที่ 4.1 พบว่า อัตราการรู้จำเฉลี่ยของกลุ่มผู้ใช้อิสระในตัวอักษรระดับบนร้อยละ 84.3 ในตัวอักษรระดับบนล่างร้อยละ 89.1 แต่เมื่อพิจารณาการอัตราจำของแต่ละตัวอักษรในระดับบนและกลางจากตารางที่ 4.2 ถึงตารางที่ 4.5 พบว่ามีตัวอักษรบางตัวนั้นมีอัตราการรู้จำที่ต่ำ และมักตัดสั่นใจผิดไปเป็นตัวอักษรที่มีลักษณะใกล้เคียงกัน เช่น ตัวอักษร ‘ฎ’ มีอัตราการรู้จำร้อยละ 76.84 ตามตารางที่ 4.3 ตัวที่ระบบตัดสั่นใจผิดมักตัดสั่นใจเป็น ‘ฏ’ ตัวอักษร ‘ท’ มีอัตราการรู้จำร้อยละ 72.11 ตามตารางที่ 4.3 ตัวที่ระบบตัดสั่นใจผิดมักตัดสั่นใจเป็น ‘ฑ’ ตัวอักษร ‘บ’ มีอัตราการรู้จำร้อยละ 66.87 ตามตารางที่ 4.3 ตัวที่ระบบตัดสั่นใจผิดมักตัดสั่นใจเป็น ‘ข’ สระ ‘๑’ มีอัตราการรู้จำร้อยละ 40.00 ตามตารางที่ 4.5 ตัวที่ระบบตัดสั่นใจผิดมักตัดสั่นใจเป็น ‘๒’ เป็นต้น อีกทั้งในกรณีของวรรณยุกต์ ‘โ’ นั้นไม่สามารถฝึกฝนแบบจำลองได้เนื่องจากข้อมูลของลายมือเขียนส่วนใหญ่มีเพียง 1 เส้นย่อย สำหรับตัวอักษรในระดับล่างมี สระ ‘ุ’ และ ‘ู’ เพียง 2 ตัวนั้น มีอัตราการรู้จำเฉลี่ยสูงถึงร้อยละ 96.84 ในกรณีลายมือเขียนของผู้เขียนอิสระ ในหัวข้อต่อไปจะวิเคราะห์ถึงลักษณะความผิดพลาดที่พบในระบบรู้จำตัวอักษรลายมือเขียนภาษาไทยด้วยแบบจำลองฮิตเดนมาร์คอฟ

#### 4.4 วิเคราะห์การทดสอบประสิทธิภาพของการรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟ

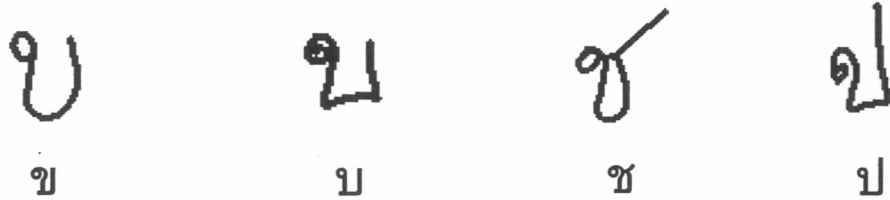
ความผิดพลาดที่เกิดขึ้นในการทดสอบประสิทธิภาพของการรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟส่วนใหญ่จะตัดสินใจผิดพลาดไปเป็นตัวอักษรที่มีความคล้ายคลึงกันดังที่กล่าวไปแล้ว ถึงแม้้อตราการรู้จำเฉลี่ยจะค่อนข้างสูงแต่จะมีตัวอักษรบางตัวที่มี้อตราการรู้จำต่ำ เนื่องจากกระบวนการการฝึกฝนแบบจำลองฮิดเดนมาร์คอฟพยายามสร้างความเป็นภายหลัง  $P(O|\lambda)$  ของแต่ละแบบจำลองให้สูงสุด แต่ไม่ได้มีการเปรียบเทียบกันในแต่ละตัวอักษร ดังนั้นคำตอบของระบบจึงมีโอกาสที่จะเป็นตัวที่มีลักษณะคล้ายกับตัวอักษรที่ถูกต้องค่อนข้างมาก ความผิดพลาดที่เกิดขึ้นในการทดลองที่ผ่านมาสามารถแบ่งกลุ่มของตัวอักษรที่คล้ายกันได้ 6 กลุ่มตามประเภทของลักษณะบ่งความต่างที่แบ่งแยกตัวอักษรที่คล้ายกันนั้นได้ดังนี้

1. กลุ่มตัวอักษรที่มีลักษณะคล้ายกัน ต่างกันที่ทิศทางการเขียนของหัว กลุ่มนี้ประกอบไปด้วย 2 กลุ่มย่อยคือ ('ก' 'ถ' และ 'ภ') และ ('ร' 'โ' และ 'ธ') ลักษณะโดยรวมของกลุ่มนี้เมื่อเป็นลายมือเขียนจะคล้ายคลึงกันมากดังแสดงในรูปที่ 4.5 ซึ่งเป็นตัวอักษรจากฐานข้อมูลลายมือเขียนที่ใช้ทดสอบโดยจะมีลักษณะการเขียนของช่วงต้นตัวอักษรต่างกันเท่านั้น



รูปที่ 4.5 ตัวอย่างลายมือเขียนตัวอักษรที่แตกต่างกันที่ทิศทางการเขียนหัว

2. กลุ่มตัวอักษรที่มีลักษณะคล้ายกัน ต่างกันที่อัตราส่วนความกว้างต่อความสูง กลุ่มนี้ประกอบไปด้วย 2 คู่ตัวอักษรที่มี้อตราการรู้จำต่ำคือ ('บ' กับ 'ข') และ ('ป' กับ 'ช') ลักษณะลายมือเขียนของกลุ่มนี้ยกตัวอย่างไว้ในรูปที่ 4.6 ลักษณะที่บอกความแตกต่างของตัวอักษรแต่ละตัวคือความกว้างของตัวอักษร
3. กลุ่มตัวอักษรที่มีลักษณะคล้ายกัน ต่างกันที่ตำแหน่ง จุดต้นและ จุดปลายของตัวอักษร คู่ของตัวอักษรที่มีตำแหน่งจุดเริ่มต้น หรือตำแหน่งจุดปลายต่างกัน และแบบจำลองฮิดเดนมาร์คอฟ



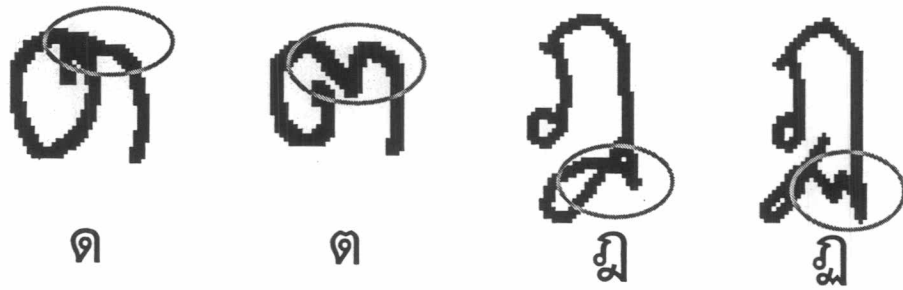
รูปที่ 4.6 ตัวอย่างลายมือเขียนตัวอักษรที่แตกต่างกันที่อัตราส่วนความกว้างต่อความสูง

มักตัดสินใจผิดประกอบไปด้วย ('ม' กับ 'ฒ' และ 'ท') ('ร' กับ 'ภ') และ ('ถ' กับ 'ธ' และ 'ฒ')  
 ดังแสดงตัวอย่างลายมือเขียนของตัวอักษรในกลุ่มนี้ในรูปที่ 4.7

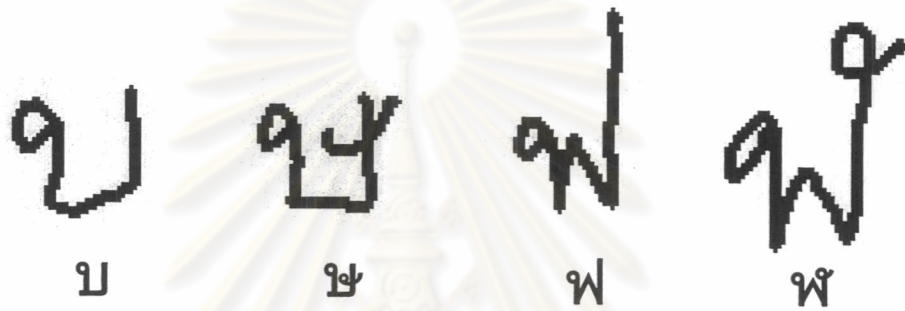


รูปที่ 4.7 ตัวอย่างลายมือเขียนตัวอักษรที่แตกต่างกันที่ตำแหน่งจุดต้น และจุดปลายของตัวอักษร

4. กลุ่มตัวอักษรที่มีลักษณะคล้ายกัน แตกกันที่รอยหยัก ตัวอักษรที่มีรอยหยักในภาษาไทยมักพบคู่กับตัวอักษรที่มีรูปร่างเหมือนกันแต่ไม่มีรอยหยัก โดยกลุ่มที่มีอัตราการเรียนรู้จำต่ามีดังนี้ ('ฎ' กับ 'ฏ') ('ม' กับ 'ฒ') ('ท' กับ 'ฑ') ('ด' กับ 'ต') ('ช' กับ 'ฌ') และ ('ย' กับ 'พ') ดังแสดงตัวอย่างลายมือเขียนของตัวอักษรในกลุ่มนี้ในรูปที่ 4.8
5. กลุ่มตัวอักษรที่มีลักษณะคล้ายกัน แตกกันที่ตำแหน่งจุดตัดบนตัวอักษร คู่ของตัวอักษรที่มีตำแหน่งจุดตัดกันของเส้นต่างกัน และแบบจำลองฮิตเดนมาร์คอฟมักตัดสินใจผิดประกอบไปด้วย ('พ' กับ 'ฟ' และ 'ภ') ('บ' กับ 'ม' กับ 'ษ') ดังแสดงตัวอย่างลายมือเขียนของตัวอักษรในกลุ่มนี้ในรูปที่ 4.9
6. กลุ่มตัวอักษรที่มีลักษณะคล้ายกัน แตกกันที่ลักษณะและจำนวนขีดด้านท้ายตัวอักษร สระ '๕' '๖' และ '๗' มีลักษณะการเขียนเริ่มต้นเหมือนกับสระ '๑' เพียงแต่มีขีด หรือวงกลมต่อท้ายเข้าไป จึงมีโอกาสสับสนจำผิดพลาดได้มาก สำหรับตัวอย่างของลายมือเขียนสระกลุ่มนี้แสดงได้ในรูปที่ 4.10



รูปที่ 4.8 ตัวอย่างลายมือตัวอักษรที่แตกต่างกันที่รอยหยัก



รูปที่ 4.9 ตัวอย่างลายมือตัวอักษรที่แตกต่างกันที่ตำแหน่งจุดตัดกันของเส้น

ศูนย์วิทยพัทยาการ  
จุฬาลงกรณ์มหาวิทยาลัย





ตารางที่ 4.3 ผลการทดสอบระบบรับรู้จำด้วยแบบจำลองอิตเดนมาร์คอฟที่ใช้ข้อมูลจากตัวอักษรและสระระดับกลางของกลุ่มผู้เขียนอิสระ

ก	ข																										ค	จ	ช	ฉ	ส	ห	ฬ	อ	ะ	า	า	อ	ะ	า	อ	ะ	า	อ	ะ	า	อ	ะ	า	อ	ะ	%	Tot	Correct
	ก	ข	ค	จ	ช	ฉ	ส	ห	ฬ	อ	ะ	า	า	อ	ะ	า	อ	ะ	า	อ	ะ	า	อ	ะ	า	อ																												
ก	158	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	83.16	158																								
ข	0	138	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	72.63	138																								
ค	1	0	183	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	96.32	183																								
จ	0	0	0	164	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	86.32	164																								
ช	0	0	0	2	182	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	95.79	182																								
ฉ	0	0	0	0	1	171	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	90.00	171																								
ส	0	0	0	0	0	4	184	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	96.84	184																								
ห	0	0	0	0	0	0	0	136	42	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	71.58	136																								
ฬ	0	0	0	3	0	0	5	166	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	87.37	166																								
อ	0	0	0	0	0	0	0	0	0	184	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	96.84	184																								
ะ	0	0	0	1	0	0	0	0	176	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	92.63	176																								
า	0	0	0	0	0	0	0	1	146	43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	76.84	146																								
า	0	0	0	0	0	0	0	0	0	1	13	171	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	90.00	171																								
อ	0	0	0	0	0	0	0	0	0	0	188	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	96.95	188																								
ะ	0	0	0	12	0	0	0	0	0	0	0	0	171	3	0	0	0	0	0	0	0	0	0	0	0	0	0	190	90.00	171																								
า	0	0	0	1	0	0	0	0	0	0	184	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	96.84	184																								
อ	0	0	0	0	0	0	0	0	0	0	181	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	95.26	181																								
อ	0	0	2	0	0	0	0	0	0	0	0	162	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	85.26	162																								
อ	0	0	0	0	0	0	0	0	3	0	0	0	1	180	0	0	0	0	0	0	0	0	0	0	0	0	0	190	94.74	180																								
อ	0	0	1	0	0	0	0	0	0	0	0	2	1	172	0	0	0	0	0	0	0	0	0	0	0	0	0	190	90.53	172																								
อ	0	0	0	0	0	0	0	0	0	0	0	1	0	137	0	0	0	0	0	0	0	0	0	0	0	0	0	190	72.11	137																								
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	179	0	0	0	0	0	0	0	0	0	0	0	0	190	94.21	179																								
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	163	0	0	0	0	0	0	0	0	0	0	0	0	190	85.79	163																								
อ	13	10	0	0	0	0	0	0	0	0	0	127	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	66.84	127																								
อ	0	0	0	0	0	0	0	0	0	0	0	0	1	140	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	73.68	140																							
อ	0	0	1	0	0	0	0	0	0	0	0	0	0	0	186	2	0	0	0	0	0	0	0	0	0	0	0	0	190	97.90	186																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	177	0	0	0	0	0	0	0	0	0	0	0	190	96.42	177																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	4	127	0	0	0	0	0	0	0	0	0	190	66.84	127																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	181	0	0	0	0	0	0	0	0	190	95.26	181																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	167	0	0	0	0	0	0	0	190	87.90	167																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	176	0	0	0	0	0	0	190	92.63	176																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	190	88.42	168																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	94.74	168																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	82.11	156																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	97.37	185																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	88.95	169																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	94.21	179																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	93.68	178																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	90.53	172																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	93.16	177																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	83.66	178																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	94.21	179																							
อ	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	96.32	183																								
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	76.32	145																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	88.95	169																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	95.47	189																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	91.58	174																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	89.09	182																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	91.20	180																							
อ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	89.09	182																							

ตารางที่ 4.4 ผลการทดสอบระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟโดยใช้ข้อมูลจากสระและวรรณยุกต์ระดับบนสำหรับทดสอบของกลุ่มผู้เขียนที่ใช้น

การฝึกฝนระบบ

	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	Tot	%	Correct	
อิ	207	5	7	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	223	92.83	207	อิ
อี	1	183	14	29	1	0	1	0	0	0	1	0	0	0	0	0	0	0	230	79.57	183	อี
อ	1	1	199	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	226	88.05	199	อ
อิอ	1	18	9	205	0	0	0	0	0	0	0	0	0	0	0	0	0	0	233	87.98	205	อิอ
อ่า	5	0	0	0	196	8	5	0	0	0	0	0	0	0	0	0	4	218	89.90	196	อ่า	
อัน	0	0	0	0	1	225	0	0	0	0	3	1	0	0	0	0	0	230	97.83	225	อัน	
อิ	0	1	0	0	0	0	221	5	0	0	0	0	0	0	0	0	0	227	97.36	221	อิ	
อิ	0	0	0	0	0	0	10	212	1	2	0	0	0	0	0	0	0	225	94.22	212	อิ	
อิอ	0	0	0	0	0	0	0	0	0	221	1	0	0	0	0	0	0	222	99.55	221	อิอ	
อิอ	0	0	0	0	0	0	0	0	0	1	0	230	0	0	0	0	0	231	99.57	230	อิอ	
อ	0	0	0	1	0	0	0	0	0	0	0	1	237	0	0	0	0	239	99.16	237	อ	
																		รวม	2504	93.29	2336	
	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ				

ตารางที่ 4.5 ผลการทดสอบบรรณารักษ์จำด้วยแบบจำลองฮิตเดนมาร์คอฟที่ใช้ข้อมูลจากระยะเวลาการยกฐานะและวารรณยุกต์ระดับบนของกลุ่มผู้เขียนอิสระ

	อิ	อี	อ	อึ	อื	อึ	อื	อึ	อื	อึ	อื	อึ	อื	อึ	อื	อึ	อื	อึ	อื	Tot	%	Correct	
อิ	147	14	16	6	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	190	77.37	147	อิ
อี	7	76	13	85	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	40.00	76	อี
อ	1	0	176	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	92.63	176	อ
อึ	4	28	27	122	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	64.21	122	อึ
อื	11	1	0	0	173	4	0	0	0	0	0	0	0	0	0	0	0	0	0	190	91.05	173	อื
อึ	0	0	0	0	2	186	0	0	2	0	0	0	0	0	0	0	0	0	0	190	97.90	186	อึ
อื	0	4	0	1	1	0	182	2	0	0	0	0	0	0	0	0	0	0	0	190	95.79	182	อื
อึ	0	0	0	0	0	0	19	166	0	2	3	0	0	0	0	0	0	0	0	190	87.37	166	อึ
อื	2	0	0	0	1	1	0	0	179	1	6	0	0	0	0	0	0	0	0	190	94.21	179	อื
อึ	0	3	10	0	2	0	0	0	0	0	175	0	0	0	0	0	0	0	0	190	92.11	175	อึ
อื	0	0	0	0	8	0	0	0	2	0	0	180	0	0	0	0	0	0	0	190	94.74	180	อื
																				2090	84.31	1762	
	อิ	อี	อ	อึ	อื	อึ	อื	อึ	อื	อึ	อื	อึ	อื	อึ	อื	อึ	อื	อึ	อื				



รูปที่ 4.10 ตัวอย่างลายมือตัวอักษรที่แตกต่างกันที่จำนวนขีดด้านท้ายตัวอักษร

ตารางที่ 4.6 ผลการทดสอบระบบรู้จำที่นำเสนอโดยใช้ข้อมูลจากระดับล่างสำหรับทดสอบของกลุ่มผู้เขียนที่ใช้ในการฝึกฝนระบบ

	อ	อุ	Tot	%	Correct
อ	223	0	223	100.00	223
อุ	0	229	229	100.00	229
	รวม		452	100.00	452

#### 4.5 การทดสอบประสิทธิภาพของการรู้จำด้วยระบบรู้จำที่นำเสนอ

เมื่อทดสอบประสิทธิภาพของระบบรู้จำลายมือเขียนภาษาไทยแบบออนไลน์ด้วยแบบจำลองฮิดเดน มาร์คอฟและพีชชีโลจิกซึ่งมีขั้นตอนการทำงานดังรูปที่ 3.1 พบว่ามีอัตราการรู้จำเฉลี่ยในชุดข้อมูลทดสอบทั้งของกลุ่มผู้เขียนที่ใช้ในการฝึกฝนระบบ และกลุ่มผู้เขียนอิสระดังตารางที่ 4.8 โดยมีรายละเอียดในการรู้จำแต่ละกลุ่มข้อมูล และระดับของตัวอักษรตามตารางที่ 4.15 ถึงตารางที่ 4.18 การรู้จำด้วยวิธีที่นำเสนอนี้ทำให้อัตราการรู้จำส่วนใหญ่ของแต่ละตัวอักษรดีขึ้น โดยเฉพาะกลุ่มที่เคยมีอัตราการรู้จำต่ำดังที่เคยกล่าวมาในหัวข้อที่แล้ว สำหรับอัตราการรู้จำเฉลี่ยของตัวอักษรระดับกลางและระดับบน

ตารางที่ 4.7 ผลการทดสอบระบบรู้จำที่นำเสนอโดยใช้ข้อมูลจากระดับล่างของกลุ่มผู้เขียนอิสระ

	อ	อุ	Tot	%	Correct
อ	187	3	190	98.42	187
อุ	9	181	190	95.26	181
	รวม		380	96.84	368

ตารางที่ 4.8 อัตราการรู้จำเฉลี่ยของระบบรู้จำตัวอักษรภาษาไทยด้วยแบบจำลองฮิดเดนมาร์คอฟและฟัซซีโลจิก

	อัตราการรู้จำตัวอักษรระดับต่าง ๆ (%)		
	ระดับบน	ระดับกลาง	ระดับล่าง
ชุดข้อมูล ของ กลุ่ม ผู้เขียนที่ใช้ในการฝึกฝนระบบ	94.32	95.66	100.00
ชุดข้อมูลของกลุ่มผู้เขียนอิสระ	90.96	93.27	96.84

ตารางที่ 4.9 อัตราการรู้จำผิดพลาดภายในตัวอักษรภายในกลุ่มที่มีความแตกต่างที่ทิศทางการเขียนหัว และอัตราการรู้จำเฉลี่ยของตัวอักษรเหล่านั้น

ชนิดของระบบรู้จำ		กลุ่มที่ 1			กลุ่มที่ 2		
		ก	ถ	ฏ	ฐ	ร	โร
อัตราการรู้จำผิดเป็นตัวอักษรภายในกลุ่ม (%)	HMM	14.21	1.58	2.63	0.53	5.26	4.74
	HMM+FUZZY	8.42	2.11	2.11	0.53	3.16	3.16
อัตราการรู้จำเฉลี่ย (%)	HMM	83.16	90.53	95.26	94.21	88.42	88.95
	HMM+FUZZY	89.47	94.74	95.79	96.32	92.11	94.21

ของข้อมูลจากผู้เขียนที่ใช้ฝึกฝนระบบร้อยละ 95.66 และ 94.32 ตามลำดับ เพิ่มขึ้นจากการรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟร้อยละ 2.13 และ 1.03 อัตราการรู้จำเฉลี่ยของตัวอักษรระดับกลางและระดับบนของข้อมูลจากผู้เขียนอิสระร้อยละ 93.27 และ 90.96 ตามลำดับ เพิ่มขึ้นจากการรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟร้อยละ 4.15 และ 5.75 ในขณะที่เวลาในการทำงานของระบบรู้จำที่นำเสนอเท่ากับ 768 วินาทีเมื่อทดสอบกับตัวอักษร 8804 ตัว หรือใช้เวลาทำงาน 0.09 วินาทีต่อตัวอักษรโดยเฉลี่ย ซึ่งมากกว่าการรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟที่มีเวลาทำงาน 0.08 วินาทีต่อตัวอักษรโดยเฉลี่ยเพียงเล็กน้อยเมื่อทดสอบด้วยชุดข้อมูลเดียวกัน

ส่วนผลการทดสอบเปรียบเทียบอัตราการรู้จำเฉลี่ยระหว่างการรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟกับการรู้จำที่นำเสนอในกลุ่มตัวอักษรที่คล้ายกันสามารถแสดงได้ดังต่อไปนี้

1. กลุ่มตัวอักษรที่มีลักษณะคล้ายกันต่างกันที่ทิศทางการเขียนของหัว ผลการทดสอบด้วยชุดข้อมูลตัวอักษรจากผู้เขียนอิสระในตัวอักษรกลุ่มนี้แสดงรายละเอียดได้ดังตารางที่ 4.9 พบว่าระบบรู้จำ

ตารางที่ 4.10 อัตราการรู้จำผิดพลาดภายในตัวอักษรภายในกลุ่มที่มีความแตกต่างที่อัตราส่วนความกว้างต่อความสูง และอัตราการรู้จำเฉลี่ยของตัวอักษรเหล่านั้น

	ชนิดของระบบรู้จำ	กลุ่มที่ 1		กลุ่มที่ 2	
		ข	บ	ช	ป
อัตราการรู้จำผิดเป็น	HMM	8.95	6.84	0.00	7.89
ตัวอักษรภายในกลุ่ม(%)	HMM+FUZZY	8.95	10.00	0.00	13.16
อัตราการรู้จำเฉลี่ย(%)	HMM	72.63	66.84	71.58	73.68
	HMM+FUZZY	78.95	76.84	90.00	78.42

ที่นำเสนอแก้ไขการตัดสินใจผิดพลาดของแบบจำลองฮิดเดนมาร์คอฟได้เล็กน้อย ในกลุ่มตัวอักษร ('ก' 'ถ' และ 'ภ') มีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 4.2 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 6.1 โดยตัวอักษร 'ก' อัตราการรู้จำผิดพลาดในกลุ่มดีขึ้นพอสมควร ขณะที่ ตัวอักษร 'ถ' และตัวอักษร 'ภ' มีอัตราการรู้จำผิดพลาดมากขึ้นเพียงเล็กน้อย

ในกลุ่มตัวอักษร ('ธ' 'ร' และ 'ไ') มีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 2.3 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 3.5 โดยตัวอักษร 'ร' และ 'ไ' มีอัตราการรู้จำผิดพลาดในกลุ่มดีขึ้นเล็กน้อย

2. กลุ่มตัวอักษรที่มีลักษณะคล้ายกัน ต่างกันที่อัตราส่วนความกว้างต่อความสูง ผลการทดสอบด้วยชุดข้อมูลตัวอักษรจากผู้เขียนอิสระในตัวอักษรกลุ่มนี้แสดงรายละเอียดได้ดังตารางที่ 4.10 พบว่าระบบรู้จำที่นำเสนอไม่สามารถแก้ไขการตัดสินใจผิดพลาดของแบบจำลองฮิดเดนมาร์คอฟ ในกลุ่มตัวอักษรที่มีความแตกต่างกันที่ความกว้างของตัวอักษร ในคู่ตัวอักษร ('บ' กับ 'ช') มีอัตราการรู้จำผิดพลาดภายในคู่เฉลี่ยร้อยละ 9.15 เพิ่มขึ้นจากการใช้ระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีอัตราการรู้จำผิดพลาดภายในคู่เฉลี่ยร้อยละ 7.89

ในคู่ตัวอักษร ('ป' กับ 'ช') มีอัตราการรู้จำผิดพลาดภายในคู่เฉลี่ยร้อยละ 6.58 เพิ่มขึ้นจากการใช้ระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีอัตราการรู้จำผิดพลาดภายในคู่เฉลี่ยร้อยละ 3.95

3. กลุ่มตัวอักษรที่มีลักษณะคล้ายกันต่างกันที่ตำแหน่ง จุดต้นและ จุดปลายของตัวอักษร ผลการทดสอบด้วยชุดข้อมูลตัวอักษรจากผู้เขียนอิสระในตัวอักษรกลุ่มนี้แสดงรายละเอียดได้ดังตารางที่ 4.11

ตารางที่ 4.11 อัตราการรู้จำผิดพลาดภายในตัวอักษรภายในกลุ่มที่มีความแตกต่างที่ตำแหน่งจุดต้น และจุดปลายของตัวอักษร และอัตราการรู้จำเฉลี่ยของตัวอักษรเหล่านั้น

	ชนิดของระบบรู้จำ	กลุ่มที่ 1			กลุ่มที่ 2		กลุ่มที่ 3		
		ตม	ชม	จท	ร	ภ	ถ	ณ	ณ
อัตราการรู้จำผิดเป็น	HMM	0.53	7.89	6.32	2.66	0.56	5.79	0.00	0.00
ตัวอักษรภายในกลุ่ม(%)	HMM+FUZZY	1.58	1.58	0.00	1.58	0.00	1.05	0.00	0.00
อัตราการรู้จำเฉลี่ย(%)	HMM	96.84	86.32	90.00	88.42	95.26	90.53	95.26	96.84
	HMM+FUZZY	95.79	93.16	92.11	92.11	95.79	94.74	97.89	99.47

ตารางที่ 4.12 อัตราการรู้จำผิดพลาดภายในตัวอักษรภายในกลุ่มที่มีความแตกต่างที่รอยหยัก และอัตราการรู้จำเฉลี่ยของตัวอักษรเหล่านั้น

	ชนิดของระบบรู้จำ	กลุ่มที่ 1		กลุ่มที่ 2		กลุ่มที่ 3		กลุ่มที่ 4		กลุ่มที่ 5		กลุ่มที่ 6	
		ฎ	ฏ	ม	ฆ	ท	ฑ	ด	ต	ช	ช	ย	ผ
อัตราการรู้จำผิดเป็น	HMM	22.57	6.84	11.05	3.68	18.42	1.05	11.58	0.53	22.11	2.63	6.30	0.00
ตัวอักษรภายในกลุ่ม(%)	HMM+FUZZY	3.68	9.47	2.63	3.68	0.53	2.63	0.00	0.00	3.68	2.63	4.74	0.53
อัตราการรู้จำเฉลี่ย(%)	HMM	76.84	90.00	87.90	86.32	72.11	90.00	85.26	94.73	71.57	87.37	92.63	97.89
	HMM+FUZZY	94.73	87.36	94.73	93.16	90.53	92.11	96.31	95.79	90.00	88.47	93.68	97.36

พบว่าระบบรู้จำที่นำเสนอแก้ไขการตัดสินใจผิดพลาดของแบบจำลองฮิดเดนมาร์คอฟได้ดีมาก ในกลุ่มตัวอักษร ('ช' กับ 'ตม' และ 'จท') มีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยเพียงร้อยละ 1.05 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 4.91 โดยตัวอักษร 'ช' และ 'จท' มีอัตราการรู้จำผิดพลาดในกลุ่มดีขึ้นมาก

ในคู่ตัวอักษร ('ร' กับ 'ภ') มีอัตราการรู้จำผิดพลาดภายในคู่เฉลี่ยร้อยละ 0.69 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีอัตราการรู้จำผิดพลาดภายในคู่เฉลี่ยร้อยละ 1.58 โดยตัวอักษร 'ร' และ 'ภ' มีอัตราการรู้จำผิดพลาดในกลุ่มดีขึ้นเล็กน้อย

ในกลุ่มตัวอักษร ('ถ' กับ 'ณ' และ 'ณ') มีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยเพียงร้อยละ 0.35 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 1.93 โดยตัวอักษร 'ถ' มีอัตราการรู้จำผิดพลาดในกลุ่มดีขึ้นมาก

4. กลุ่มตัวอักษรที่มีลักษณะคล้ายกันต่างกันอย่างที่รอยหยัก ผลการทดสอบด้วยชุดข้อมูลตัวอักษรจากผู้เขียนอิสระในตัวอักษรกลุ่มนี้แสดงรายละเอียดได้ดังตารางที่ 4.12 พบว่าระบบรู้จำที่นำเสนอ

แก้ไขการตัดสินใจผิดพลาดของแบบจำลองฮิตเดนมาร์คอฟได้ดีมาก ในคู่ตัวอักษร ('ฎ' กับ 'ฏ') มีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 6.55 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิตเดนมาร์คอฟซึ่งมีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 14.74 โดยตัวอักษร 'ฎ' มีอัตราการรู้จำผิดพลาดในคู่ดีขึ้นมาก ส่วนตัวอักษร 'ฏ' มีอัตราการเรียนรู้ผิดพลาดในคู่มากขึ้นเล็กน้อย

ในคู่ตัวอักษร ('ม' กับ 'ฌ') มีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 3.16 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิตเดนมาร์คอฟซึ่งมีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 7.37 โดยตัวอักษร 'ม' มีอัตราการเรียนรู้ผิดพลาดในคู่ดีขึ้นมาก

ในคู่ตัวอักษร ('ท' กับ 'ฑ') มีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 1.58 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิตเดนมาร์คอฟซึ่งมีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 9.75 โดยตัวอักษร 'ท' มีอัตราการเรียนรู้ผิดพลาดในคู่ดีขึ้นมาก

ในคู่ตัวอักษร ('ด' กับ 'ต') ไม่มีการรู้จำผิดพลาดภายในคู่ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิตเดนมาร์คอฟซึ่งมีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 6.05 โดยตัวอักษร 'ด' มีอัตราการรู้จำผิดพลาดในคู่ดีขึ้นมาก

ในคู่ตัวอักษร ('ช' กับ 'ฌ') มีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 3.15 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิตเดนมาร์คอฟซึ่งมีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 12.35 โดยตัวอักษร 'ช' มีอัตราการเรียนรู้ผิดพลาดในคู่ดีขึ้นมาก

ในคู่ตัวอักษร ('ย' กับ 'ฬ') มีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 2.11 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิตเดนมาร์คอฟซึ่งมีอัตราการเรียนรู้ผิดพลาดภายในคู่เฉลี่ยร้อยละ 3.15 โดยตัวอักษร 'ย' มีอัตราการเรียนรู้ผิดพลาดในคู่ดีขึ้นเล็กน้อย

5. กลุ่มตัวอักษรที่มีลักษณะคล้ายกันต่างกันที่ตำแหน่งจุดตัดบนตัวอักษร ผลการทดสอบด้วยชุดข้อมูลตัวอักษรจากผู้เขียนอิสระในตัวอักษรกลุ่มนี้แสดงรายละเอียดได้ดังตารางที่ 4.13 พบว่าระบบรู้จำที่นำเสนอแก้ไขการตัดสินใจผิดพลาดของแบบจำลองฮิตเดนมาร์คอฟได้ดี ในกลุ่มตัวอักษร ('ฟ' กับ 'ฟ' และ 'ฟ') มีอัตราการเรียนรู้ผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 2.11 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิตเดนมาร์คอฟซึ่งมีอัตราการเรียนรู้ผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 11.91 โดยตัวอักษร 'ฟ' อัตราการเรียนรู้ผิดพลาดในกลุ่มดีขึ้นมาก

ในกลุ่มตัวอักษร ('บ' กับ 'ม' และ 'ษ') มีอัตราการเรียนรู้ผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 3.86 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิตเดนมาร์คอฟซึ่งมีอัตราการเรียนรู้ผิดพลาดภายในกลุ่ม



ตารางที่ 4.13 อัตราการรู้จำผิดพลาดภายในตัวอักษรภายในกลุ่มที่มีความแตกต่างที่ตำแหน่ง จุดตัด และอัตราการรู้จำเฉลี่ยของตัวอักษรเหล่านั้น

	ชนิดของระบบรู้จำ	กลุ่มที่ 1			กลุ่มที่ 2		
		พ	ฟ	ฟ	บ	ม	ช
อัตราการรู้จำผิดพลาดเป็น	HMM	0.53	28.42	6.78	11.58	0.00	3.68
ตัวอักษรภายในกลุ่ม(%)	HMM+FUZZY	0.00	3.16	3.16	4.21	0.00	7.37
อัตราการรู้จำเฉลี่ย(%)	HMM	93.16	66.84	90.53	66.84	87.89	88.95
	HMM+FUZZY	95.26	92.68	91.58	76.84	94.74	85.79

ตารางที่ 4.14 อัตราการรู้จำผิดพลาดภายในตัวอักษรภายในกลุ่มที่มีความแตกต่างที่ลักษณะ และ จำนวนขีดด้านท้ายตัวอักษร และอัตราการรู้จำเฉลี่ยของตัวอักษรเหล่านั้น

	ชนิดของระบบรู้จำ	อิ	อี	อี	อีอ
อัตราการรู้จำผิดพลาดเป็น	HMM	18.95	55.26	7.37	33.68
ตัวอักษรภายในกลุ่ม(%)	HMM+FUZZY	11.05	12.63	12.63	14.21
อัตราการรู้จำเฉลี่ย(%)	HMM	77.37	40.00	92.63	64.21
	HMM+FUZZY	86.32	84.74	87.37	82.63

เฉลี่ยร้อยละ 5.10 โดยตัวอักษร 'บ' มีอัตราการรู้จำผิดพลาดในกลุ่มดีขึ้น แต่ตัวอักษร 'ช' มีอัตราการรู้จำผิดพลาดในกลุ่มมากขึ้นเล็กน้อย

6. กลุ่มตัวอักษรที่มีลักษณะคล้ายกันต่างกันที่ลักษณะและจำนวนขีดด้านท้ายตัวอักษร ผลการทดสอบ ด้วยชุดข้อมูลตัวอักษรจากผู้เขียนอิสระในตัวอักษรกลุ่มนี้แสดงรายละเอียดได้ดังตารางที่ 4.14 พบว่าระบบรู้จำที่นำเสนอแก้ไขการตัดสินใจผิดพลาดของแบบจำลองฮิดเดนมาร์คอฟได้ดี ในกลุ่มตัวอักษร ('๕' '๖' '๗' และ '๘') มีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 12.63 ลดลงจากการใช้ระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คอฟซึ่งมีอัตราการรู้จำผิดพลาดภายในกลุ่มเฉลี่ยร้อยละ 28.81 โดยตัวอักษร '๖' '๗' และ '๘' อัตราการรู้จำผิดพลาดในกลุ่มดีขึ้นมาก ขณะที่ตัวอักษร '๕' มีอัตราการรู้จำผิดพลาดมากขึ้นเล็กน้อย

ตารางที่ 4.15 ผลการทดสอบระบบรู้จักที่นำเสนอด้วยข้อมูลจากตัวอักษรและสระระดับกลางสำหรับทดสอบของกลุ่มผู้เขียนที่ใช้ในการฝึกหัดระบบ

Table with 27 rows and 100+ columns. Columns are labeled with Thai characters for each letter (ก, ข, ค, etc.). Rows contain counts of correct/incorrect answers for each letter, followed by a 'Tot' row with percentages, and a final 'Correct' row with totals. Includes a 'TH' subtotal at the bottom right.



ตารางที่ 4.17 ผลการทดสอบระบบรู้จำที่นำเสนอโดยข้อมูลจากสระและวรรณยุกต์ระดับสำหรับทดสอบของกลุ่มผู้เขียนที่ใช้ในการฝึกฝนระบบ

	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	Tot	%	Correct	
อิ	207	7	5	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	223	92.83	207	อิ
อี	1	213	9	4	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	230	92.61	213	อี
อ	1	12	185	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	226	88.05	199	อ
อิอ	1	27	5	180	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	233	85.84	200	อิอ
อ่า	5	0	0	0	196	8	5	0	0	0	0	0	3	1	0	0	0	0	0	0	1	217	90.32	196	อ่า
อัน	0	0	0	0	1	225	0	0	0	3	1	0	0	0	0	0	0	0	0	0	0	230	97.83	225	อัน
อี	0	1	0	0	0	0	221	5	0	0	0	0	0	0	0	0	0	0	0	0	0	227	97.36	221	อี
อ	0	0	0	0	0	0	10	212	1	2	0	0	0	0	0	0	0	0	0	0	0	225	94.22	212	อ
อ้อ	0	0	0	0	0	0	0	0	0	221	1	0	0	0	0	0	0	0	0	0	0	222	99.55	221	อ้อ
อีอ	0	0	0	0	0	0	0	1	0	0	230	0	0	0	0	0	0	0	0	0	0	231	99.57	230	อีอ
อ้อ	0	1	0	0	0	0	0	0	0	0	1	232	5	0	0	0	0	0	0	0	0	234	99.15	232	อ้อ
อ่อ	0	0	0	0	0	0	0	0	0	0	0	6	214	0	0	0	0	0	0	0	220	97.27	214	อ่อ	
												รวม										2498	94.32	2356	
	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ	อิ	อี	อ				

ตารางที่ 4.18 ผลการทดสอบระบบรู้จักที่นำเสนอโดยใช้ข้อมูลจากระบบและวารณยุกต์ระดับของกลุ่มผู้เขียนอิสระ

	อิ	อี	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	Tot	%	Correct	
อิ	164	0	14	7	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	86.32	164	อิ
อี	9	161	15	0	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	190	84.74	161	อี
อ	7	10	166	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190	87.37	166	อ
อิ	5	21	1	157	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	190	82.63	157	อิ
อี	15	0	0	1	167	7	0	0	0	0	0	0	0	0	0	0	0	0	0	190	87.90	167	อี
อ	0	1	0	0	0	189	0	0	0	0	0	0	0	0	0	0	0	0	0	190	99.47	189	อ
อิ	0	1	0	3	1	0	182	3	0	0	0	0	0	0	0	0	0	0	0	190	95.79	182	อิ
อี	0	2	0	0	1	0	10	175	0	2	0	0	0	0	0	0	0	0	0	190	92.11	175	อี
อ	2	0	2	0	12	0	0	0	0	174	0	0	0	0	0	0	0	0	0	190	91.58	174	อ
อิ	1	1	6	0	0	0	0	1	0	181	0	0	0	0	0	0	0	0	0	190	95.26	181	อิ
อี	0	0	0	0	8	0	0	3	0	0	179	3	0	0	0	0	0	0	0	190	94.21	179	อี
อ	1	0	0	0	1	0	0	0	0	1	8	179	0	0	0	0	0	0	0	190	94.21	179	อ
																				2280	90.96	2074	
	อิ	อี	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ	อิ	อ				

#### 4.6 วิเคราะห์การทดสอบประสิทธิภาพของการรู้จำด้วยระบบรู้จำที่นำเสนอ

จากผลการทดสอบประสิทธิภาพของระบบรู้จำที่นำเสนอพบว่าแก้ไขกรณีที่มีอัตราการรู้จำต่ำในการรู้จำด้วยแบบจำลองฮิดเดนมาร์คอปได้ส่วนใหญ่ มีเพียงในกรณีของการตัดสินใจผิดในกลุ่มที่มีความแตกต่างกันที่ความกว้างของตัวอักษรที่ยังไม่สามารถแก้ไขได้ด้วยวิธีการที่นำเสนอ เนื่องจากลายมือเขียนของแต่ละบุคคลมีความกว้างของตัวอักษรโดยเฉลี่ยไม่เท่ากันจึงยากที่จะสามารถรู้จำได้ถูกต้องโดยใช้ข้อมูลทางด้านอัตราส่วนความกว้างต่อความสูงมาช่วยในการรู้จำ และยังมีความผิดพลาดเนื่องจากกรณีต่างๆ ดังที่จะกล่าวต่อไป

##### 4.6.1 ความผิดพลาดจากการวัดค่าความน่าจะเป็นภายหลังของแบบจำลองฮิดเดนมาร์คอป

ในบางตัวอักษรที่ตัดสินใจผิดพลาดอาจเนื่องมาจากการที่ความน่าจะเป็นภายหลังของแบบจำลองฮิดเดนมาร์คอปของแบบจำลองที่ถูกต้องนั้นมีความห่างจากความน่าจะเป็นภายหลังสูงสุดมาก ทำให้ค่าตัวแปรพีชชีของความน่าจะเป็นภายหลังมีค่าน้อยมาก หรือเท่ากับศูนย์ ทำให้ข้อมูลจากลักษณะบ่งความต่างไม่สามารถแก้ไขได้

##### 4.6.2 ความผิดพลาดจากการวัดตัวแปรพีชชีของหัวตัวอักษร

ความผิดพลาดที่เกิดจากการวัดหัวของตัวอักษรสามารถแบ่งได้เป็น 2 ประเภทคือ

1. การแบ่งส่วนหัวตัวอักษรที่ผิดพลาด การตัดส่วนหัวออกมาพิจารณามีการผิดพลาดอยู่ 2 แบบ คือ การตัดส่วนที่เป็นหัวเกินออกมาซึ่งเรามักพบในการตัดส่วนหัวของตัวอักษร 'ก' ซึ่งวางสมมติฐานไว้ว่าเส้นย่อยแรกของ 'ก' จะเป็นเส้นตรงยาว ในลายมือเขียนบางครั้งอาจพบเส้นตรงแรกของตัวอักษรนั้นสั้นมากทำให้ตัดส่วนของหัวผิด อีกกรณีหนึ่งคือการเขียนหัวเล็กเกินไปเมื่อผ่านกระบวนการก่อนหน้ามา ทำให้ส่วนที่เป็นหัวหายไปหมดทำให้ส่วนของหัวไม่มีในตัวอักษรนั้น ตัวอย่างของลายมือเขียนที่ตัดส่วนของหัวผิดพลาดสามารถแสดงได้ดังรูปที่ 4.11



รูปที่ 4.11 การแบ่งส่วนหัวตัวอักษรที่ผิดพลาด

2. การจรดปากกาโดยบังเอิญที่ไม่สามารถลบออกด้วยกระบวนการก่อนหน้า การเขียนตัวอักษรบนกระดาษอิเล็กทรอนิกส์ในตอนที่จรดปากกา และยกปากกาขึ้นมักเกิดเส้นที่ไม่ต้องการขึ้น ตามปกติจะถูกกำจัดออกในกระบวนการก่อนหน้า แต่บางกรณีเส้นที่ไม่ต้องการนั้นมีขนาดยาวจึงไม่ถูกกำจัดออก จึงถูกวัดว่าพบหัว ทำให้ตัวแปรพีชชีของการไม่มีหัวมีค่าต่ำ ดังแสดงรูปตัวอักษรตัวอย่างที่มีเส้นเกินจากการจรดปากกาในรูปที่ 4.12



รูปที่ 4.12 ตัวอักษร 'ก' ที่มีเส้นจากการจรดปากกาเกินมา และไม่ถูกกำจัดออกในกระบวนการก่อนหน้า

#### 4.6.3 ความผิดพลาดจากการวัดตัวแปรพีชชีของอัตราส่วนความกว้างต่อความสูงของตัวอักษร

จากการทดลองพบว่าอัตราส่วนความกว้างต่อความสูงไม่สามารถปรับปรุงตัวอักษรคู่ ('บ' กับ 'ข') และ ('ป' กับ 'ช') ให้มีอัตราการจัดขึ้นได้ เนื่องจากลายมือเขียนของแต่ละบุคคลมีความกว้างของตัวอักษรไม่เหมือนกัน บางคนเขียนตัวอักษรแคบ บางคนเขียนตัวอักษรกว้าง จนกระทั่งบางครั้งตัวลายมือเขียนอักษร 'ข' ของคนหนึ่ง กว้างกว่าตัวอักษร 'บ' ของอีกคนหนึ่ง ดังแสดงตัวอย่างในรูปที่ 4.13 การใช้กฎเพื่อแยกแยะตัวอักษรเหล่านี้เมื่อปรับช่วงอัตราส่วนความกว้างต่อความสูงของตัวแปรพีชชีตัวอักษรแคบมีค่าเป็น 1 กว้างขึ้น ตัวอักษรของผู้เขียนที่มีความกว้างลายมือเขียนแคบมีการตัดสินใจผิดมากขึ้น ในทางตรงกันข้ามหากลดช่วงอัตราส่วนความกว้างต่อความสูงให้แคบลง จะทำให้ตัวอักษรของผู้เขียนที่มีความกว้างลายมือเขียนกว้างมีการตัดสินใจผิดมากขึ้น ดังนั้นการตัดสินใจโดยใช้อัตราส่วนความกว้างต่อความสูงเป็นข้อมูลประกอบจึงไม่มีประสิทธิภาพที่ดีขึ้นสำหรับกลุ่มตัวอักษรเหล่านี้



รูปที่ 4.13 ตัวอักษร 'ข' (รูปทางซ้าย) ที่มีความกว้างใกล้เคียงหรือมากกว่า ตัวอักษร 'บ' (รูปทางขวา)

#### 4.6.4 ความผิดพลาดจากการวัดตัวแปรฟิซซีของตำแหน่งจุดต้นและจุดปลายของตัวอักษร

ตำแหน่งจุดต้นและจุดปลายของตัวอักษรมักไม่พบความผิดพลาดเกิดขึ้น จะมีก็เพียงเล็กน้อยเช่น จุดต้นของตัวอักษร ‘ฒ’ ไปอยู่ด้านบนสุดของตัวอักษร ดังรูปที่ 4.14 จึงไม่สามารถจะแก้ไขได้หาก ตัวแปรฟิซซีของความน่าจะเป็นเป็นภายหลังจากแบบจำลองฮิดเดนมาร์คอฟของ ‘ฒ’ มีค่าสูงกว่าของ ‘ฒ’ อย่างไรก็ตามจากการทดลองพบกรณีเช่นนี้น้อยมาก



รูปที่ 4.14 จุดต้นของตัวอักษร ‘ฒ’ ที่อยู่ด้านบนสุดของตัวอักษร

#### 4.6.5 ความผิดพลาดจากการวัดตัวแปรฟิซซีของรอยหยักของตัวอักษร

ความผิดพลาดจากการวัดรอยหยักส่วนเกิดขึ้นได้ 2 กรณีคือ

1. การไม่เขียนรอยหยักหรือรูปแบบผิดเนื่องจากตัวอักษรมีความซับซ้อนมาก กรณีนี้มักเกิดกับตัวอักษร ‘ฎ’ ที่มีรายละเอียดของตัวอักษรมากกว่าตัวอื่น ในบางครั้งจึงพบการเขียนที่ผิดไปจากรูปแบบปกติได้บ่อยดังเช่นตัวอย่างในรูปที่ 4.15



รูปที่ 4.15 ตัวอักษร ‘ฎ’ ที่เขียนผิดรูปแบบทำให้หารอยหยักด้านล่างไม่พบ

2. การเขียนเส้นหักกลายเป็นวงรอบ จากอัลกอริทึมในการตรวจจับเส้นหยัก การตรวจจับเส้นที่มีการเปลี่ยนมุมระหว่างเส้นย่อยเป็นเครื่องหมายตรงกันข้ามกับมุมระหว่างเส้นย่อยก่อนหน้า โดยในที่นี้จะเรียกจุดหักมุม ในการเขียนตัวอักษรจริงๆ บางครั้งจุดที่ควรจะเป็นจุดหักมุมในตัวอักษร อาจเขียนเป็นวงรอบเล็กๆแทน ดังลายมือเขียนตัวอย่างในรูปที่ 4.16 การตรวจจับเส้นหยักจึงหาไม่พบทำให้มีการวัดที่ผิดพลาดไป



รูปที่ 4.16 วงรอบในตัวอักษรที่แทนเส้นหัก

#### 4.6.6 ความผิดพลาดจากการวัดตัวแปรพีชชีของตำแหน่งของจุดตัดในตัวอักษร

ความผิดพลาดนี้มักเกิดดังแสดงในรูปที่ 4.17 คือการเขียนวงรอบโดยไม่ตั้งใจ และการเขียนที่ไม่เกิดจุดตัด

รูปที่ 4.17 การเขียนวงรอบโดยไม่ตั้งใจ และการไม่เกิดจุดตัด

#### 4.6.7 ความผิดพลาดจากการวัดตัวแปรพีชชีของจำนวนขีดด้านท้ายตัวอักษร

ความผิดพลาดแบ่งได้เป็น 2 กรณีคือ

1. เส้นตรงท้ายที่ไม่ถูกตรวจจับ เส้นตรงท้ายที่ควรจะถูกตรวจจับบางครั้งอาจวัดออกมาผิดพลาดเนื่องจากเส้นท้ายนั้นเอียงมากเกินไป ค่าตัวแปรพีชชีจึงออกมามีค่าน้อยหรือเท่ากับศูนย์ อีกกรณีหนึ่งคือการที่เส้นตรงท้ายบางครั้งพบว่าเขียนมีลักษณะเป็นจุด และกำจัดออกไปตั้งแต่กระบวนการก่อนหน้าทำให้ตรวจจับไม่ได้ ดังแสดงตัวอย่างลายมือเขียนทั้ง 2 กรณีในรูปที่ 4.18

รูปที่ 4.18 เส้นที่ไม่ถูกนับ

2. การตรวจจับเส้นท้ายเกิน



รูปที่ 4.19 การนับเส้นเกิน

ตารางที่ 4.19 เปรียบเทียบสมรรถนะของระบบรู้จำที่นำเสนอกับระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์ค

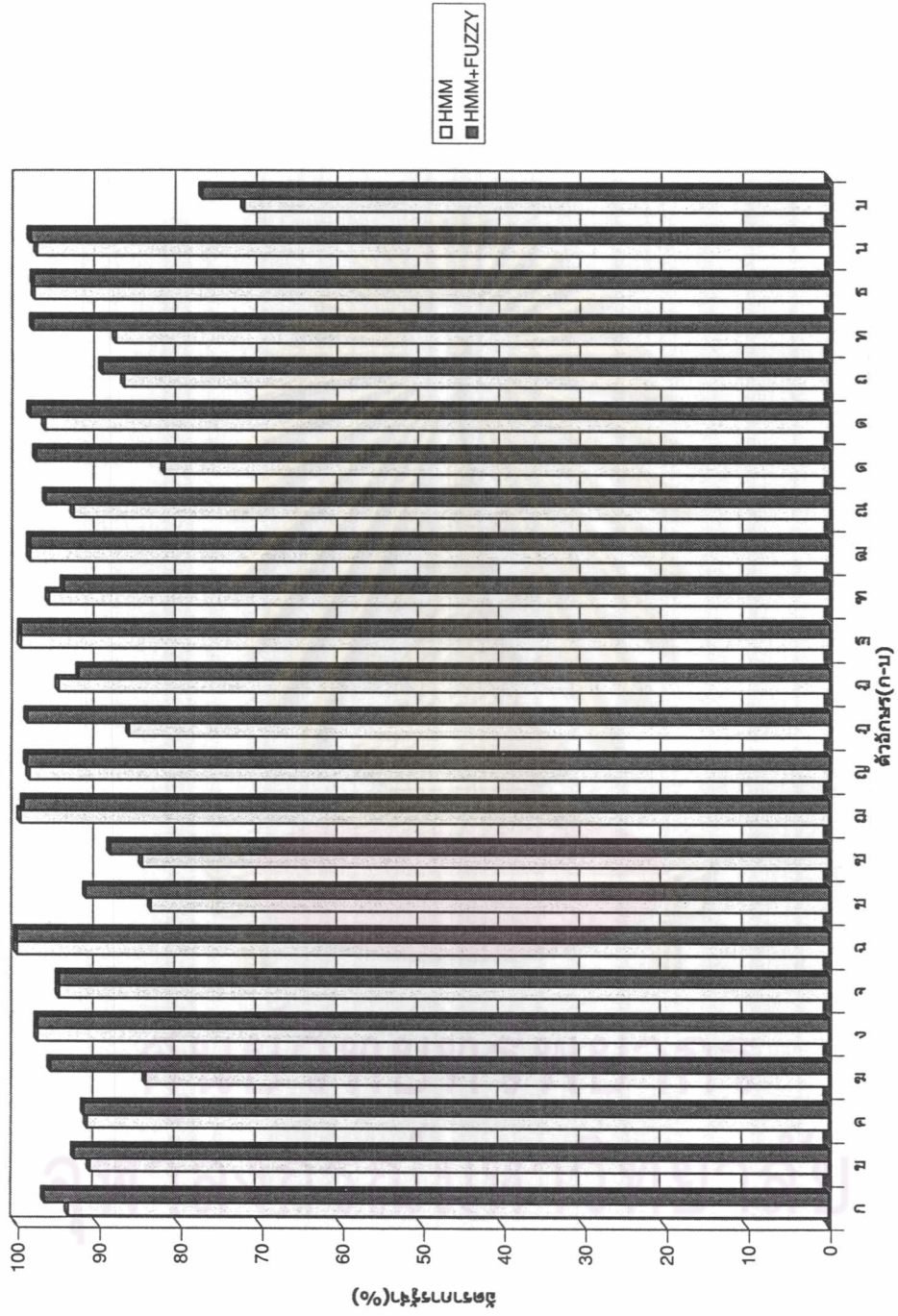
คอฟ

	อัตราการเรียนรู้ตัวอักษรระดับบน(%)		อัตราการเรียนรู้ตัวอักษรระดับกลาง(%)		เวลาในการคำนวณ (วินาที/ตัวอักษร)
	ผู้เขียนที่ฝึกฝนระบบ	ผู้เขียนอิสระ	ผู้เขียนที่ฝึกฝนระบบ	ผู้เขียนอิสระ	
HMM	93.29	84.31	93.53	89.09	0.08
HMM+FUZZY	94.32	90.96	95.66	93.27	0.09

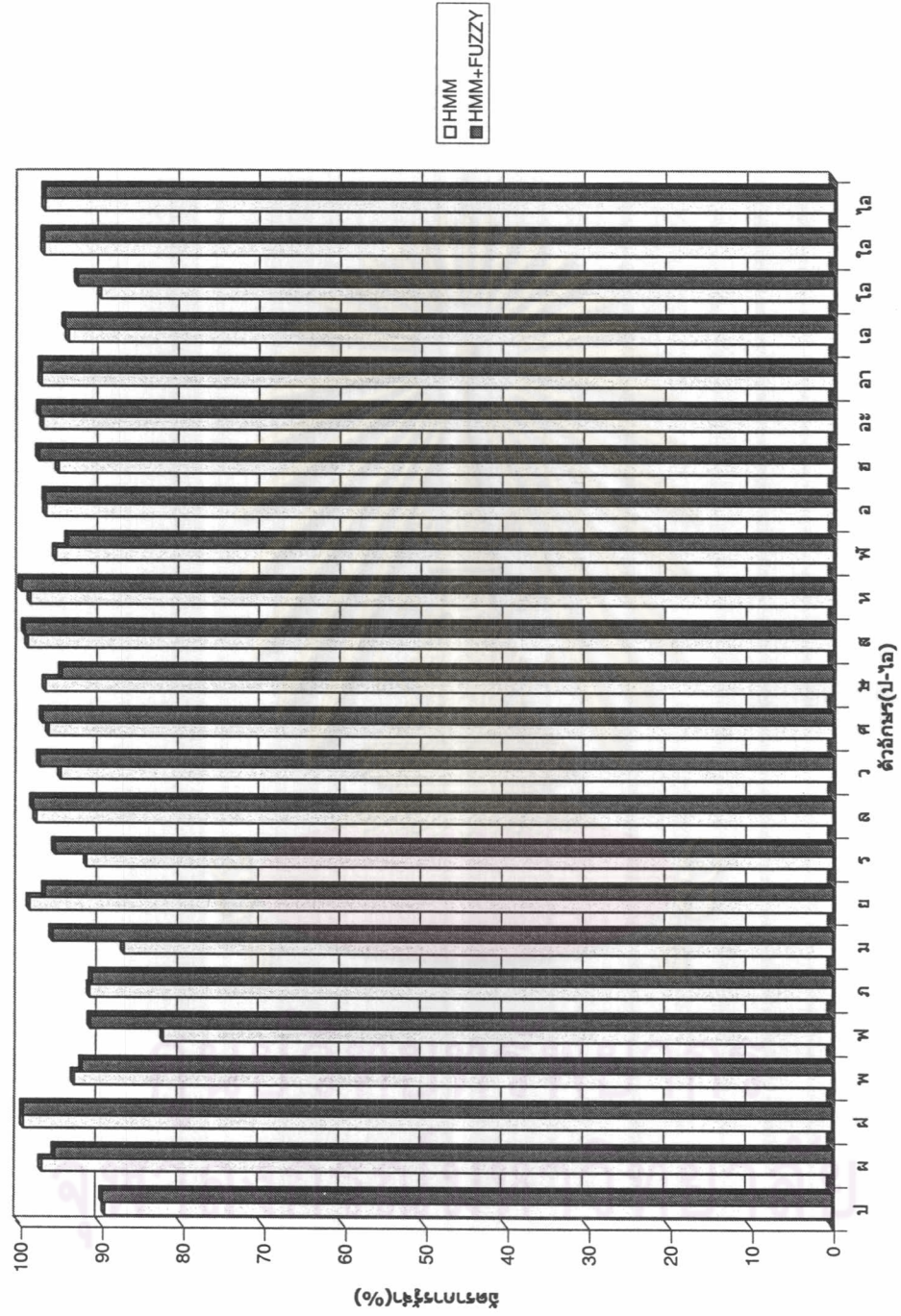
เส้นในตัวอักษรบางครั้งมีคุณสมบัติคล้ายเส้นทแยงของตัวอักษรจึงถูกตรวจจับขึ้นมาทำให้จำนวนเส้นทแยงมีค่าเกินกว่าที่ควรจะเป็นดังแสดงในรูปตัวอย่างที่ 4.19 จึงเกิดการรู้จำที่ผิดพลาดขึ้นได้

#### 4.7 การเปรียบเทียบประสิทธิภาพของระบบรู้จำทั้งสอง

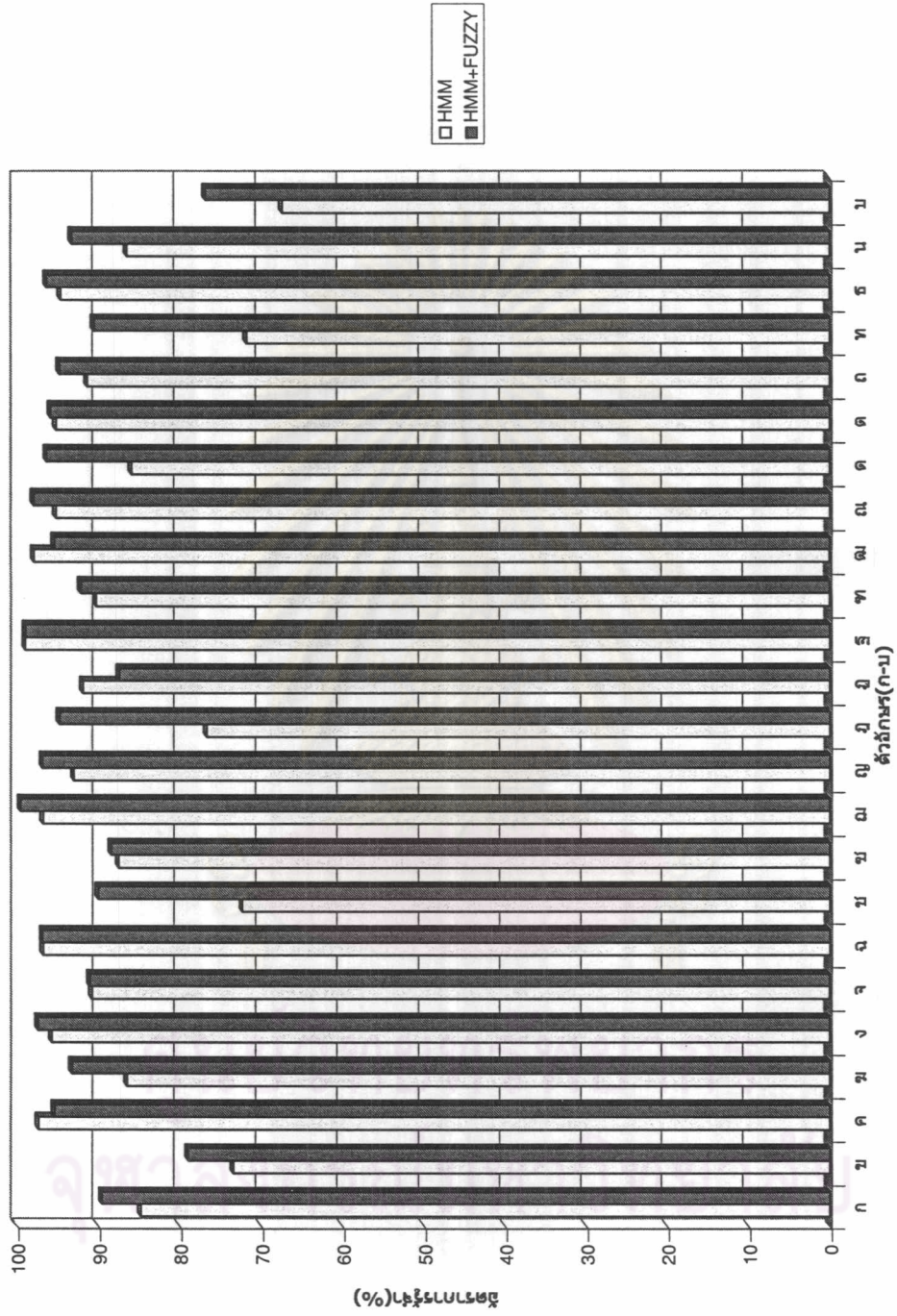
จากผลการทดสอบประสิทธิภาพในหัวข้อก่อนหน้านี้แสดงให้เห็นว่าระบบรู้จำที่นำเสนอสามารถรู้จำตัวอักษรที่มีลักษณะใกล้เคียงกับตัวอักษรอื่นได้ดีกว่าระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คคอฟมาก ในขณะที่ตัวอักษรบางตัวนั้นม้อัตรการเรียนรู้จาลดลงเพียงเล็กน้อย และยังคงอยู่ในเกณฑ์ที่ดี จึงทำให้อัตราการเรียนรู้เฉลี่ยของระบบรู้จำที่นำเสนอสูงกว่าระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คคอฟร้อยละ 4.17 ในตัวอักษรระดับกลางของกลุ่มผู้เขียนอิสระ และร้อยละ 6.65 ในตัวอักษรระดับบนของกลุ่มผู้เขียนอิสระ สำหรับเวลาในการคำนวณของระบบรู้จำที่นำเสนอเพิ่มขึ้นจากระบบรู้จำด้วยแบบจำลองฮิดเดนมาร์คคอฟประมาณร้อยละ 14.10 และอยู่ในระดับที่จัดว่าทำงานได้เร็วมากเพราะสามารถรู้จำตัวอักษรได้ประมาณ 11 ตัวอักษรต่อวินาที หรือ 666 ตัวอักษรต่อนาที ซึ่งเพียงพอต่อการใช้งานเนื่องจากเทียบจากความเร็วในการเขียนของคนทั่วไปประมาณ 2 ตัวอักษรต่อวินาที ในตารางที่ 4.19 แสดงการเปรียบเทียบสมรรถนะโดยรวมของระบบรู้จำทั้งสอง ส่วนอัตราการเรียนรู้เปรียบเทียบของตัวอักษรแต่ละตัวในแต่ละชุดข้อมูลสามารถแสดงได้ในรูปที่ 4.20 ถึงรูปที่ 4.25



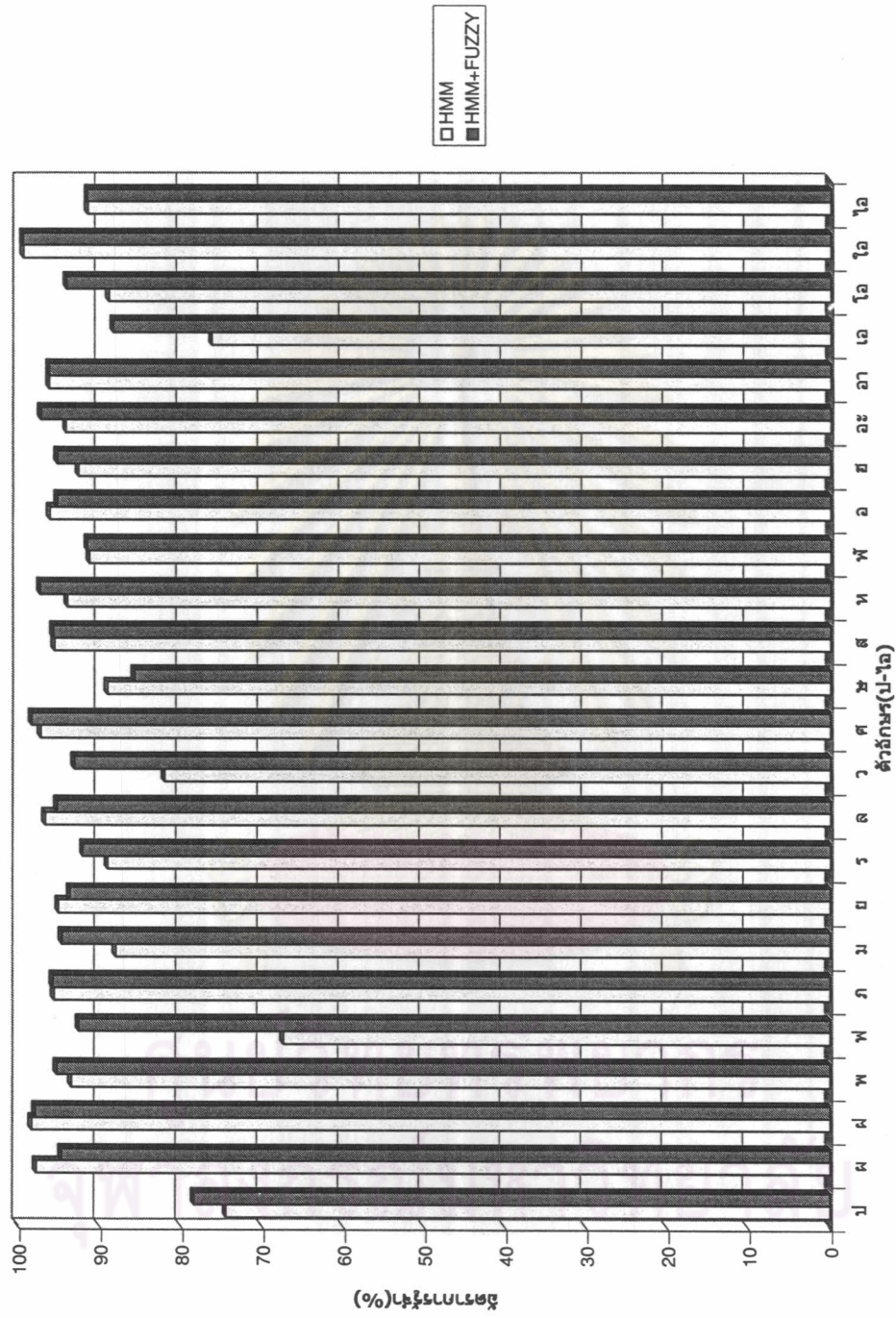
รูปที่ 4.20 แผนภูมิเปรียบเทียบอัตราการเรียนรู้จำตัวอักษรระดับกลาง (ก-บ) ของข้อมูลตัวอักษรจากกลุ่มผู้เขียนที่ใช้ในการฝึกฝนระบบ



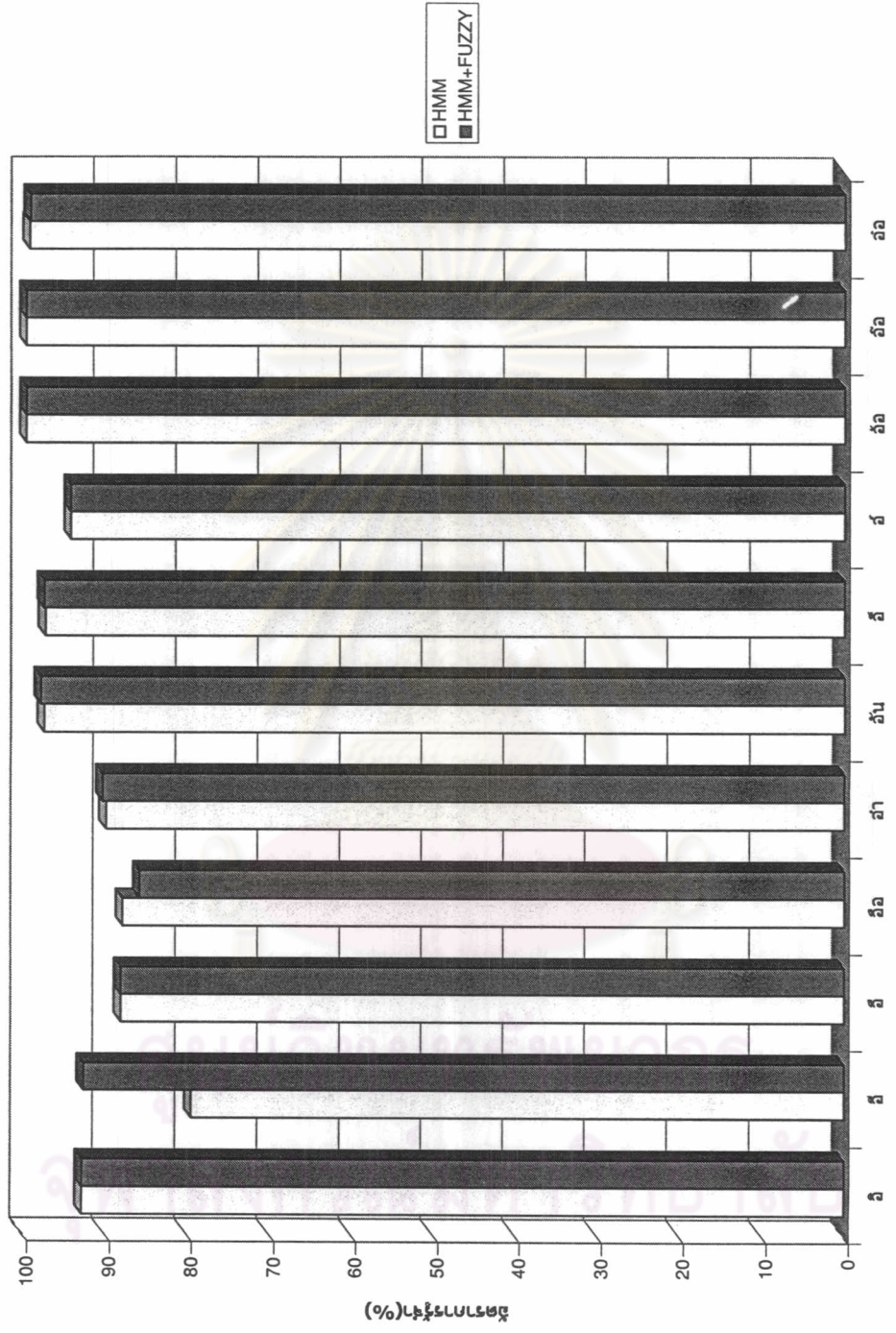
รูปที่ 4-21 แผนภูมิเปรียบเทียบอัตราการจัดจำตัวอักษรระดับกลาง (ป-ไอ) ของข้อมูลตัวอักษรจากกลุ่มผู้เชี่ยวชาญที่ใช้ในการฝึกฝนระบบ



รูปที่ 4.22 แผนภูมิเปรียบเทียบอัตราการรู้จำตัวอักษรระดับกลาง (ก-ข) ของข้อมูลตัวอักษรจากกลุ่มผู้เขียนอิสระ



รูปที่ 4.23 แผนภูมิเปรียบเทียบอัตราการรู้จำตัวอักษรระดับกลาง (ป-ไอ) ของข้อมูลตัวอักษรจากผู้เขียนอิสระ



รูปที่ 4.24 แผนภูมิเปรียบเทียบอัตราการรู้จำตัวอักษรระดับบนของข้อมูลตัวอักษรจากกลุ่มผู้เขียนที่ใช้ในการฝึกในระบบ



รูปที่ 4.25 แผนภูมิเปรียบเทียบอัตราการรู้จำตัวอักษรระดับบนของข้อมูลตัวอักษรจากกลุ่มผู้เขียนอิสระ