

ตารางที่ 4.3 การวิเคราะห์ประสิทธิภาพระบบในคำถามข้อที่ 2

ข้อที่ 2. คำถาม : องค์ประกอบของระบบคอมพิวเตอร์มีอะไรบ้าง

คำตอบ : บุคลากร, ขั้นตอน, ฮาร์ดแวร์, ซอฟต์แวร์, ข้อมูล

ลำดับ ผู้เข้าสอบ	จำนวนคำสำคัญ		ค่าความถูกต้อง (ร้อยละ) (Precision)	ค่าความครบถ้วน (ร้อยละ) (Recall)	ประสิทธิภาพ (ร้อยละ) (F-Measure)
	คำตอบ เฉลี่ย	ระบบ ค้นพบ			
1	5	5	100.0	100.0	100.0
2	5	3	100.0	60.0	75.0
3	5	4	100.0	100.0	100.0
...					
169	5	2	100.0	40.0	57.1
ค่าเฉลี่ย	5	2.91	96.4	58.3	72.7

กรณีที่คำถามระบุให้กรอกคำตอบครบทั้ง 5 คำตอบ การเปรียบเทียบค่าความครบถ้วน (Recall) จึงบังคับให้ผู้สอบต้องตอบให้ครบตามจำนวนของคำตอบเฉลี่ย จึงส่งผลให้ค่า F-Measure มีค่าร้อยละ 72.7 (ผู้ไม่ตอบหรือตอบไม่ถูกต้อง จำนวน 6 คน)

ตารางที่ 4.4 การวิเคราะห์ประสิทธิภาพระบบในคำถามข้อที่ 3

ข้อที่ 3. คำถาม : หน่วยประมวลผลกลางคืออะไร

คำตอบ : ซีพียู, สมอของคอมพิวเตอร์

ลำดับ ผู้เข้าสอบ	จำนวนคำสำคัญ		ค่าความถูกต้อง (ร้อยละ) (Precision)	ค่าความครบถ้วน (ร้อยละ) (Recall)	ประสิทธิภาพ (ร้อยละ) (F-Measure)
	คำตอบ เฉลี่ย	ระบบ ค้นพบ			
1	2	1	100.0	50.0	66.7
2	2	1	100.0	50.0	66.7
3	2	0	0	0	0
...					
169	2	1	100.0	50.0	66.7
ค่าเฉลี่ย	2	1.05	98.2	53.0	68.8

กรณีที่คำตอบที่มีจำนวนน้อย และตอบได้หลากหลาย ผู้สอบจะตอบเพียงคำเดียว การเปรียบเทียบค่าความครบถ้วน (Recall) จึงบังคับให้ผู้สอบต้องตอบให้ครบตามจำนวนของคำตอบเฉลี่ย จึงส่งผลให้ค่า F-Measure มีค่าร้อยละ 68.8 (ผู้ไม่ตอบหรือตอบไม่ถูกต้อง จำนวน 3 คน)

ตารางที่ 4.5 การวิเคราะห์ประสิทธิภาพระบบในคำถามข้อที่ 4

ข้อที่ 4. คำถาม : ประเภทของซอฟต์แวร์มีอะไรบ้าง

คำตอบ : ซอฟต์แวร์ระบบ, ซอฟต์แวร์ประยุกต์

ลำดับ ผู้เข้าสอบ	จำนวนคำสำคัญ		ค่าความถูกต้อง (ร้อยละ) (Precision)	ค่าความครบถ้วน (ร้อยละ) (Recall)	ประสิทธิภาพ (ร้อยละ) (F-Measure)
	คำตอบ เฉลย	ระบบ ค้นพบ			
1	2	2	100.0	100.0	100.0
2	2	1	100.0	50.0	66.7
3	2	2	100.0	100.0	100.0
...					
169	2	2	100.0	100.0	100.0
ค่าเฉลี่ย	2	1.60	85.8	80.2	82.9

กรณีที่คำตอบกำหนดให้ผู้สอบตอบครบจำนวนที่กำหนด และระบบช่วยขยายขอบเขตของคำตอบ ตัวอย่างเช่น ซอฟต์แวร์ระบบ {โปรแกรมระบบ}, ซอฟต์แวร์ประยุกต์ {โปรแกรมประยุกต์} ทำให้การเปรียบเทียบค่าความครบถ้วน (Recall) จึงส่งผลให้ค่า F-Measure มีค่าร้อยละ 82.9 (ผู้ไม่ตอบหรือตอบไม่ถูกต้อง จำนวน 24 คน)

ตารางที่ 4.6 การวิเคราะห์ประสิทธิภาพระบบในคำถามข้อที่ 5

ข้อที่ 5. คำถาม : ระบบปฏิบัติการคืออะไร

คำตอบ : โปรแกรม, ควบคุม, จัดการ, บริหาร, ตรวจสอบ, ฮาร์ดแวร์

ลำดับ ผู้เข้าสอบ	จำนวนคำสำคัญ		ค่าความถูกต้อง (ร้อยละ) (Precision)	ค่าความครบถ้วน (ร้อยละ) (Recall)	ประสิทธิภาพ (ร้อยละ) (F-Measure)
	คำตอบ เฉลย	ระบบ ค้นพบ			
1	6	5	100.0	83.3	90.9
2	6	5	100.0	83.3	90.9
3	6	3	100.0	50.0	66.7
...					
169	6	4	100.0	66.7	80.0
ค่าเฉลี่ย	6	3.20	96.4	53.4	68.7

กรณีที่คำตอบที่สามารถตอบได้อย่างหลากหลาย ไม่จำเป็นต้องตอบอย่างครบถ้วน แต่สามารถครอบคลุมกับคำตอบได้ นอกจากนั้นขอบเขตการขยายของคำตอบที่มีจำนวนเพิ่มขึ้น ตัวอย่างเช่น โปรแกรม {ซอฟต์แวร์}, ตรวจสอบ {ตรวจเช็ค, ตรวจ, เช็ค}, ฮาร์ดแวร์ {อุปกรณ์} ทำให้การเปรียบเทียบค่าความครบถ้วน (Recall) จึงส่งผลให้ค่า F-Measure มีค่าร้อยละ 68.7 (ผู้ไม่ตอบหรือตอบไม่ถูกต้อง จำนวน 6 คน)

ตารางที่ 4.7 แสดงค่าประมาณการค่าเฉลี่ยของการวัดประสิทธิภาพ

ลำดับ ข้อสอบ	ค่าความถูกต้อง (ร้อยละ) (Precision)	ค่าความครบถ้วน (ร้อยละ) (Recall)	ประสิทธิภาพ (ร้อยละ) (F-Measure)
1	95.3	54.7	69.5
2	96.4	58.3	72.7
3	98.2	53.0	68.8
4	85.8	80.2	82.9
5	96.4	53.4	68.7
ค่าเฉลี่ย	94.4	59.9	72.5

จากการสุ่มข้อสอบจำนวน 5 ข้อเพื่อทดสอบสมมุติฐาน พบว่าค่าเฉลี่ยของค่าความถูกต้อง (Precision) มีค่าร้อยละ 94.4, ค่าความครบถ้วน (Recall) มีค่าร้อยละ 59.9 และวัดประสิทธิภาพ (F-Measure) มีค่าร้อยละ 72.5

ตารางที่ 4.8 การแสดงผลการสืบค้นคำสำคัญของคำตอบระหว่างเทคนิคการค้นหา

ข้อที่	จำนวนคำสำคัญ จากคำตอบเฉลี่ย	เทคนิคการค้นหา					การเพิ่มโมดูลตาม แนวทางออนโทโลยี				
		รวมกับการพยากรณ์คำศัพท์									
		คน	คน	คน	คน	คนที่	คน	คน	คน	คน	คนที่
		1	2	3	...	169	1	2	3	...	169
1	5	3	4	4	...	3	4	5	4	...	3
2	5	4	5	3	...	4	5	3	4	...	2
3	2	1	1	0	...	1	1	1	0	...	1
4	2	2	1	2	...	2	2	1	2	...	2
5	6	4	3	3	...	3	5	5	3	...	4

หมายเหตุ การเชื่อมต่อตามแนวคิดของออนโทโลยีทำให้สืบค้นคำสำคัญเพิ่มขึ้น/ลดลง เช่น บุคลากร { คน, เจ้าหน้าที่, พนักงาน }, ขั้นตอน { กระบวนการ, วิธีการ }, ซอฟต์แวร์ระบบ { โปรแกรมระบบ, ระบบปฏิบัติการ }, ซอฟต์แวร์ประยุกต์ { โปรแกรมประยุกต์ }, โปรแกรม { ซอฟแวร์ }, ตรวจสอบ { ตรวจเช็ค, ตรวจ, เช็ค }, ฮาร์ดแวร์ { อุปกรณ์ } เป็นต้น

ตารางที่ 4.9 ผลเปรียบเทียบผลการสืบค้นคำศัพท์

คนที่	การเทคนิคการค้นหา ร่วมกับการพยากรณ์คำศัพท์					การเพิ่มโมดูล ตามแนวทางออนโทโลยี				
	ข้อที่ 1	ข้อที่ 2	ข้อที่ 3	ข้อที่ 4	ข้อที่ 5	ข้อที่ 1	ข้อที่ 2	ข้อที่ 3	ข้อที่ 4	ข้อที่ 5
1	3	4	1	2	4	4	5	1	2	5
2	4	5	1	1	3	5	3	1	1	5
3	4	3	0	2	3	4	4	0	2	3
4	4	3	1	2	3	4	4	2	2	3
5	4	4	0	1	4	4	4	1	1	5
...										
169	3	3	0	1	2	2	3	1	2	4

การประเมินให้คะแนนด้วยเกณฑ์ 3 ระดับ Point-Score Method ระหว่างผู้เชี่ยวชาญและระบบที่พัฒนาขึ้น โดยใช้วิธีเลือกกลุ่มตัวอย่างผู้ทำการสอบจำนวน 5 คนในลักษณะที่ต่างกัน

ตารางที่ 4.10 การเปรียบเทียบการประเมินคะแนนคำถามข้อที่ 1 ระหว่างผู้เชี่ยวชาญและระบบ

ข้อที่ 1. คำถาม : จงยกตัวอย่างอุปกรณ์นำเข้า (อย่างน้อย 4 อุปกรณ์)

คำตอบของผู้สอบ	คะแนนผู้เชี่ยวชาญ					ค่าเฉลี่ย	คะแนนจากระบบ	
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		ร้อยละ	เกณฑ์
คนที่ 1 : จอภาพกับ คีย์บอร์ด	0	0	0	0	0	0.0	16	0
คนที่ 2 : เมาส์ คีย์บอร์ด สแกนเนอร์และ ไมโครโฟน	2	2	2	2	2	2.0	80	2
คนที่ 3 : เมาส์ คีย์บอร์ด กล้อง ไมโครโฟน	0	1	1	1	1	0.8	60	1
คนที่ 4 : ไมโครโฟน คีย์บอร์ด เมาส์	2	1	2	1	1	1.4	60	1
คนที่ 5 : คีย์บอร์ด เมาส์	1	1	1	1	1	1.0	40	1

ตารางที่ 4.11 การเปรียบเทียบการประเมินคะแนนคำถามข้อที่ 2 ระหว่างผู้เชี่ยวชาญและระบบ

ข้อที่ 2. คำถาม : องค์ประกอบของระบบคอมพิวเตอร์มีอะไรบ้าง

คำตอบของผู้สอบ	คะแนนจากผู้เชี่ยวชาญ					ค่าเฉลี่ย	คะแนนจากระบบ	
	คนที่	คนที่	คนที่	คนที่	คนที่		ร้อยละ	เกณฑ์
	1	2	3	4	5			
คนที่ 1 : บุคลากร ขั้นตอน ฮาร์ดแวร์ ซอฟต์แวร์และข้อมูล	2	2	2	2	2	2.0	100	2
คนที่ 2 : ข้อมูล วิธีการ คน ฮาร์ดแวร์ ซอฟต์แวร์	2	2	2	2	2	2.0	100	2
คนที่ 3 : ซอฟต์แวร์และ ฮาร์ดแวร์	0	0	1	1	1	0.6	40	1
คนที่ 4 : ระบบสื่อสาร และคอมพิวเตอร์	0	0	0	0	0	0.0	0	0
คนที่ 5 : ซอฟต์แวร์ คน และคอมพิวเตอร์	1	0	0	1	1	0.6	40	1

ตารางที่ 4.12 การเปรียบเทียบการประเมินคะแนนคำถามข้อที่ 3 ระหว่างผู้เชี่ยวชาญและระบบ

ข้อที่ 3. คำถาม : หน่วยประมวลผลกลางคืออะไร

คำตอบของผู้สอบ	คะแนนจากผู้เชี่ยวชาญ					ค่าเฉลี่ย	คะแนนจากระบบ	
	คนที่	คนที่	คนที่	คนที่	คนที่		ร้อยละ	เกณฑ์
	1	2	3	4	5			
คนที่ 1 : ซีพียู	2	2	2	2	2	2.0	100	2
*คนที่ 2 :	0	0	0	0	0	0.0	0	0
คนที่ 3 : ศูนย์กลางการ ประมวลผล	0	0	1	0	0	0.2	0	0
คนที่ 4 : ซีพียู, สมอง ของคอมพิวเตอร์	2	2	2	2	2	2.0	100	2
*คนที่ 5 : หน่วยหลักใช้ สั่งการทำงานประมวลผล ของคอมพิวเตอร์	2	1	0	0	1	0.8	0	0

กรณีคนที่ 2 ไม่ตอบคำถาม, คนที่ 5 ผู้เชี่ยวชาญมีความคิดเห็นว่าเป็นคำตอบที่ถูกต้องเช่นกัน และบางคนเห็นว่าเป็นการตอบโดยล้าจากคำถาม

ตารางที่ 4.13 การเปรียบเทียบการประเมินคะแนนคำถามข้อที่ 4 ระหว่างผู้เชี่ยวชาญและระบบ

ข้อที่ 4. คำถาม : ประเภทของซอฟต์แวร์มีอะไรบ้าง

คำตอบของผู้สอบ	คะแนนจากผู้เชี่ยวชาญ					ค่าเฉลี่ย	คะแนนจากระบบ	
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		ร้อยละ	เกณฑ์
คนที่ 1 : ระบบปฏิบัติการและ โปรแกรมประยุกต์	2	1	1	2	2	1.6	100	2
คนที่ 2 : โปรแกรม ระบบและโปรแกรม ทั่วไป	1	1	1	2	1	1.2	50	1
* คนที่ 3 : ระบบ และประยุกต์	2	1	1	1	2	0.0	0	0
คนที่ 4 : วินโดว์ ลิ นุก และไอโอเอส	0	0	0	0	0	0.0	0	0
คนที่ 5 : ซอฟต์แวร์ ระบบกับซอฟต์แวร์ ประยุกต์	2	2	2	2	2	2.0	100	2

กรณีคำตอบคนที่ 3 ผู้เชี่ยวชาญมีความคิดเห็นว่าเป็นคำตอบที่ถูกต้องแต่เขียนไม่ชัดเจน และระบบไม่พิจารณาให้คะแนน ทั้งนี้ผู้เชี่ยวชาญแนะนำให้ดำเนินการแก้ไขโดยเพิ่มคำสำคัญลงฐานข้อมูลหรือเพิ่มโหนดเชื่อมภายในโครงสร้างให้เหมาะสม

ตารางที่ 4.14 การเปรียบเทียบการประเมินคะแนนคำถามข้อที่ 5 ระหว่างผู้เชี่ยวชาญและระบบ

ข้อที่ 5. คำถาม : ระบบปฏิบัติการคืออะไร

คำตอบของผู้สอบ	คะแนนจากผู้เชี่ยวชาญ					ค่าเฉลี่ย	คะแนนจากระบบ	
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		ร้อยละ	เกณฑ์
คนที่ 1 : โปรแกรมที่ใช้สำหรับควบคุมและตรวจสอบอุปกรณ์ก่อนใช้งาน	2	2	2	2	2	2.0	66	2
คนที่ 2 : ซอฟต์แวร์ชนิดหนึ่ง	0	0	0	0	1	0.2	10	0
คนที่ 3 : ซอฟต์แวร์ที่ใช้บริหารจัดการคอมพิวเตอร์	2	2	2	2	2	2.0	50	1
*คนที่ 4 : อุปกรณ์ในการควบคุมบริหารจัดการคอมพิวเตอร์	0	0	0	0	0	0.0	66	2
คนที่ 5 : เป็นระบบซอฟต์แวร์ที่ทำหน้าที่จัดการอุปกรณ์คอมพิวเตอร์และแหล่งซอฟต์แวร์และบริหารโปรแกรมคอมพิวเตอร์	2	2	2	2	2	2.0	66	2

กรณีคำตอบคนที่ 4 ผู้เชี่ยวชาญมีความคิดเห็นว่าผู้สอบอาจมีความรู้ที่ผิดหรืออาจตั้งใจเขียนสลับตำแหน่งประโยค จึงไม่พิจารณาให้คะแนน ทั้งนี้ผู้เชี่ยวชาญแนะนำให้ควรมีระบบตรวจสอบจัดลำดับคำสำคัญ

การเปลี่ยนแปลงค่าให้อยู่ในรูปแบบ เกณฑ์การประเมิน (Rubric Assessment) เพื่อใช้ในเกณฑ์การประเมิน 3 ระดับ

ตารางที่ 4.15 การประเมินให้คะแนนด้วยเกณฑ์ 3 ระดับ Point-Score Method

คนที่	ค่าเฉลี่ยคะแนนแต่ละข้อจากผู้เชี่ยวชาญ					คะแนนแต่ละข้อที่ระบบฯ				
	ข้อที่ 1	ข้อที่ 2	ข้อที่ 3	ข้อที่ 4	ข้อที่ 5	ข้อที่ 1	ข้อที่ 2	ข้อที่ 3	ข้อที่ 4	ข้อที่ 5
1	0.00	2.00	2.00	1.60	2.00	0.00	2.00	2.00	2.00	2.00
2	2.00	2.00	0.00	1.20	0.20	2.00	2.00	0.00	1.00	0.00
3	0.80	0.60	0.20	0.00	2.00	1.00	1.00	0.00	0.00	1.00
4	1.40	0.00	2.00	0.00	0.00	1.00	0.00	2.00	0.00	2.00
5	1.00	0.60	0.80	2.00	2.00	1.00	1.00	0.00	2.00	2.00
...										
25	2.00	2.00	0.80	1.60	0.00	2.00	2.00	1.00	2.00	0.00

ตารางที่ 4.16 ผลการเปรียบเทียบ t-test การให้คะแนนข้อสอบ 5 ข้อข้อละ 25 ฉบับ

ข้อที่	t Stat	t Critical one-tail	P(T<=t) one-tail	Pearson Correlation
1	1.41	1.71	0.09	0.54
2	0.90	1.71	0.19	0.54
3	0.70	1.71	0.25	0.55
4	1.07	1.71	0.15	0.40
5	1.44	1.71	0.08	0.74

สมมติฐานหลัก (H0): คะแนนจากระบบคล้ายกับคะแนนจากผู้เชี่ยวชาญ เมื่อเปรียบเทียบค่า t-test ของผลการตรวจคะแนนทั้ง 5 ข้อ พบว่า มีค่าแปรผลไม่แตกต่างกัน ค่า t stat น้อยกว่าหรือเท่ากับ t critical one-tail อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

ตารางที่ 4.17 ผลการวิเคราะห์ผลด้วย T-Test (t-Test: Paired Two Sample for Means)

	คะแนนจากระบบ	ค่าเฉลี่ยจากผู้เชี่ยวชาญ
Mean	1.08	1.056
Variance	0.74	0.71
Observations	25.0	25.0
Pearson Correlation	0.81	
Hypothesized Mean Difference	0.00	
df	24.0	
t Stat	0.23	
P(T<=t) one-tail	0.40	
t Critical one-tail	1.71	
P(T<=t) two-tail	0.81	
t Critical two-tail	2.06	

สมมติฐานหลัก (H0): ค่าเฉลี่ยคะแนนจากระบบคล้ายกับค่าเฉลี่ยคะแนนจากผู้เชี่ยวชาญ เมื่อเปรียบเทียบมีค่าแปรผลไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05

4.2 ผลการประเมินประสิทธิภาพโดยผู้เชี่ยวชาญ

ตารางที่ 4.18 ผลการวิเคราะห์หาความพึงพอใจของระบบ

ข้อความคำถามของแบบประเมินคุณภาพ	\bar{x}	S.D.
1.ด้านกระบวนการ/ขั้นตอนการใช้งานระบบ		
1.1. ทำงานตามอัลกอริทึมอย่างถูกต้อง	4.60	0.49
1.2. มีลำดับทำงานแต่ละโปรแกรมครบถ้วน	4.00	0.63
ผลรวม	4.30	0.56
2.ด้านประสิทธิภาพของระบบ		
2.1. ความถูกต้องแม่นยำของระบบ	3.80	0.75
2.2. ตรงตามวัตถุประสงค์ที่ต้องการ	4.60	0.49
2.3. รายงานข้อมูลเป็นปัจจุบัน	4.20	0.75
2.4. ระบบมีความน่าเชื่อถือ	4.80	0.40
ค่าเฉลี่ยผลรวม	4.35	0.60
3.ด้านความสะดวก สบายงาม		
3.1. ให้ความสนใจกับระบบได้ง่าย	4.80	0.40
3.2. โปรแกรมมีความเหมาะสมน่าใช้	4.60	0.49
ค่าเฉลี่ยผลรวม	4.70	0.45
4.ด้านคุณภาพของระบบ		
4.1. รายงานมีความถูกต้อง	4.40	0.49
4.2. ความพึงพอใจในผลรายงาน	4.80	0.40
ค่าเฉลี่ยผลรวม	4.60	0.45
ค่าเฉลี่ยรวมทั้งหมด	4.49	0.51

ความพึงพอใจโดยรวมของระบบมีค่า \bar{x} เท่ากับ 4.49 ซึ่งอยู่ในระดับพึงพอใจมาก

4.3 ผลการประเมินความพึงพอใจจากกลุ่มผู้ใช้

ตารางที่ 4.19 การประเมินความพึงพอใจจากกลุ่มผู้ใช้

ข้อความคำถามของแบบประเมินคุณภาพ	\bar{x}	S.D.
1. เชื่อถือได้ของระบบ (Reliable)	4.10	0.82
2. ระบบเข้าใจง่าย (Simple)	4.10	0.70
3. รายงานผลทันต่อเวลา (Timely)	4.36	0.55
4. คำนึงราคาต่อการใช้งาน (Economical)	4.17	0.62
5. ผลรายงานตรวจสอบได้ (Verifiable)	4.35	0.48
6. ระบบมีความยืดหยุ่นสามารถใช้ได้กับอุปกรณ์ที่หลากหลาย (Flexible)	4.21	0.61
7. สอดคล้องกับความต้องการขององค์กร (Relevant)	4.23	0.58
8. ระบบมีสะดวกในการเข้าถึงและใช้งาน (Accessible)	4.15	0.55
9. ระบบมีความปลอดภัย (Secure)	4.19	0.38
ค่าเฉลี่ยรวมทั้งหมด	4.21	0.70

ผู้มีความพึงพอใจในการใช้งานระบบตรวจสอบอัตโนมัติภาษาไทยอัตโนมัติ อยู่ในระดับพึงพอใจมาก โดยมีค่า \bar{x} เท่ากับ 4.21