

บทที่ 4

การทดลองและผลการทดลอง

ในบทนี้ได้กล่าวถึงการกำหนดสิ่งแวดล้อมในการทดลอง การกำหนด Benchmark และผลการทดลอง สามารถอธิบายได้ดังนี้

4.1 การกำหนดสิ่งแวดล้อมในการทดลอง (Experiment setting)

การกำหนดสิ่งแวดล้อมในการทดลอง (Experiment setting) มีจุดประสงค์เพื่อกำหนดจำนวนและประเภทแฟ้มเสียงที่นำมาทดลองค้นหาคำหลักด้วยโมเดลคำหลักที่สร้างขึ้น เพื่อทดลองหาคำหลัก 10 และ 15 คำ ทั้ง 3 โดเมน ดังแสดงไว้ในตารางที่ 4.1

ตารางที่ 4.1

การกำหนดสิ่งแวดล้อมในการทดลอง

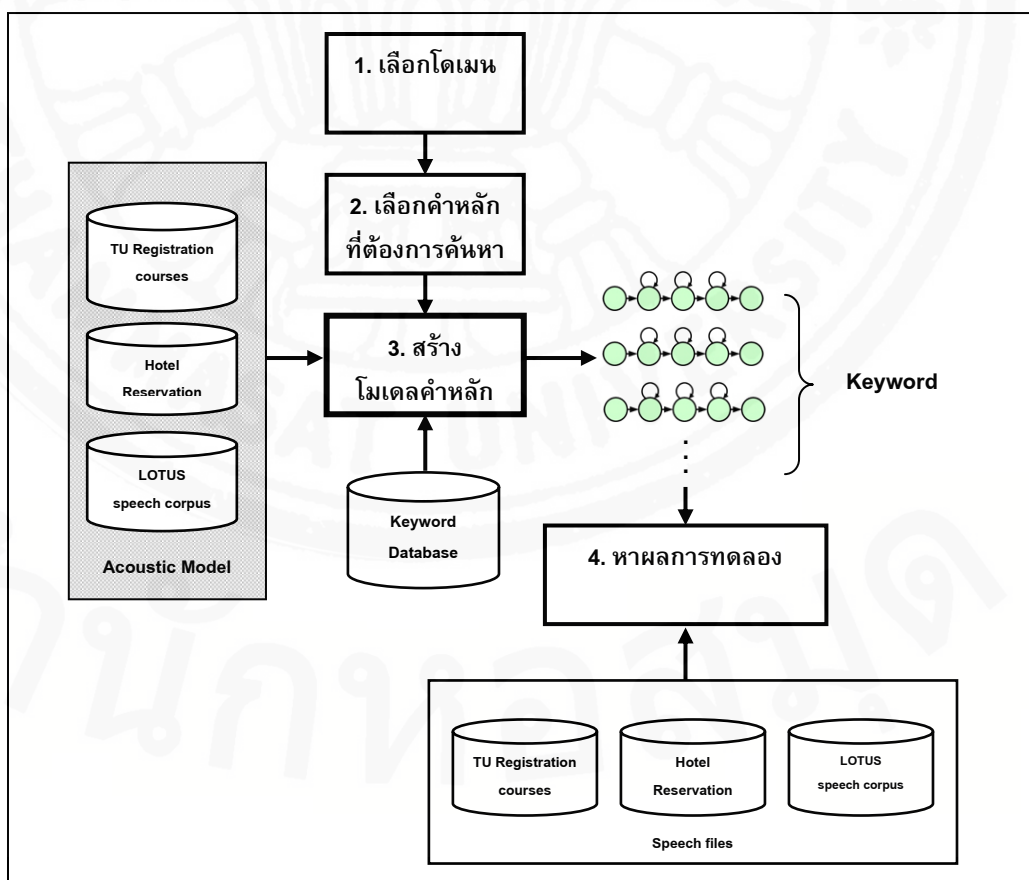
โดเมน	จำนวนแฟ้มที่ใช้ทดลอง			
	แฟ้มเสียงที่ใช้ฝึกฝน (Seen files)	แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน (Unseen files)	แฟ้มเสียงที่ใช้และแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน (Seen and Unseen files)	
			แฟ้มเสียงที่ใช้ฝึกฝน (Seen files)	แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน (Unseen files)
ลงทะเบียนนักศึกษา	100	100	50	50
จองห้องพัก ในโรงแรม	100	100	50	50
LOTUS speech corpus	366	366	168	168

การทดลองได้ทำการสร้างโมเดลคำหลัก จากโมเดลเสียงที่ฝึกฝนไว้ ดังนี้

1. สร้างโมเดลคำหลักจากโมเดลเสียงที่ฝึกฝนในแต่ละโดเมน (Dependent Domain) ดังภาพที่ 4.1 โดยมีขั้นตอน ดังนี้
 - 1.1 เลือกโดเมนที่ต้องการ
 - 1.2 เลือกรายการคำหลักที่ต้องการค้นหาที่ละ 10 และ 15 คำ ตามลำดับ จากโดเมนที่เลือก
 - 1.3 โปรแกรมทำการสร้างโมเดลคำหลักจากโมเดลเสียงจากโดเมนที่ได้ฝึกฝนไว้แล้ว ตามโดเมนที่เลือกไว้ในข้อ 1.
 - 1.4 หาผลการลองโดยนำแฟ้มเสียงที่จัดเตรียมเพื่อทดสอบ (ดังตารางที่ 4.1) ตามโดเมนที่เลือกไว้ในข้อ 1.

ภาพที่ 4.1

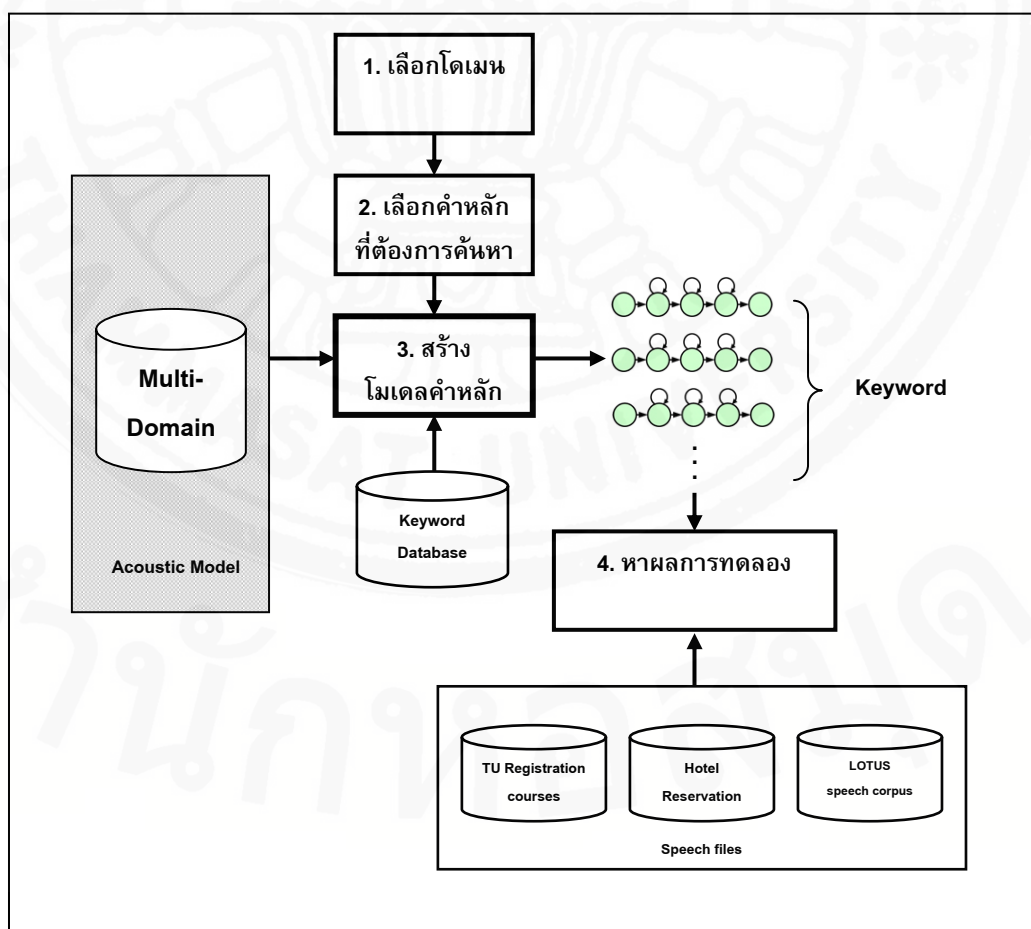
สร้างโมเดลคำหลักจากโมเดลเสียงที่ฝึกฝนในแต่ละโดเมน



2. สร้างโมเดลคำหลักจากโมเดลเสียงที่ฝึกฝนแบบหลายโดเมน (Multi-Domain) ดังภาพที่ 4.2 โดยมีขั้นตอน ดังนี้
 - 2.1 เลือกโดเมน
 - 2.2 เลือกรายการคำหลักที่ต้องการค้นหาที่ละ 10 และ 15 คำ ตามลำดับ ที่ละโดเมน
 - 2.3 โปรแกรมทำการสร้างโมเดลคำหลักจากโมเดลเสียงแบบหลายโดเมนที่ฝึกฝนไว้
 - 2.4 หาผลการทดลอง โดยนำแฟ้มเสียงที่จัดเตรียมเพื่อทดสอบ (ดังตารางที่ 4.1) หาคำหลักจากแฟ้มเสียงตามโดเมนที่เลือกในข้อ 1.

ภาพที่ 4.2

การสร้างโมเดลคำหลักจากโมเดลเสียงที่ฝึกฝนในแต่ละโดเมน



4.2 การกำหนด Benchmark

เพื่อวัดประสิทธิภาพของโมเดลคำหลักแบบอัตโนมัติ ในวิทยานิพนธ์ได้ทดลองค้นหาคำหลักจากโมเดลคำหลักที่ได้ถูกฝึกฝนไว้แล้ว ด้วยโปรแกรม HTK toolkit 3.3 โดยได้ทดลองค้นหาคำหลักด้วยแฟ้มเสียง ดังตารางที่ 4.1 ซึ่งการทดลองนี้ได้ทำการทดลองค้นหาคำหลักจากโมเดลคำหลักที่ฝึกฝนไว้แล้วในแต่ละโดเมน และค้นหาคำหลักจากโมเดลคำหลักที่ฝึกฝนไว้แล้วแบบหลายโดเมน ผลการทดลอง ดังตารางที่ 4.2 ถึง ตารางที่ 4.7 ซึ่งผลการทดลองที่ได้จะถูกใช้เป็นเกณฑ์วัดประสิทธิภาพในการค้นหาคำหลักด้วยการสร้างโมเดลคำหลักแบบอัตโนมัติ

ตารางที่ 4.2

ตารางผลการทดลองการค้นหาคำหลักในโดเมนลงทะเบียนนักศึกษา

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงลงทะเบียนนักศึกษา

เพื่อกำหนด Benchmark

ประเภทแฟ้มเสียง ที่ใช้ในการทดสอบ	โดเมน ลงทะเบียนนักศึกษา							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall-Out	Precision	Recall	F Measure	Fall out
แฟ้มเสียงที่ใช้ฝึกฝน	0.624	0.093	0.162	0.066	0.914	0.137	0.238	0.015
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.606	0.088	0.153	0.067	0.933	0.135	0.236	0.011
แฟ้มเสียงที่ใช้ฝึกฝน และแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.599	0.093	0.162	0.074	0.939	0.146	0.253	0.011

ตารางที่ 4.3

ตารางผลการทดลองการค้นหาคำหลักในโดเมนจองห้องพักในโรงแรม
ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงจองห้องพักในโรงแรม
เพื่อกำหนด Benchmark

ประเภทเพิ่มเสียงที่ ใช้ในการทดสอบ	โดเมน จองห้องพักในโรงแรม							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall out
เพิ่มเสียงที่ใช้ฝึกฝน	0.626	0.235	0.341	0.225	0.927	0.348	0.506	0.044
เพิ่มเสียงที่ไม่ได้ใช้ฝึกฝน	0.544	0.256	0.348	0.404	0.760	0.357	0.486	0.213
เพิ่มเสียงที่ใช้ฝึกฝน และเพิ่มเสียงที่ไม่ได้ใช้ฝึกฝน	0.642	0.239	0.348	0.212	0.883	0.329	0.479	0.070

ตารางที่ 4.4

ตารางผลการทดลองการค้นหาคำหลักในโดเมน LOTUS speech corpus
ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียง LOTUS speech corpus
เพื่อกำหนด Benchmark

ประเภทเพิ่มเสียงที่ ใช้ในการทดสอบ	โดเมน LOTUS speech corpus							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall out
เพิ่มเสียงที่ใช้ฝึกฝน	0.892	0.062	0.117	0.008	0.985	0.069	0.129	0.001
เพิ่มเสียงที่ไม่ได้ใช้ฝึกฝน	0.827	0.053	0.099	0.012	0.964	0.061	0.116	0.002
เพิ่มเสียงที่ใช้ฝึกฝน และเพิ่มเสียงที่ไม่ได้ใช้ฝึกฝน	0.800	0.051	0.096	0.014	0.967	0.062	0.116	0.002

ตารางที่ 4.5

ตารางผลการทดลองการค้นหาคำหลักในโดเมนลงทะเบียนนักศึกษา

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน

เพื่อกำหนด Benchmark

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน ลงทะเบียนนักศึกษา							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall out
แฟ้มเสียงที่ใช้ฝึกฝน	0.671	0.100	0.175	0.058	0.971	0.145	0.253	0.005
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.594	0.086	0.151	0.069	0.928	0.135	0.235	0.012
แฟ้มเสียงที่ใช้ฝึกฝนและแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.627	0.098	0.169	0.069	0.943	0.147	0.255	0.010

ตารางที่ 4.6

ตารางผลการทดลองการค้นหาคำหลักในโดเมนจองห้องพักในโรงแรม

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน

เพื่อกำหนด Benchmark

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน จองห้องพักในโรงแรม							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall out
แฟ้มเสียงที่ใช้ฝึกฝน	0.609	0.229	0.332	0.235	0.888	0.333	0.485	0.067
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.544	0.256	0.348	0.404	0.760	0.357	0.486	0.213
แฟ้มเสียงที่ใช้ฝึกฝนและแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.636	0.237	0.345	0.216	0.883	0.329	0.479	0.070

ตารางที่ 4.7

ตารางผลการทดลองการค้นหาคำหลักในโดเมน LOTUS speech corpus

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน

เพื่อกำหนด Benchmark

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน LOTUS speech corpus							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall out
แฟ้มเสียงที่ใช้ฝึกฝน	0.892	0.062	0.117	0.008	0.985	0.069	0.129	0.001
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.809	0.052	0.097	0.013	0.973	0.062	0.117	0.002
แฟ้มเสียงที่ใช้ฝึกฝนและแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.778	0.050	0.093	0.015	0.978	0.062	0.117	0.002

จากนั้นทำการหาทดลองค้นหาคำหลักด้วยโมเดลคำหลักทั้ง 2 ประเภท โดยผลการทดลองแสดงไว้ในหัวข้อต่อไป

4.3 ผลการทดลอง

จากการทดลองสร้างโมเดลคำหลักแบบอัตโนมัติ เพื่อใช้สำหรับค้นหาคำหลักในแฟ้มเสียงของโดเมนทั้ง 3 โดเมน โดยไม่จำเป็นต้องฝึกฝนโมเดลเสียงให้รู้จำคำหลักที่เพิ่มเข้ามาใหม่ได้ผลการทดลอง ดังนี้

1. ผลการทดลองค้นหาคำหลักโดยใช้โมเดลคำหลักที่สร้างจากโมเดลเสียงที่ฝึกฝนในแต่ละโดเมน ได้ผลการวัดประสิทธิภาพในการทดลองตามที่แสดงไว้ในตารางที่ 4.8 ถึง ตารางที่ 4.10 โดยผลการค้นหาคำหลักที่ดีที่สุดในแต่ละโดเมนแสดงด้วยลักษณะแบบอักษรตัวหนา
2. ผลการทดลองค้นหาคำหลักโดยใช้โมเดลคำหลักที่สร้างจากโมเดลเสียงที่ฝึกฝนแบบหลายโดเมน ได้ผลการวัดประสิทธิภาพในการทดลองตามที่แสดงไว้ในตาราง

ที่ 4.11 ถึง ตารางที่ 4.13 โดยผลการค้นหาคำหลักที่ดีที่สุดในแต่ละโดเมนแสดงด้วยลักษณะแบบอักษรตัวหนา

3. ผลการทดลองเพื่อเปรียบเทียบประสิทธิภาพการค้นหาคำหลักในข้อ 1 และข้อ 2 เปรียบเทียบกับ Benchmark ที่ตั้งไว้
4. ผลการทดลองเมื่อเพิ่มคำหลักใหม่และค้นหาคำหลัก โดเมนลงทะเบียนนักศึกษาที่เป็น Benchmark เปรียบเทียบกับโมเดลคำหลักที่ถูกสร้างขึ้นแบบอัตโนมัติจากโมเดลเสียงแบบหลายโดเมน

สำหรับผลทดลองในข้อ 1 และ ข้อ 2 นั้นมีจุดประสงค์ในการทดลองเพื่อนำผลลัพธ์ที่ได้มาเปรียบเทียบประสิทธิภาพระหว่างการสร้างโมเดลคำหลักแบบอัตโนมัติจากโมเดลเสียงที่ถูกฝึกฝนด้วยแฟ้มเสียงของแต่ละโดเมน กับโมเดลคำหลักแบบอัตโนมัติจากโมเดลเสียงที่ถูกฝึกฝนด้วยแฟ้มเสียงแบบหลายโดเมน ส่วนการผลทดลองในข้อ 3 มีจุดประสงค์ในการทดลองเพื่อนำผลลัพธ์ที่ได้มาเปรียบเทียบประสิทธิภาพการค้นหาคำหลักจากโมเดลคำหลักที่ถูกฝึกฝน กับโมเดลคำหลักที่สร้างแบบอัตโนมัติ และผลการทดลองในข้อ 4 มีจุดประสงค์ในการทดลองเพื่อทดสอบการสร้างโมเดลคำหลักแบบอัตโนมัติโดยไม่ต้องฝึกฝนระบบใหม่ ที่ได้นำเสนอในวิทยานิพนธ์นี้

ตารางที่ 4.8

ตารางผลการทดลองการค้นหาคำหลักในโดเมนลงทะเบียนนักศึกษา
ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงลงทะเบียนนักศึกษา

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน ลงทะเบียนนักศึกษา							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall-out
แฟ้มเสียงที่ใช้ฝึกฝน	0.624	0.093	0.162	0.066	0.914	0.137	0.238	0.015
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.606	0.088	0.153	0.067	0.933	0.135	0.236	0.011
แฟ้มเสียงที่ใช้ฝึกฝนและแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.599	0.093	0.162	0.074	0.939	0.146	0.253	0.011

ตารางที่ 4.9

ตารางผลการทดลองการค้นหาคำหลักในโดเมนของห้องพักในโรงแรม

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงของห้องพักในโรงแรม

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน ของห้องพักในโรงแรม							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall-out
แฟ้มเสียงที่ใช้ฝึกฝน	0.626	0.236	0.341	0.225	0.927	0.348	0.506	0.044
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.544	0.256	0.348	0.404	0.760	0.357	0.486	0.213
แฟ้มเสียงที่ใช้ฝึกฝน และแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.642	0.239	0.348	0.212	0.883	0.329	0.479	0.070

ตารางที่ 4.10

ตารางผลการทดลองการค้นหาคำหลักในโดเมน LOTUS speech corpus

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียง LOTUS speech corpus

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน LOTUS speech corpus							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall-out
แฟ้มเสียงที่ใช้ฝึกฝน	0.892	0.062	0.117	0.008	0.985	0.069	0.129	0.001
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.827	0.053	0.099	0.012	0.964	0.061	0.116	0.002
แฟ้มเสียงที่ใช้ฝึกฝน และแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.800	0.051	0.096	0.014	0.967	0.062	0.116	0.002

ตารางที่ 4.11

ตารางผลการทดลองการค้นหาคำหลักในโดเมนลงทะเบียนนักศึกษา

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน ลงทะเบียนนักศึกษา							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall-out
แฟ้มเสียงที่ใช้ฝึกฝน	0.671	0.100	0.175	0.058	0.971	0.145	0.253	0.005
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.594	0.086	0.151	0.069	0.928	0.135	0.235	0.012
แฟ้มเสียงที่ใช้ฝึกฝนและแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.627	0.098	0.169	0.069	0.943	0.147	0.255	0.010

ตารางที่ 4.12

ตารางผลการทดลองการค้นหาคำหลักในโดเมนจองห้องพักในโรงแรม

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน จองห้องพักในโรงแรม							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall-out
แฟ้มเสียงที่ใช้ฝึกฝน	0.609	0.229	0.332	0.235	0.888	0.333	0.485	0.067
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.544	0.256	0.348	0.404	0.760	0.357	0.486	0.213
แฟ้มเสียงที่ใช้ฝึกฝนและแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.636	0.237	0.345	0.216	0.883	0.329	0.479	0.070

ตารางที่ 4.13

ตารางผลการทดลองการค้นหาคำหลักในโดเมน LOTUS speech corpus

ด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน

ประเภทเพิ่มเสียงที่ใช้ในการทดสอบ	โดเมน LOTUS speech corpus							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall out
เพิ่มเสียงที่ใช้ฝึกฝน	0.892	0.062	0.117	0.008	0.985	0.069	0.129	0.001
เพิ่มเสียงที่ไม่ได้ใช้ฝึกฝน	0.809	0.052	0.097	0.013	0.973	0.062	0.117	0.002
เพิ่มเสียงที่ใช้ฝึกฝน และเพิ่มเสียงที่ไม่ได้ใช้ฝึกฝน	0.778	0.050	0.093	0.015	0.978	0.062	0.117	0.002

4.3.1 ผลการเปรียบเทียบประสิทธิภาพการค้นหาคำหลักจากโมเดลคำหลักแบบอัตโนมัติกับโมเดลคำหลักใน Benchmark

เมื่อนำผลการทดลองการค้นหาคำหลักจากโมเดลคำหลักแบบอัตโนมัติ ตารางที่ 4.8 ถึง 4.13 มาเปรียบเทียบกับตาราง Benchmark ตารางที่ 4.2 ถึง 4.7 ในแต่ละกรณีตามลำดับ จากเปรียบเทียบประสิทธิภาพพบว่าการสร้างโมเดลคำหลักแบบอัตโนมัติที่นำเสนอในวิทยานิพนธ์นี้ มีประสิทธิภาพในการค้นหาคำหลักได้ดีเทียบเท่ากับ Benchmark 100%

4.3.2 ผลการเปรียบเทียบประสิทธิภาพการค้นหาคำหลักระหว่างโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน กับโมเดลคำหลักที่สร้างจากโมเดลเสียงในแต่ละโดเมน

การเปรียบเทียบประสิทธิภาพการค้นหาคำหลักระหว่างโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน กับโมเดลคำหลักที่สร้างจากโมเดลเสียงในแต่ละโดเมนนั้น ได้นำผลการทดลองจากการค้นหาคำหลักด้วยโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน มาลบกับโมเดลคำหลักที่สร้างจากโมเดลเสียงในแต่ละโดเมน ซึ่งค่าความแตกต่างที่ได้ คือ ประสิทธิภาพที่เพิ่มขึ้น (ค่า Precision ค่า Recall ค่า F-Measure มีค่าเป็นบวก และค่า Fall-out มีค่าเป็นลบ)

หรือลดลง (ค่า Precision ค่า Recall ค่า F-Measure มีค่าเป็นลบ และค่า Fall-out มีค่าเป็นบวก) ของโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน

ตารางที่ 4.14

ตารางเปรียบเทียบประสิทธิภาพในการค้นหาคำหลักระหว่างโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมนกับโมเดลคำหลักที่สร้างจากโมเดลเสียงในโดเมนลงทะเบียนนักศึกษา

ประเภทแฟ้มเสียงที่ใช้ ในการทดสอบ	โดเมน ลงทะเบียนนักศึกษา							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall-out
แฟ้มเสียงที่ใช้ฝึกฝน	0.048	0.007	0.012	-0.008	0.057	0.009	0.015	-0.010
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	-0.011	-0.002	-0.003	0.002	-0.006	-0.001	-0.001	0.001
แฟ้มเสียงที่ใช้ฝึกฝนและ แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.028	0.004	0.008	-0.005	0.005	0.001	0.001	-0.001

จากตารางที่ 4.14 แสดงให้เห็นว่าการค้นหาคำหลักจากโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน มีประสิทธิภาพโดยรวมเพิ่มขึ้นเมื่อเปรียบเทียบกับโมเดลคำหลักที่สร้างจากโมเดลเสียงในโดเมนลงทะเบียนนักศึกษา

ตารางที่ 4.15

ตารางเปรียบเทียบประสิทธิภาพในการค้นหาคำหลักระหว่างโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมนกับโมเดลคำหลักที่สร้างจากโมเดลเสียงในโดเมนจองห้องพักในโรงแรม

ประเภทแฟ้มเสียงที่ใช้ ในการทดสอบ	โดเมน จองห้องพักในโรงแรม							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall out
แฟ้มเสียงที่ใช้ฝึกฝน	-0.017	-0.006	-0.009	0.010	-0.039	-0.015	-0.021	0.023
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
แฟ้มเสียงที่ใช้ฝึกฝนและ แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	-0.006	-0.002	-0.003	0.004	0.000	0.000	0.000	0.000

จากตารางที่ 4.15 แสดงให้ทราบว่าผลการค้นหาคำหลักจากโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน มีประสิทธิภาพโดยรวมลดลงเมื่อเปรียบเทียบกับโมเดลคำหลักที่สร้างจากโมเดลเสียงในโดเมนของห้องพักในโรงแรม

ตารางที่ 4.16

ตารางเปรียบเทียบประสิทธิภาพในการค้นหาคำหลักระหว่างโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมนกับโมเดลคำหลักที่สร้างจากโมเดลเสียงในโดเมน LOTUS speech corpus

ประเภทแฟ้มเสียงที่ใช้ในการทดสอบ	โดเมน LOTUS speech corpus							
	ค้นหาคำหลัก 10 คำ				ค้นหาคำหลัก 15 คำ			
	Precision	Recall	F Measure	Fall Out	Precision	Recall	F Measure	Fall out
แฟ้มเสียงที่ใช้ฝึกฝน	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
แฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	-0.018	-0.001	-0.002	0.001	0.009	0.001	0.001	-0.001
แฟ้มเสียงที่ใช้ฝึกฝนและแฟ้มเสียงที่ไม่ได้ใช้ฝึกฝน	-0.022	-0.001	-0.003	0.002	0.011	0.001	0.001	-0.001

จากตารางที่ 4.16 แสดงให้ทราบว่าผลการค้นหาคำหลักจากโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน มีประสิทธิภาพโดยรวมลดลงเมื่อเปรียบเทียบกับโมเดลคำหลักที่สร้างจากโมเดลเสียงในโดเมน LOTUS speech corpus

4.3.3 สรุปการวิเคราะห์ผลจากตารางเปรียบเทียบประสิทธิภาพในการค้นหาคำหลักระหว่างโมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน กับ โมเดลคำหลักที่สร้างจากโมเดลเสียงในแต่ละโดเมน ได้ดังนี้

- จำนวนคำหลักที่ใช้ในการทดสอบมีผลต่อประสิทธิภาพ เนื่องจากในกระบวนการเปรียบเทียบคำที่ค้นหาได้กับกลุ่มของคำหลักที่กำหนดไว้ในโมเดลภาษา โดยถ้าคำที่ค้นหาได้อยู่ในกลุ่มของคำหลักที่กำหนด มีผลทำให้ค่าความน่าจะเป็นในการที่คำที่ค้นหาได้นั้นจะเป็นคำหลักมีค่าสูงขึ้น ซึ่งจะทำให้ได้ถูกต้องโดยรวมเพิ่มขึ้นด้วย แต่เมื่อทำการพิจารณา ระหว่างการค้นหาคำหลัก 10 คำ และ 15 คำ ในโดเมนของห้องพักใน

โรงแรม (แสดงไว้ใน ภาคผนวก ค ผลการทดลอง 8 กับ ผลการทดลอง 9) ผลของการค้นหาคำหลัก “จอง” เมื่อทำการค้นหาคำหลัก 10 คำ สามารถค้นหาได้มากกว่าการค้นหาคำหลัก 15 คำ เนื่องจาก ในการค้นหาคำหลัก 15 คำ ได้มีการเพิ่มคำหลัก “สอง” ซึ่ง คำทั้ง 2 มีหน่วยเสียงคล้ายคลึงกัน ดังภาพที่ 4.3

ภาพที่ 4.3

ตัวอย่างลำดับหน่วยเสียง คำว่า “จอง” และ “สอง”

จอง	ch	@@	ng [^]
สอง	s	@@	ng [^]

ดังนั้น เมื่อนำแฟ้มเสียงที่มีการออกเสียงคำว่า “จอง” และ “สอง” ไม่ชัดเจนมาค้นหาคำหลักที่กำหนดแล้ว จะทำให้ค่าความน่าจะเป็นของคำที่ได้จากโมเดลเสียงมีค่าใกล้เคียง เป็นผลให้โมเดลภาษาไม่สามารถพิจารณาคำนั้นได้ว่า เป็นคำว่า “จอง” หรือ “สอง” จึงทำให้คำนั้นถูกโมเดลภาษาปฏิเสธออกไปว่าไม่ใช่คำที่อยู่ในภาษา ซึ่งทำให้ค้นหาคำนั้นไม่พบ

- จากการเปรียบเทียบประสิทธิภาพของการค้นหาคำหลักในโดเมนลงทะเบียนนักศึกษา และโดเมน LOTUS speech corpus จะได้ประสิทธิภาพในการค้นหาคำหลัก ดีกว่าโดเมนจองห้องพักในโรงแรม โดยเฉพาะการค้นหาคำหลักจำนวน 15 คำ เนื่องจากแฟ้มเสียงในโดเมนจองห้องพักในโรงแรมเป็นแฟ้มเสียงที่ประกอบด้วยคำอุทาน (spontaneous speech) ซึ่งต่างกับแฟ้มเสียงที่ใช้ทดสอบในโดเมนลงทะเบียนนักศึกษา และโดเมน LOTUS speech corpus ที่เป็นแฟ้มเสียงแบบอ่าน (reading speech) ทำให้สามารถวิเคราะห์ได้ว่าวิธีการที่นำเสนอในวิทยานิพนธ์ฉบับนี้เหมาะสมกับการสร้างโมเดลคำหลัก เพื่อค้นหาคำหลักในแฟ้มเสียงแบบอ่าน มากกว่าหาคำหลักในแฟ้มเสียงแบบที่มีคำอุทาน (spontaneous speech) สำหรับการค้นหาคำหลักในแฟ้มเสียงที่มีคำอุทานประกอบอยู่นั้น จำเป็นต้องใช้โมเดลพิเศษเพื่อกำจัดคำอุทานออกจากแฟ้มเสียง ซึ่งมีได้นำเสนอในวิทยานิพนธ์ฉบับนี้
- โมเดลเสียงที่ใช้สร้างคำหลักมีผลต่อประสิทธิภาพของโมเดลคำหลักที่สร้างขึ้น โดยโมเดลคำหลักที่สร้างขึ้นจากโมเดลเสียงของแต่ละโดเมนให้ประสิทธิภาพโดยรวมดีกว่า

โมเดลคำหลักที่สร้างจากโมเดลเสียงแบบหลายโดเมน แสดงไว้ในตารางที่ 4.14 ถึง ตารางที่ 4.15 (โดยพิจารณาประสิทธิภาพจาก ค่า Precision ค่า Recall ค่า F-Measure มีค่าเป็นลบ และค่า Fall-out มีค่าเป็นบวก ในแต่ละตาราง) เนื่องจากคุณภาพของโมเดลเสียงแบบหลายโดเมนจะแปรผันตามแฟ้มเสียงที่นำมาสร้างโมเดลเสียง ดังนั้นเมื่อนำแฟ้มเสียงจาก Hotel Reservation corpus ซึ่งเป็นแฟ้มเสียงแบบที่มีคำอุทาน (spontaneous speech) มาร่วมฝึกฝนด้วย จึงทำให้คุณภาพโดยรวมของโมเดลเสียงที่ใช้สร้างโมเดลคำหลักลดลงไปด้วย

- จากการทดสอบค้นหาคำหลักในแฟ้มเสียงที่ได้จัดกลุ่มไว้นั้น การค้นหาคำหลักในแฟ้มเสียงที่ใช้ฝึกฝนโมเดลเสียงนั้น ได้ผลถูกต้องมากกว่า (ค่า Precision ที่แสดงไว้ในตารางที่ 4.8 ถึง ตารางที่ 4.13) แฟ้มเสียงที่ไม่ได้ฝึกฝนบางแฟ้มบันทึกมาจากผู้พูดคนละคนกัน จึงทำให้โมเดลคำหลักที่สร้างจากโมเดลเสียง พิจารณาความน่าจะเป็นของหน่วยเสียงผิดพลาดได้ เนื่องจากคุณลักษณะเฉพาะของเสียงพูดในคนแต่ละคน แม้ว่า จะพูดคำเดียวกัน ก็ให้ผลในการรู้จำคำนั้นต่างกัน

4.3.4 ผลการทดลองเมื่อเพิ่มคำหลักใหม่เข้าไปในโดเมน

ในทดสอบการค้นหาคำหลักหลักที่เพิ่มขึ้น ได้มีการบันทึกแฟ้มเสียงของประโยค “ทดสอบการค้นหาคำหลักในเสียงพูด” ซึ่งได้ทำการกำหนดแฟ้มออกเสียง (Transcription File) ดังภาพที่ 4.4 และได้ทำการกำหนดคำหลักเพิ่มขึ้นใหม่ ดังนี้ “คำหลัก” “ทดสอบ” “เสียงพูด” “การ” “ค้นหา”

ภาพที่ 4.4

แฟ้มการออกเสียง ประโยค “ทดสอบการค้นหาคำหลักในเสียงพูด”

```
"*/test.lab"
ทดสอบ
การ
ค้นหา
คำหลัก
ใน
เสียงพูด
.
```

จากนั้นทำการค้นหาคำหลักใหม่จากแฟ้มเสียงดังกล่าวเพียงแฟ้มเดียว ด้วยโมเดลคำหลักที่ฝึกฝนในโดเมนลงทะเบียนนักศึกษาที่เป็น Benchmark ผลปรากฏว่าโดเมนที่เป็น Benchmark ไม่สามารถค้นหาคำหลักที่เพิ่มขึ้นได้ ดังภาพที่ 4.5 สาเหตุที่โดเมนลงทะเบียนนักศึกษาที่ใช้เป็น Benchmark แสดงข้อผิดพลาดว่าไม่สามารถพบโมเดลของคำหลักคำว่า “คำหลัก” ได้ เนื่องจากในขั้นตอนการฝึกฝนไม่ได้ฝึกฝนโมเดลเสียงของโดเมนลงทะเบียนนักศึกษารู้จำคำว่า “คำหลัก” แต่เมื่อทดลองสร้างคำหลักแบบอัตโนมัติด้วยเทคนิคที่เสนอในวิทยานิพนธ์นี้ โดยทำการกำหนดคำหลักใหม่ 5 คำ คือ “คำหลัก” “ทดสอบ” “เสียงพูด” “การ” “ค้นหา” ในโดเมนลงทะเบียนนักศึกษา ดังภาพที่ 4.6 เมื่อทำการทดลองค้นหาคำหลักที่เพิ่มใหม่นั้น สามารถค้นหาคำหลักใหม่ในแฟ้มเสียงที่กำหนดได้ถูกต้อง ดังภาพที่ 4.7

ภาพที่ 4.5

ระบบ Benchmark แสดงข้อผิดพลาด
ในการค้นหาคำหลัก คำว่า “คำหลัก”

```
C:\WINDOWS\system32\cmd.exe
C:\DomainSelector\DomainSelector>HVite -T 1 -C c:\thasr\scp\config2 -H C:\thasr\
\register\hmm20\hmmdef -S C:\DomainSelector\DomainSelector\scp\demo.scp -l '*' -
w C:\thasr\register\dict\wdnet -i C:\thasr\register\hviteout.mlf -p 0.0 -s 5.0 C
:\thasr\register\dict\keyworddict C:\thasr\register\lab\tiedlist
Read 92 physical / 1685 logical HMMs
Read lattice with 10 nodes / 18 arcs
Performing network expansion with word internal contexts
Defined 67 contexts
@ a e i q u v x @@ aa ee ia ii oo qq ua uu vv xx iia uua
vva kh m^ l k^ th o t^ s p^ ng^ ph k n^ h t ch j^ w^ w n p
c j b m kr pl d ng khr f r z pr dr kl thr tr phl f^ phr kw
khw l^ fl
Building word !NULL
Building word !NULL
Building word silence
Building word คำหลัก
ERROR [+8231] GetHCIModel: Cannot find hmm [m^-]l[+a]
FATAL ERROR - Terminating program HVite

C:\DomainSelector\DomainSelector>C:\DomainSelector\DomainSelector\HResult2 -T 1
-t -I C:\DomainSelector\DomainSelector\lab\demo.mlf C:\thasr\register\lab\tiedli
st C:\thasr\register\hviteout.mlf 1>C:\thasr\register\dict\hresultout.txt

C:\DomainSelector\DomainSelector>pause
Press any key to continue . . .
```

แสดงข้อผิดพลาดว่าค้นหา
โมเดลคำว่า “คำหลัก” ไม่พบ

ภาพที่ 4.6

การกำหนดคำหลักเพิ่มในโดเมนลงทะเบียนนักศึกษา
เพื่อสร้างโมเดลคำหลักแบบอัตโนมัติ

<DOMAIN> ลงทะเบียนนักศึกษา
<KEYWORD> ลง, ทะเบียน, ถอน, วิชา, เพิ่ม, ฝึกอบรม, บัณฑิต, เลือก, เสรี, เฉพาะ, หน่วยกิต,
เทอม, หลักสูตร, ภาค, ศึกษา, คำหลัก, ทดสอบ, เสียชีวิต, การ, ค้นหา
.

ภาพที่ 4.7

การค้นหาคำหลักที่เพิ่มขึ้นใหม่ในโดเมนลงทะเบียนนักศึกษา

ทดสอบ = 1
การ = 1
ค้นหา = 1
คำหลัก = 1
ใน = 1
เสียชีวิต = 1
Summary : 6

4.3.5 สรุปการทดลอง

การค้นหาคำหลักด้วยโมเดลคำหลักแบบอัตโนมัติ สามารถแก้ปัญหาของระบบค้นหาคำหลักที่ต้องทำการฝึกฝนระบบใหม่ เมื่อมีคำหลักเพิ่มเข้ามา โดยประสิทธิภาพของโมเดลคำหลักที่สร้างขึ้นนั้นจะขึ้นอยู่กับคุณภาพของโมเดลเสียงกับคุณภาพของแฟ้มเสียงที่ใช้ค้นหา เช่นโมเดลเสียงถูกฝึกฝนด้วยแฟ้มเสียงที่ปราศจากเสียงรบกวน โมเดลคำหลักแบบอัตโนมัติที่สร้างขึ้นจะมีความเหมาะสมในการค้นหาแฟ้มเสียงที่ปราศจากเสียงรบกวนเช่นกัน

สำหรับการสร้างโมเดลคำหลักในวิทยานิพนธ์นี้ ใช้วิธีการสร้างจากพจนานุกรมคำหลักที่ประกอบด้วยคำหลัก และลำดับหน่วยเสียง โดยไม่คำนึงถึงวรรณยุกต์ เช่นคำว่า “ข้าว” และ คำว่า “ขาว” จะมีลำดับของหน่วยเสียงคือ “kh aa w[^]” เหมือนกัน ดังนั้นในการพิจารณาว่าคำหลักที่ค้นหาได้จากแฟ้มเสียงจะพิจารณาจากโดเมนที่กำหนดไว้ เช่น คำว่า “ข้าว” คือคำหลักในโดเมน

กสิกรรม และ คำว่า “ขาว” คือ คำหลักในโดเมนสี ถ้าหากมีการกำหนดคำหลัก คำว่า “ข้าว” และ คำว่า “ขาว” ในโดเมนเดียวกันแล้ว จะทำให้ระบบไม่สามารถค้นหาคำหลักที่กำหนดไว้ได้ถูกต้อง

