

บทที่ 4

ผลการวิเคราะห์ข้อมูล

4.1 บทนำ

ในบทนี้จะแสดงผลและวิเคราะห์เปรียบเทียบทดสอบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ด้วยตัวแบบทั้ง 2 ตัวแบบ เพื่อนำมาตอบวัตถุประสงค์ของงานวิจัยที่กล่าวไปในบทที่ 3 ซึ่งได้แก่ 1) เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทาง 2) เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบในสถานการณ์การป้องกันการเกิดข้อผิดพลาด และ 3) เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว และในส่วนท้ายของบทนี้เป็นการศึกษาและวิเคราะห์ข้อมูลเพิ่มเติม

4.2 ผลการทดลอง

การทดลองนี้มีวัตถุประสงค์เพื่อวัดประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ของตัวแบบ 2 ตัวแบบในสถานการณ์ของการพัฒนาซอฟต์แวร์ที่ต่างกัน 3 สถานการณ์ โดยใช้ค่าเอฟเมสเซอร์เป็นค่าประเมินประสิทธิภาพของการทำเหมืองข้อมูลบนข้อมูลซอฟต์แวร์อาร์ไคฟ์ในสถานการณ์การนำทาง และสถานการณ์การป้องกันข้อผิดพลาด และใช้ค่าผลสะท้อนกลับเป็นค่าประเมินประสิทธิภาพของการทำเหมืองข้อมูลบนข้อมูลซอฟต์แวร์อาร์ไคฟ์ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว

ตารางแสดงชุดทดสอบจำนวน 60 ตารางแสดงชั้นที่เป็นตัวแทนของตารางแสดงชั้นการเปลี่ยนแปลงแก้ไขจากโครงการพัฒนาซอฟต์แวร์ทางการบัญชีชื่อเคมายมันนี่ (KMyMoney) ที่เลือกมาตามขั้นตอนย่อยภายในขั้นตอนการสร้างข้อสอบถามหัวข้อที่ 3.6.2 และภาคผนวก ก สามารถนำมาสร้างเป็นข้อสอบถามได้ทั้งหมด 962 ข้อสอบถาม ประกอบด้วย ข้อสอบถามสำหรับทดสอบในสถานการณ์การนำทางทั้งหมด 451 ข้อสอบถาม ข้อสอบถามสำหรับทดสอบใน

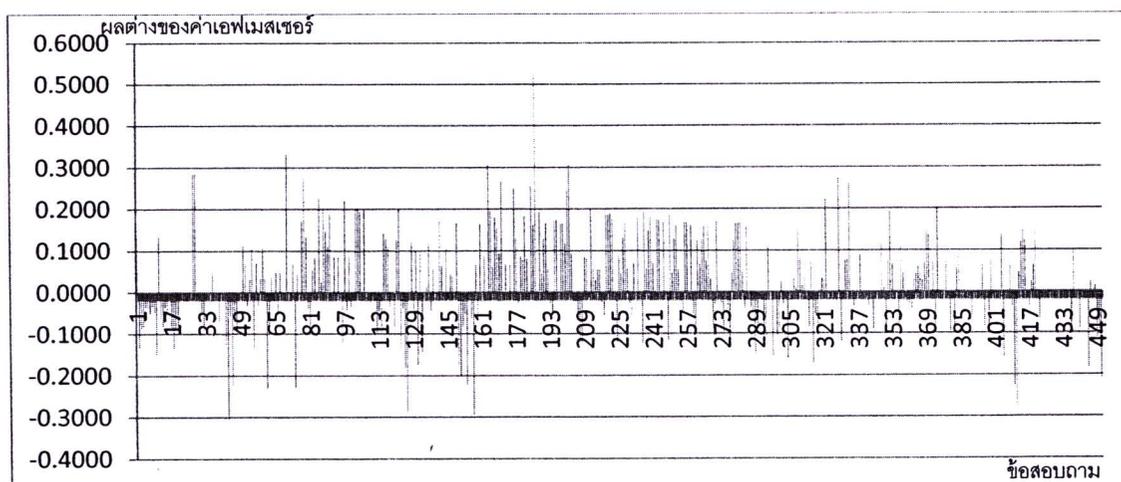
สถานการณ์การป้องกันการเกิดข้อผิดพลาดทั้งหมด 451 ข้อสอบถาม และข้อสอบถามสำหรับทดสอบในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้วทั้งหมด 60 ข้อสอบถาม ค่าเอฟเมสเซอร์ที่ได้จากการทดสอบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบสำหรับสถานการณ์การนำทางทั้งหมด 451 ค่าและสำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาดทั้งหมด 451 ค่าแสดงในตารางที่ ข-1 และตารางที่ ข-2 ตามลำดับ ค่าผลสะท้อนกลับที่ได้จากการทดสอบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบสำหรับสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้วทั้งหมด 60 ค่าแสดงในตารางที่ ข-3

จากการทดสอบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ด้วยตัวแบบทั้ง 2 ตัวแบบคือ การค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 และ การค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 นั้น ในสถานการณ์ของการพัฒนาซอฟต์แวร์ที่ต่างกัน 3 สถานการณ์คือ สถานการณ์การนำทาง สถานการณ์การป้องกันการเกิดข้อผิดพลาด และสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว สามารถสรุปได้ดังนี้

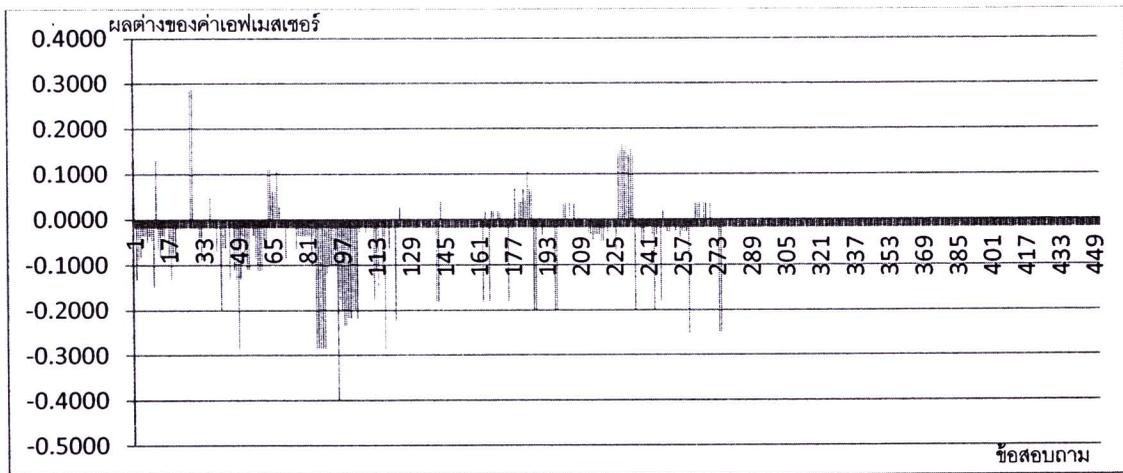
- การทดสอบสถานการณ์การนำทาง
 - ข้อสอบถามที่ได้เขตของคำแนะนำที่ให้ค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 มากกว่าค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 มีจำนวน 192 ข้อสอบถาม
 - ข้อสอบถามที่ได้เขตของคำแนะนำที่ให้ค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 เท่ากับค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 มีจำนวน 111 ข้อสอบถาม
 - ข้อสอบถามที่ได้เขตของคำแนะนำที่ให้ค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 น้อยกว่าค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 มีจำนวน 148 ข้อสอบถาม
- การทดสอบสถานการณ์การป้องกันการเกิดข้อผิดพลาด
 - ข้อสอบถามที่ได้เขตของคำแนะนำที่ให้ค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 มากกว่าค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 มีจำนวน 46 ข้อสอบถาม

- ข้อสอบถามที่ได้เขตของคำแนะนำที่ให้ค่าประสิทธิภาพของการค้นหาจากความสัมพันธ์ด้วยตัวแบบที่ 2 เท่ากับค่าประสิทธิภาพของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 มีจำนวน 279 ข้อสอบถาม
- ข้อสอบถามที่ได้เขตของคำแนะนำที่ให้ค่าประสิทธิภาพของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 น้อยกว่าค่าประสิทธิภาพของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 มีจำนวน 126 ข้อสอบถาม
- การทดสอบสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว
 - ข้อสอบถามที่ได้เขตของคำแนะนำเป็นเขตว่าง มีจำนวน 19 ข้อสอบถาม
 - ข้อสอบถามที่ได้เขตของคำแนะนำไม่เป็นเขตว่าง มีจำนวน 41 ข้อสอบถาม

จากผลการทดสอบค่าเอฟเมสเซอร์ของสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ในตารางที่ ข-1 และ ข-2 ตามลำดับ ผู้วิจัยสามารถแสดงกราฟเพื่อเปรียบเทียบประสิทธิภาพของการค้นหาความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์โคพีด้วยตัวแบบทั้ง 2 ตัวแบบโดยใช้กราฟแท่ง (Column Chart) ที่แสดงผลต่างค่าเอฟเมสเซอร์ของการค้นหาความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์โคพีด้วยตัวแบบทั้ง 2 ตัวแบบของสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดดังรูปที่ 4.1 และ 4.2 ตามลำดับ และตารางที่ 4-1 แสดงตารางสรุปผลการทดสอบของทั้ง 3 สถานการณ์



รูปที่ 4-1 แสดงกราฟผลต่างค่าเอฟเมสเซอร์ของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 กับการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง



รูปที่ 4-2 แสดงกราฟผลต่างค่าเอฟแมสเซอร์ของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 กับการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

ตารางที่ 4-1 แสดงตารางผลการทดสอบทั้ง 3 สถานการณ์

สถานการณ์การนำทาง (ค่าเฉลี่ยของค่าเอฟแมสเซอร์)		สถานการณ์การป้องกันการเกิดข้อผิดพลาด (ค่าเฉลี่ยของค่าเอฟแมสเซอร์)		สถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว (ค่าผลสะท้อนกลับ)	
ตัวแบบที่ 1	ตัวแบบที่ 2	ตัวแบบที่ 1	ตัวแบบที่ 2	ตัวแบบที่ 1	ตัวแบบที่ 2
0.3013	0.3245	0.1353	0.1135	$\frac{41}{60}$	$\frac{41}{60}$

จากตารางสรุปข้างต้นแสดงให้เห็นว่าค่าประสิทธิภาพของสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้นมีค่าแตกต่างกัน ค่าประสิทธิภาพของสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้วนั้นไม่มีความแตกต่างกัน แต่ไม่สามารถสรุปได้ว่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ในสถานการณ์ต่างๆนั้นแตกต่างกันอย่างมีนัยสำคัญ ผู้วิจัยจึงจำเป็นต้องวิเคราะห์ผลการทดสอบความแตกต่างกันอย่างมีนัยสำคัญ ดังรายละเอียดในหัวข้อต่อไป

4.3 ผลการวิเคราะห์ข้อมูล

สำหรับสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด การตรวจสอบเงื่อนไขพื้นฐานขั้นแรก ผู้วิจัยต้องตรวจสอบการแจ่มแจ้งของประชากรว่าการแจ่มแจ้ง

ปกติหรือไม่ เพื่อเลือกทางเลือกในการทดสอบสมมติฐานว่าจะใช้วิธีการทดสอบสมมติฐานแบบอิงพารามิเตอร์ (Parametric Test) หรือแบบไม่อิงพารามิเตอร์ (Non Parametric Test) ถ้าผลการทดสอบพบว่าประชากรมีการแจกแจงแบบปกติ จึงใช้การวิเคราะห์โดยสถิติทดสอบที (t-test) เพื่อทดสอบสมมติฐานของผลต่างระหว่างค่าเฉลี่ยของค่าเอฟเมสเซอร์ของหน่วยทดลอง 2 กลุ่ม ถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05 และค่าสถิติที่มากกว่า 0 จึงจะสามารถปฏิเสธ H_0 ได้ แต่ถ้าผลการแจกแจงประชากรพบว่าการแจกแจงไม่ปกติ จึงใช้การวิเคราะห์โดยสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซันสำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) เพื่อทดสอบสมมติฐานของผลต่างระหว่างค่าเฉลี่ยของค่าเอฟเมสเซอร์ของหน่วยทดลอง 2 กลุ่ม ถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05 และค่าสถิติที่มากกว่า 0 ในกรณีที่ผลการวิเคราะห์ตั้งอยู่บนพื้นฐานทางบวก (Based on positive ranks) จึงจะสามารถปฏิเสธ H_0 ได้

4.3.1 การวิเคราะห์การแจกแจงข้อมูล

ในงานวิจัยนี้ผู้วิจัยสนใจตัวแปร คือประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ของตัวแบบ 2 ตัวแบบ ดังนั้นจึงตรวจสอบการแจกแจงของข้อมูลที่ได้จากหน่วยทดลอง นั่นคือค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ ซึ่งได้แก่ค่าเอฟเมสเซอร์ที่ได้มาจากผลการทดสอบ 451 ข้อสอบถาม สำหรับสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ดังนั้นผู้วิจัยจะตรวจสอบว่าค่าเอฟเมสเซอร์ทั้งหมด 451 ค่าของทั้ง 2 สถานการณ์มีการแจกแจงแบบปกติหรือไม่ โดยตั้งสมมติฐานของการทดสอบค่าประสิทธิภาพแต่ละสถานการณ์มีการแจกแจงแบบปกติหรือไม่ภายใต้สมมติฐานทางสถิติ ดังนี้

- 1) ทดสอบการแจกแจงของข้อมูลค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 สถานการณ์การนำทาง
 - H_0 : ข้อมูลค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 สถานการณ์การนำทาง มีการแจกแจงแบบปกติ
 - H_1 : ข้อมูลค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 สถานการณ์การนำทาง ไม่แจกแจงแบบปกติ
- 2) ทดสอบการแจกแจงของข้อมูลค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 สถานการณ์การนำทาง

H_0 : ข้อมูลค่าประสิทธิภาพของการค้นหาคำความสัมพันธ์ตัวแบบที่ 2 สถานการณ์การนำทาง มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าประสิทธิภาพของการค้นหาคำความสัมพันธ์ตัวแบบที่ 2 สถานการณ์การนำทาง ไม่แจกแจงแบบปกติ

- 3) ทดสอบการแจกแจงของข้อมูลค่าประสิทธิภาพของการค้นหาคำความสัมพันธ์ตัวแบบที่ 1 สถานการณ์การป้องกันการเกิดข้อผิดพลาด

H_0 : ข้อมูลค่าประสิทธิภาพของการค้นหาคำความสัมพันธ์ตัวแบบที่ 1 สถานการณ์การป้องกันการเกิดข้อผิดพลาด มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าประสิทธิภาพของการค้นหาคำความสัมพันธ์ตัวแบบที่ 1 สถานการณ์การป้องกันการเกิดข้อผิดพลาด ไม่แจกแจงแบบปกติ

- 4) ทดสอบการแจกแจงของข้อมูลค่าประสิทธิภาพของการค้นหาคำความสัมพันธ์ตัวแบบที่ 2 สถานการณ์การป้องกันการเกิดข้อผิดพลาด

H_0 : ข้อมูลค่าประสิทธิภาพของการค้นหาคำความสัมพันธ์ตัวแบบที่ 2 สถานการณ์การป้องกันการเกิดข้อผิดพลาด มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าประสิทธิภาพของการค้นหาคำความสัมพันธ์ตัวแบบที่ 2 สถานการณ์การป้องกันการเกิดข้อผิดพลาด ไม่แจกแจงแบบปกติ

ผู้วิจัยเลือกใช้สถิติทดสอบ Kolmogorov-Sminov เนื่องจากมีขนาดตัวอย่างมากกว่า 50 หน่วย โดยจะยอมรับสมมติฐาน H_0 เมื่อค่า Sig. มีค่ามากกว่าค่า α ซึ่งกำหนดให้เท่ากับ 0.05 ดังตารางต่อไปนี้

ตารางที่ 4-2 แสดงค่าสถิติทดสอบการแจกแจงปกติ (Normality Test) ของค่าประสิทธิภาพของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

สถานการณ์	ตัวแบบที่	Kolmogorov-Smirnov		
		Statistic	df	Sig.
สถานการณ์การนำทาง	1	0.070	451	0.000
	2	0.075	451	0.000
สถานการณ์การป้องกันการเกิดข้อผิดพลาด	1	0.313	451	0.000
	2	0.326	451	0.000

ผลการทดสอบในตารางที่ 4-2 ชี้ให้เห็นพบว่าค่า Sig. ของตัวแปรค่าประสิทธิภาพการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดเป็นดังนี้

- 1) สำหรับสถานการณ์การนำทาง การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 2) สำหรับสถานการณ์การนำทาง การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 3) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 4) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0

ดังนั้นสรุปได้ว่าการแจกแจงของตัวแปรค่าประสิทธิภาพของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบสำหรับ 2 สถานการณ์นั้นไม่เป็นแบบปกติ

4.3.2 การวิเคราะห์เปรียบเทียบประสิทธิภาพการค้นหากฎความสัมพันธ์ทั้ง 2 ตัวแบบ

จากการวิเคราะห์การแจกแจงข้อมูลข้างต้นพบว่าค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดไม่แจกแจงแบบปกติ ดังนั้นผู้วิจัยจึงเลือกใช้สถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) กับการทดสอบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด โดยที่จะปฏิเสธสมมติฐาน H_0 ได้เมื่อถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05 และค่าสถิติ Z มากกว่า 0 โดยที่ผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) และผู้วิจัยเลือกใช้การทดสอบอัตราส่วน 2 กลุ่มด้วยสถิติทดสอบ Z (Two Proportion Z Tests) กับการทดสอบในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว โดยที่จะปฏิเสธสมมติฐาน H_0 ได้เมื่อถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05

การวิเคราะห์เปรียบเทียบนี้เป็นการวิเคราะห์เปรียบเทียบประสิทธิภาพระหว่างการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 กับการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ในสถานการณ์ 3 สถานการณ์คือ สถานการณ์การนำทาง สถานการณ์การป้องกันการเกิดข้อผิดพลาด สถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว ซึ่งสามารถตั้งสมมติฐานได้ดังนี้

- 1) วิเคราะห์เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์โคฟว์ทั้ง 2 ตัวแบบในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 2) วิเคราะห์เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

M_2 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 3) วิเคราะห์เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว

M_2 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

การทดสอบสมมติฐานข้อ 1 และ 2 ด้วยสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซันสำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) แสดงดังตารางที่ 4-3 สำหรับสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว การทดสอบสมมติฐานข้อ 3 ด้วยสถิติทดสอบ Z (Two Proportion Z Tests) แสดงดังตารางที่ 4-4

ตารางที่ 4-3 แสดงค่าสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ของค่าประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ตัวแบบที่ 2 เทียบกับตัวแบบที่ 1 ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

	ค่าประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ตัวแบบที่ 2 - ค่าประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ตัวแบบที่ 1	
	สถานการณ์การนำทาง	สถานการณ์การป้องกันการเกิดข้อผิดพลาด
Z	-4.374 ^a	-6.055 ^b
Asymp. Sig. (2 tailed)	.000	.000

a. Based on negative ranks.

b. Based on positive ranks.

จากตารางที่ 4-3 การวิเคราะห์เปรียบเทียบประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทาง ได้สถิติทดสอบค่า Z เท่ากับ -4.374 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ เนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางลบ (Based on negative ranks) ดังนั้นจึงสามารถปฏิเสธสมมติฐาน H_0 ได้ นั่นคือค่าประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ด้วยตัวแบบที่ 2 มากกว่าค่าประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05 และสำหรับการวิเคราะห์เปรียบเทียบประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ทั้ง 2 ตัวแบบในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ได้สถิติทดสอบค่า Z เท่ากับ -6.055 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ เนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติฐาน H_0 ได้ นั่นคือค่าประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าค่าประสิทธิภาพของการค้นหาภูมิตามสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

ตารางที่ 4-4 แสดงการทดสอบอัตราส่วน 2 กลุ่มด้วยสถิติทดสอบ Z ของค่าประสิทธิภาพของการค้นหาความสัมพันธ์ตัวแบบที่ 2 เทียบกับตัวแบบที่ 1 ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.000 ^a	1	1.000		
Continuity Correction ^b	.000	1	1.000		
Likelihood Ratio	.000	1	1.000		
Fisher's Exact Test				1.000	.578
N of Valid Cases	120				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 19.00.

จากตารางที่ 4.4 ค่าสถิติ Z เท่ากับรากที่สองของค่า Pearson Chi-Square นั้นค่า 0.000 และมีเครื่องหมายเดียวกับผลต่างของอัตราส่วนที่ 2 กับอัตราส่วนที่ 1 ($\frac{41}{60} - \frac{41}{60} = 0.000$) นั้นค่าเครื่องหมายบวก และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $1.000 / 2 = 0.500$ ซึ่งมากกว่าค่า $\alpha = 0.05$ ดังนั้นจึงไม่สามารถปฏิเสธสมมติฐาน H_0 ได้ นั่นคือค่าประสิทธิภาพของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ไม่ต่างกับค่าประสิทธิภาพของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้วที่ระดับนัยสำคัญ 0.05

4.3.3 สรุปผลการวิเคราะห์ข้อมูล

จากการวิเคราะห์ผลการทดสอบประสิทธิภาพของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด โดยใช้ค่าเอฟเมสเซอร์ และสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว โดยใช้ค่าผลสะท้อนกลับ ผู้วิจัยสามารถสรุปได้ว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีประสิทธิภาพมากกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีประสิทธิภาพไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด และสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว

การค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 มีประสิทธิภาพไม่ต่างกับการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1

จากผลสรุปดังกล่าวแสดงให้เห็นว่า การใช้ตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นใหม่ในการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์บนข้อมูลซอฟต์แวร์อาร์โคฟว์แสดงประสิทธิภาพที่ดีกว่าการใช้ตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นเพียงแคในสถานการณ์การนำทางเท่านั้น

4.4 ผลการศึกษาเพิ่มเติม

จากผลทดสอบการเปรียบเทียบประสิทธิภาพการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์โคฟว์ด้วยตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นและตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นใหม่ข้างต้น ผู้วิจัยมีความต้องการทดสอบเพิ่มเติมเพื่ออธิบายสาเหตุที่ทำให้การทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์โคฟว์ด้วยตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นใหม่แสดงประสิทธิภาพที่ดีกว่าเฉพาะในสถานการณ์ของการนำทาง แต่ให้ประสิทธิภาพที่ไม่ต่างกันหรือน้อยกว่าในสถานการณ์การป้องกันการเกิดข้อผิดพลาดและสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว จากการดำเนินการทดสอบข้างต้นผู้วิจัยมีข้อสังเกตหลายประการ เช่น การค้นหากฎความสัมพันธ์ด้วยตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นใหม่นั้นมักจะให้เซตของคำแนะนำที่ใหญ่กว่าการค้นหากฎความสัมพันธ์ด้วยตัวแบบค่าสับสนุน-ค่าความเชื่อมั่น เนื่องจากกฎความสัมพันธ์ 10 อันดับแรกที่ได้จากการค้นหากฎความสัมพันธ์ด้วยตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นนั้นมักจะเป็นกฎความสัมพันธ์ที่มีเซตรายการที่ตามมาขนาด 1 รายการเท่านั้น ในขณะที่กฎความสัมพันธ์ 10 อันดับแรกที่ได้จากการค้นหากฎความสัมพันธ์ด้วยตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นใหม่มักจะเป็นกฎความสัมพันธ์ที่มีเซตรายการที่ตามมาขนาดมากกว่า 1 รายการ เป็นต้น จากข้อสังเกตดังกล่าวผู้วิจัยคิดว่าเซตของคำแนะนำที่มีขนาดใหญ่อาจเป็นสาเหตุที่ทำให้ค่าความถูกต้อง (Precision) ของการใช้ตัวแบบค่าสับสนุน-ค่าความเชื่อมั่นใหม่มีค่าน้อยและส่งผลให้ค่าเอฟเมเชอร์มีค่าน้อยตามมาด้วย

เนื่องจากขั้นตอนวิธีในการทดสอบของงานวิจัยมีข้อแตกต่างจากขั้นตอนวิธีที่ Zimmermann และคณะกับ Methanias และคณะ (Zimmermann et al., 2005; Methanias et

al., 2009) ใช้อยู่ขั้นตอนหนึ่ง คือนักวิจัยทั้ง 2 คนจะเน้นมีการปรับปรุงขั้นตอนวิธีหรือปริโอรี 2 ประการ เพื่อให้สามารถค้นหากฎความสัมพันธ์ที่รวดเร็วขึ้นตามที่อธิบายไว้ในบทที่ 2 คือ 1) การกำหนดให้ค้นหาความสัมพันธ์เฉพาะกฎที่มีเซตรายการที่มาก่อนที่ต้องการเท่านั้น และ 2) การกำหนดให้ทุกกฎความสัมพันธ์ที่ค้นหาจะมีเซตรายการที่ตามมาเพียง 1 รายการเท่านั้น จากข้อกำหนดข้อที่ 2 นั้นทำให้เซตของคำแนะนำที่จะได้มานั้นจะมีขนาดน้อยกว่าหรือเท่ากับ 10 รายการเสมอ แต่ในงานวิจัยนี้ผู้วิจัยไม่ได้ปรับปรุงขั้นตอนวิธีหรือปริโอรีด้วยข้อกำหนดดังกล่าว ขนาดของเซตของคำแนะนำจึงมีความหลากหลายแตกต่างกันออกไป โดยที่เซตของคำแนะนำที่ได้จากการค้นหาความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นนั้นจะมีขนาดน้อยกว่าหรือเท่ากับ 10 รายการเสมอเนื่องจากกฎความสัมพันธ์ 10 อันดับแรกมักจะมีเซตรายการที่ตามมาขนาด 1 รายการหรือเซตรายการที่ตามมาขนาดมากกว่า 1 รายการแต่มีสมาชิกที่ซ้ำกับเซตรายการที่ตามมาของกฎความสัมพันธ์ที่อยู่อันดับที่สูงกว่า แต่สำหรับเซตของคำแนะนำที่ได้จากการค้นหาความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่นั้นมักมีขนาดมากกว่า 10 เสมอเนื่องจากเนื่องจากกฎความสัมพันธ์ 10 อันดับแรกมักจะมีเซตรายการที่ตามมาขนาดมากกว่า 1 รายการและสมาชิกของเซตรายการที่ตามมาของแต่ละกฎความสัมพันธ์ใน 10 อันดับแรกมักจะไม่ซ้ำกันด้วย

ด้วยเหตุนี้ผู้วิจัยจึงมีความต้องการที่จะทดสอบเพิ่มเติมเพื่ออธิบายสาเหตุของผลการทดสอบข้างต้น โดยที่การทดสอบเพิ่มเติมนี้ผู้วิจัยทำการปรับปรุงขั้นตอนวิธีในการสร้างเซตของคำแนะนำใหม่ จากคุณลักษณะของกฎความสัมพันธ์ที่อยู่ใน 10 อันดับของการค้นหาความสัมพันธ์ด้วยตัวแบบทั้ง 2 นั้น ผู้วิจัยเห็นว่าไม่สามารถนำเซตรายการที่ตามมาของกฎความสัมพันธ์ทั้ง 10 อันดับแรกมาสร้างเป็นเซตของคำแนะนำได้โดยตรง และผู้วิจัยก็เห็นว่าไม่สามารถกำหนดให้พิจารณาเฉพาะกฎความสัมพันธ์ที่ค้นหาจะมีเซตรายการที่ตามมา 1 รายการเท่านั้นเช่นเดียวกับ Zimmermann และคณะกับ Methanias และคณะ (Zimmermann et al., 2005; Methanias et al., 2009) ด้วย ผู้วิจัยเห็นว่าการสร้างเซตของคำแนะนำที่ได้มาจากการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบเพื่อวัตถุประสงค์ในการทดสอบเปรียบเทียบควรจะต้องสร้างเซตของคำแนะนำมาจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของกฎความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ตัวอย่างเช่น ตารางด้านล่างนี้เป็นตารางแสดงกฎความสัมพันธ์ 10 อันดับแรกที่ได้จากการค้นหาความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่เรียงลำดับตามค่าความเชื่อมั่นใหม่ของกฎความสัมพันธ์ที่มีเซตรายการที่มาก่อนคือ {15}

ตารางที่ 4-5 แสดงตัวอย่างการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของกฎความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก

	กฎความสัมพันธ์	ค่านับสนับสนุน	ค่าความเชื่อมั่นใหม่
1	15 -> 29, 189	25	0.52
2	15 -> 365, 520, 521	16	0.51
3	15 -> 798	42	0.48
4	15 -> 646, 798, 799	16	0.46
5	15 -> 179, 465, 466	15	0.45
6	15 -> 646, 798, 804	15	0.45
7	15 -> 29, 678	22	0.43
8	15 -> 798, 799	22	0.43
9	15 -> 189	39	0.43
10	15 -> 225	56	0.41
เซตของคำแนะนำ คือ {29, 179, 189, 365, 465, 520, 521, 646, 798, 799 }			

ในตารางข้างต้นแสดงตัวอย่างของกฎความสัมพันธ์ ค่านับสนับสนุน และค่าความเชื่อมั่นใหม่ โดยในหลัก (column) ของกฎความสัมพันธ์นั้นใช้ตัวเลขต่างๆ แสดงแทนการเปลี่ยนแปลงแก้ไขต่างๆ เพื่อความสะดวกในการทำความเข้าใจ

ดังนั้นผู้วิจัยจึงมีความต้องการทำการทดสอบเพิ่มเติม โดยเริ่มจาก การเปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหา กฎความสัมพันธ์ของทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ต่อด้วยการเปรียบเทียบค่าประสิทธิภาพ ค่าความถูกต้อง และค่าเรียกคืน ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์ของทั้ง 2 ตัวแบบโดยสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของกฎความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ตามลำดับ เพื่อตอบข้อสงสัยที่ว่าค่าความ



ถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่มีค่าน้อยกว่า แต่ค่าเรียกคืนมีค่ามากกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่น

4.4.1 การเปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์ของทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

ผลการทดสอบประสิทธิภาพการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์ของทั้ง 2 ตัวแบบในการทดสอบข้างต้นนั้นเป็นการเปรียบเทียบประสิทธิภาพจากการใช้ค่าเอฟเมเชอร์ (F-measure) เท่านั้น ผู้วิจัยจึงมีความต้องการศึกษาเพิ่มเติมโดยการทดสอบเปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์ของทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-6 แสดงตารางค่าความถูกต้องและค่าเรียกคืนของการทดสอบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

สถานการณ์การนำทาง				สถานการณ์การป้องกันการเกิดข้อผิดพลาด			
ตัวแบบที่ 1		ตัวแบบที่ 2		ตัวแบบที่ 1		ตัวแบบที่ 2	
P_p	R_p	P_p	R_p	P_p	R_p	P_p	R_p
0.4750	0.2955	0.4554	0.3530	0.4758	0.5188	0.4620	0.4922

จากตารางสรุปข้างต้นแสดงให้เห็นว่าค่าเฉลี่ยของค่าความถูกต้อง (Precision) หรือ P_p และค่าเฉลี่ยของค่าเรียกคืน (Recall) หรือ R_p ของการทดสอบข้างต้นนี้มีค่าแตกต่างกันทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด แต่ไม่สามารถสรุปได้ว่าค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาความสัมพันธ์ของตัวแบบทั้ง 2 ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้นแตกต่างกันอย่างมีนัยสำคัญ ผู้วิจัยจึงจำเป็นต้องวิเคราะห์ผลการทดสอบความแตกต่างกันอย่างมีนัยสำคัญ ดังรายละเอียดต่อไปนี้

การวิเคราะห์การแจกแจงปกติ

เนื่องจากผู้วิจัยต้องการเปรียบเทียบค่าความถูกต้อง (Precision) หรือค่าเรียกคืน (Recall) ของการค้นหากฎความสัมพันธ์ของ 2 ตัวแบบสำหรับสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ดังนั้นผู้วิจัยจะตรวจสอบว่าค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ทั้งหมดของทั้ง 2 สถานการณ์มีการแจกแจงแบบปกติหรือไม่ โดยตั้งสมมติฐานของการทดสอบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) แต่ละสถานการณ์มีการแจกแจงแบบปกติหรือไม่ภายใต้สมมติฐานทางสถิติ ดังนี้

- 1) ทดสอบการแจกแจงของข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 ในสถานการณ์การนำทาง
 - H_0 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ
 - H_1 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 ในสถานการณ์การนำทาง ไม่แจกแจงแบบปกติ
- 2) ทดสอบการแจกแจงของข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 ในสถานการณ์การนำทาง
 - H_0 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ
 - H_1 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 ในสถานการณ์การนำทาง ไม่แจกแจงแบบปกติ
- 3) ทดสอบการแจกแจงของข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด
 - H_0 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด มีการแจกแจงแบบปกติ
 - H_1 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ไม่แจกแจงแบบปกติ
- 4) ทดสอบการแจกแจงของข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

ในการตรวจสอบการแจกแจงของข้อมูลว่าเป็นแบบปกติโดยใช้สถิติทดสอบนั้น มีสถิติทดสอบที่ใช้คือ Kolmogorov-Smirnov สำหรับหน่วยทดลองมากกว่า 50 หน่วย และ Shapiro-Wilk สำหรับหน่วยทดลองน้อยกว่า 50 หน่วย โดยจะยอมรับสมมติฐาน H_0 เมื่อค่า Sig. มีค่ามากกว่าค่า α ซึ่งกำหนดให้เท่ากับ 0.05 ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-7 แสดงค่าสถิติทดสอบการแจกแจงปกติ (Normality Test) ของค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาค่าความสัมพันธ์ทั้ง 2 ตัวแบบ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

	ตัวแบบที่	Kolmogorov-Smirnov ^a			Shapiro-Wilk			
		Statistic	df	Sig.	Statistic	df	Sig.	
สถานการณ์การนำทาง	P_{μ}	1	0.088	451	0.000	0.951	451	0.000
		2	0.097	451	0.000	0.945	451	0.000
	R_{μ}	1	0.171	451	0.000	0.852	451	0.000
		2	0.115	451	0.000	0.901	451	0.000
สถานการณ์การป้องกันการเกิดข้อผิดพลาด	P_{μ}	1	0.296	451	0.000	0.708	451	0.000
		2	0.314	451	0.000	0.699	451	0.000
	R_{μ}	1	0.351	451	0.000	0.636	451	0.000
		2	0.345	451	0.000	0.636	451	0.000

ผลการทดสอบในตารางที่ 4-7 ข้างต้นพบว่าค่า Sig. ของตัวแปรค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาค่าความสัมพันธ์ทั้ง 2 ตัวแบบ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดเป็นดังนี้

- 1) สำหรับสถานการณ์การนำทาง ค่าความถูกต้องของการค้นหาค่าความสัมพันธ์ด้วยตัวแบบที่ 1 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0

- 2) สำหรับสถานการณ์การนำทาง ค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 3) สำหรับสถานการณ์การนำทาง ค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 4) สำหรับสถานการณ์การนำทาง ค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 5) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 6) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่ามากกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 7) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 8) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่ามากกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0

ดังนั้นสรุปได้ว่าการแจกแจงของตัวแปรค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบ ใน 2 สถานการณ์นั้นไม่แจกแจงแบบปกติ

ผลการทดสอบ

จากการวิเคราะห์การแจกแจงข้อมูลข้างต้นพบว่าค่าความถูกต้องของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดไม่มีการแจกแจงแบบปกติ ดังนั้นผู้วิจัยจึงเลือกใช้การทดสอบสมมติฐานแบบไม่อิงพารามิเตอร์ (Non Parametric Test) นั่นคือสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซันสำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) กับการทดสอบต่อไปนี้ โดยที่จะปฏิเสธสมมติฐาน H_0 ได้เมื่อถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05 และค่าสถิติ Z มากกว่า 0 โดยที่ผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks)

การวิเคราะห์เปรียบเทียบนี้เป็นการวิเคราะห์เปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ระหว่างการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 กับการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การนำทาง และสถานการณ์การป้องกันการเกิดข้อผิดพลาด ซึ่งสามารถตั้งสมมติฐานได้ดังนี้

- 1) วิเคราะห์เปรียบเทียบค่าความถูกต้องของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพัทธ์กับข้อมูลซอฟต์แวร์อาร์โคฟว์ทั้ง 2 ตัวแบบ ในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าความถูกต้องของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าความถูกต้องของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 2) วิเคราะห์เปรียบเทียบค่าเรียกคืนของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพัทธ์กับข้อมูลซอฟต์แวร์อาร์โคฟว์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเรียกคืนของการค้นหาความสัมพัทธ์ด้วยตัว

แบบที่ 1 ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าเรียกคืนของการค้นหาทวิภาคความสัมพันธ์ด้วยตัว
แบบที่ 2 ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 3) วิเคราะห์เปรียบเทียบค่าความถูกต้องของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหา
ทวิภาคความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบ ในสถานการณ์การ
ป้องกันการเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าความถูกต้องของการค้นหาทวิภาคความสัมพันธ์
ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

M_2 คือ ค่ามัธยฐานของค่าความถูกต้องของการค้นหาทวิภาคความสัมพันธ์
ด้วยตัวแบบที่ 2 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

- 4) วิเคราะห์เปรียบเทียบค่าเรียกคืนของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหา
ทวิภาคความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบ ในสถานการณ์การป้องกัน
การเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเรียกคืนของการค้นหาทวิภาคความสัมพันธ์ด้วยตัว
แบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

M_2 คือ ค่ามัธยฐานของค่าเรียกคืนของการค้นหาทวิภาคความสัมพันธ์ด้วยตัว
แบบที่ 2 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

ผู้วิจัยใช้การทดสอบสมมติฐานโดยสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) กับการทดสอบต่อไปนี้ ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-8 แสดงสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ของค่าความถูกต้องและค่าเรียกคืนของการค้นหาความสัมพันธ์ตัวแบบที่ 2 เทียบกับตัวแบบที่ 1 ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

	ค่าความถูกต้องของการค้นหาความสัมพันธ์ตัวแบบที่ 2 - ค่าความถูกต้องของการค้นหาความสัมพันธ์ตัวแบบที่ 1		ค่าเรียกคืนของการค้นหาความสัมพันธ์ตัวแบบที่ 2 - ค่าเรียกคืนของการค้นหาความสัมพันธ์ตัวแบบที่ 1	
	สถานการณ์การนำทาง	สถานการณ์การป้องกันการเกิดข้อผิดพลาด	สถานการณ์การนำทาง	สถานการณ์การป้องกันการเกิดข้อผิดพลาด
Z	-2.669 ^a	-6.054 ^a	-9.234 ^b	-2.000 ^a
Asymp. Sig. (2-tailed)	0.008	0.000	0.000	0.046

a. Based on positive ranks.

b. Based on negative ranks.

จากตารางที่ 4-8 ได้ผลทดสอบดังนี้

1) การเปรียบเทียบค่าความถูกต้องในสถานการณ์การนำทาง ได้สถิติทดสอบค่า Z เท่ากับ -2.669 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบหางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.008 / 2 = 0.004$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05

2) การเปรียบเทียบค่าความถูกต้องในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ได้สถิติทดสอบค่า Z เท่ากับ -6.054 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดี่ยวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

3) การเปรียบเทียบค่าเรียกคืนในสถานการณ์การนำทาง ได้สถิติทดสอบค่า Z เท่ากับ -9.234 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดี่ยวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางลบ (Based on negative ranks) ดังนั้นจึงสามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มากกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05

4) การเปรียบเทียบค่าเรียกคืนในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ได้สถิติทดสอบค่า Z เท่ากับ -2.000 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดี่ยวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

สรุปผลการทดสอบ

จากการวิเคราะห์ผลการเปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ผู้วิจัยสามารถสรุปได้ว่าในสถานการณ์การนำทางการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ให้ค่าเรียกคืน (Recall) มากกว่าแต่ให้ค่าความถูกต้อง (Precision) ที่น้อยกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 สำหรับสถานการณ์การป้องกันการเกิด

ข้อผิดพลาดการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ให้ค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ไม่ต่างกันหรือน้อยกว่าการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1

4.4.2 การเปรียบเทียบประสิทธิภาพการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์ของทั้ง 2 ตัวแบบโดยเปลี่ยนข้อกำหนดของการสร้างเซตของคำแนะนำ

การเปรียบเทียบประสิทธิภาพการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์ของทั้ง 2 ตัวแบบโดยสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของกฎความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ผู้วิจัยกำหนดให้วิธีการสร้างเซตของคำแนะนำดังกล่าวเรียกว่า “การสร้างเซตของคำแนะนำแบบใหม่” ผลการทดสอบดังกล่าวแสดงดังต่อไปนี้

ตารางที่ 4-9 แสดงตารางผลการทดสอบเพิ่มเติมทั้ง 3 สถานการณ์

สถานการณ์การนำทาง (ค่าเฉลี่ยของค่าเอฟเมสเซอร์)		สถานการณ์การป้องกันการเกิดข้อผิดพลาด (ค่าเฉลี่ยของค่าเอฟเมสเซอร์)		สถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว (ค่าผลสะท้อนกลับ)	
ตัวแบบที่ 1	ตัวแบบที่ 2	ตัวแบบที่ 1	ตัวแบบที่ 2	ตัวแบบที่ 1	ตัวแบบที่ 2
0.3195	0.3335	0.1152	0.1057	$\frac{41}{60}$	$\frac{41}{60}$

จากตารางสรุปข้างต้นแสดงให้เห็นว่าค่าประสิทธิภาพของสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้นมีค่าแตกต่างกัน ค่าประสิทธิภาพของสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้วนั้นไม่มีความแตกต่างกัน แต่ไม่สามารถสรุปได้ว่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ในสถานการณ์ต่าง ๆ นั้นแตกต่างกันอย่างมีนัยสำคัญ ผู้วิจัยจึงจำเป็นต้องวิเคราะห์ผลการทดสอบความแตกต่างกันอย่างมีนัยสำคัญ ดังรายละเอียดต่อไปนี้

การวิเคราะห์การแจกแจงข้อมูล

เนื่องจากผู้วิจัยต้องการเปรียบเทียบประสิทธิภาพของการค้นหากฎความสัมพันธ์ของ 2 ตัวแบบสำหรับสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด โดยการ

สร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของกฎความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ดังนั้นผู้วิจัยจะตรวจสอบว่าค่าเอฟเมเชอร์ทั้งหมดมีการแจกแจงแบบปกติหรือไม่ เฉพาะในสำหรับสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด โดยตั้งสมมติฐานของการทดสอบค่าเอฟเมเชอร์แต่ละการทดสอบมีการแจกแจงแบบปกติหรือไม่ภายใต้สมมติฐานทางสถิติ ดังนี้

- 1) ทดสอบการแจกแจงของข้อมูลค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง

H_0 : ข้อมูลค่าประสิทธิภาพของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าประสิทธิภาพของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง ไม่แจกแจงแบบปกติ

- 2) ทดสอบการแจกแจงของข้อมูลค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง

H_0 : ข้อมูลค่าประสิทธิภาพของการค้นกฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าประสิทธิภาพของการค้นกฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง ไม่แจกแจงแบบปกติ

- 3) ทดสอบการแจกแจงของข้อมูลค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

H_0 : ข้อมูลค่าประสิทธิภาพของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าประสิทธิภาพของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ไม่แจกแจงแบบปกติ

- 4) ทดสอบการแจกแจงของข้อมูลค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

H_0 : ข้อมูลค่าประสิทธิภาพของการค้นหาค่าความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าประสิทธิภาพของการค้นหาค่าความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ไม่มีการแจกแจงแบบปกติ

ในการตรวจสอบการแจกแจงของข้อมูลว่าเป็นแบบปกติโดยใช้สถิติทดสอบนั้น มีสถิติทดสอบที่ใช้คือ Kolmogorov-Smirnov สำหรับหน่วยทดลองมากกว่า 50 หน่วย และ Shapiro-Wilk สำหรับหน่วยทดลองน้อยกว่า 50 หน่วย โดยจะยอมรับสมมติฐาน H_0 เมื่อค่า Sig. มีค่ามากกว่าค่า α ซึ่งกำหนดให้เท่ากับ 0.05 ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-10 แสดงค่าสถิติทดสอบการแจกแจงปกติ (Normality Test) ของค่าประสิทธิภาพของการค้นหาค่าความสัมพันธ์ทั้ง 2 ตัวแบบ โดยการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

สถานการณ์	ตัวแบบที่	Kolmogorov-Smirnov		
		Statistic	df	Sig.
สถานการณ์การนำทาง	1	0.071	451	0.000
	2	0.070	451	0.000
สถานการณ์การป้องกันการเกิดข้อผิดพลาด	1	0.296	451	0.000
	2	0.325	451	0.000

ผลการทดสอบในตารางที่ 4-10 ชี้ให้เห็นพบว่าค่า Sig. ของตัวแปรค่าประสิทธิภาพการค้นหาค่าความสัมพันธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดเป็นดังนี้

- 1) สำหรับสถานการณ์การนำทาง การค้นหาค่าความสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0

- 2) สำหรับสถานการณ์การนำทาง การค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 3) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด การค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 4) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด การค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0

ดังนั้นสรุปได้ว่าการแจกแจงของตัวแปรค่าประสิทธิภาพของการค้นหาความสัมพัทธ์ทั้ง 2 ตัวแบบสำหรับ 2 สถานการณ์นั้นไม่เป็นแบบปกติ

การวิเคราะห์เปรียบเทียบประสิทธิภาพการค้นหาความสัมพัทธ์ทั้ง 2 ตัวแบบ

จากการวิเคราะห์การแจกแจงข้อมูลข้างต้นพบว่าค่าประสิทธิภาพของการค้นหาความสัมพัทธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดโดยการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ไม่แจกแจงแบบปกติ ดังนั้นผู้วิจัยจึงเลือกใช้สถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) กับการทดสอบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด โดยที่จะปฏิเสธสมมติฐาน H_0 ได้เมื่อถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05 และค่าสถิติ Z มากกว่า 0 โดยที่ผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) และผู้วิจัยเลือกใช้การทดสอบอัตราส่วน 2 กลุ่มด้วยสถิติทดสอบ Z (Two Proportion Z Tests) กับการทดสอบในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว โดยที่จะปฏิเสธสมมติฐาน H_0 ได้เมื่อถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05

การวิเคราะห์เปรียบเทียบนี้เป็นกรวิเคราะห์เปรียบเทียบประสิทธิภาพระหว่างการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 กับการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่

ตามมาของกฎความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ในสถานการณ์ 3 สถานการณ์คือ สถานการณ์การนำทาง สถานการณ์การป้องกันการเกิดข้อผิดพลาด สถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว ซึ่งสามารถตั้งสมมุติฐานได้ดังนี้

- 1) วิเคราะห์เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 2) วิเคราะห์เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

M_2 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 3) วิเคราะห์เปรียบเทียบประสิทธิภาพของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของ

คำแนะนำแบบใหม่ ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว

M_2 คือ ค่ามัธยฐานของค่าประสิทธิภาพของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

ผู้วิจัยทดสอบสมมติฐานข้อ 1 และ 2 ด้วยสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) แสดงดังตารางที่ 4-11 สำหรับสมมติฐานข้อ 3 ผู้วิจัยเลือกใช้การทดสอบอัตราส่วน 2 กลุ่มด้วยสถิติทดสอบ Z (Two Proportion Z Tests) แสดงดังตารางที่ 4-12

ตารางที่ 4-11 แสดงค่าสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ของค่าประสิทธิภาพของการค้นหาความสัมพันธ์ตัวแบบที่ 2 เทียบกับตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

	ค่าประสิทธิภาพของการค้นหาความสัมพันธ์ตัวแบบที่ 2 - ค่าประสิทธิภาพของการค้นหาความสัมพันธ์ตัวแบบที่ 1	
	สถานการณ์การนำทาง	สถานการณ์การป้องกันการเกิดข้อผิดพลาด
Z	-3.159 ^a	-3.878 ^b
Asymp. Sig. (2 tailed)	.000	.000

a. Based on negative ranks.

b. Based on positive ranks.

จากตารางที่ 4-11 การวิเคราะห์เปรียบเทียบประสิทธิภาพของการค้นหาภูมิต้านทานทั้ง 2 ตัวแบบในสถานการณ์การนำทาง ได้สถิติทดสอบค่า Z เท่ากับ -3.159 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ เนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางลบ (Based on negative ranks) ดังนั้นจึงสามารถปฏิเสธสมมติฐาน H_0 ได้ นั่นคือค่าประสิทธิภาพของการค้นหาภูมิต้านทานด้วยตัวแบบที่ 2 มากกว่าค่าประสิทธิภาพของการค้นหาภูมิต้านทานด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05 และสำหรับการวิเคราะห์เปรียบเทียบประสิทธิภาพของการค้นหาภูมิต้านทานทั้ง 2 ตัวแบบในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ได้สถิติทดสอบค่า Z เท่ากับ -3.878 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ เนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติฐาน H_0 ได้ นั่นคือค่าประสิทธิภาพของการค้นหาภูมิต้านทานด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าค่าประสิทธิภาพของการค้นหาภูมิต้านทานด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

ตารางที่ 4-12 แสดงการทดสอบอัตราส่วน 2 กลุ่มด้วยสถิติทดสอบ Z ของค่าประสิทธิภาพของการค้นหาภูมิต้านทานด้วยตัวแบบที่ 2 เทียบกับตัวแบบที่ 1 โดยการสร้างเซตของค่าแนะนำแบบใหม่ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.000 ^a	1	1.000		
Continuity Correction ^b	.000	1	1.000		
Likelihood Ratio	.000	1	1.000		
Fisher's Exact Test				1.000	.578
N of Valid Cases	120				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 19.00.

จากตารางที่ 4-12 ค่าสถิติ Z เท่ากับรากที่สองของค่า Pearson Chi-Square นั้นค่า 0.000 และมีเครื่องหมายเดียวกับผลต่างของอัตราส่วนที่ 2 กับอัตราส่วนที่ 1 ($\frac{41}{60} - \frac{41}{60} = 0.000$) นั้นค่าเครื่องหมายบวก และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $1.000 / 2 = 0.500$ ซึ่งมากกว่าค่า $\alpha = 0.05$ ดังนั้นจึงไม่สามารถปฏิเสธสมมติฐาน H_0 ได้ นั่นคือค่าประสิทธิภาพของการค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 2 ไม่ต่างกับค่าประสิทธิภาพของการค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้วที่ระดับนัยสำคัญ 0.05

สรุปผลการวิเคราะห์ข้อมูล

จากการวิเคราะห์ผลการทดสอบประสิทธิภาพของการค้นหาทวิความสัมพันธ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของทวิความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด โดยใช้ค่าเอฟเมสเซอร์ และสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว โดยใช้ค่าผลสะท้อนกลับ ผู้วิจัยสามารถสรุปได้ว่าการค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 2 มีประสิทธิภาพมากกว่าการค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง การค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 2 มีประสิทธิภาพไม่ต่างกันหรือน้อยกว่าการค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด และสถานการณ์การเปลี่ยนแปลงแก้ไขที่สมบูรณ์แล้ว การค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 2 มีประสิทธิภาพไม่ต่างกับการค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 1

ผลสรุปดังกล่าวไม่ได้แสดงให้เห็นว่า การกำหนดการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของทวิความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรกนั้นสามารถเพิ่มค่าความถูกต้อง (Precision) ให้กับการค้นหาทวิความสัมพันธ์ด้วยตัวแบบที่ 2 อย่างที่ผู้วิจัยได้ตั้งข้อสังเกตไว้หรือไม่ ดังนั้นผู้วิจัยจึงมีความต้องการศึกษาเพิ่มเติมคือ การเปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาทวิความสัมพันธ์ด้วยตัวแบบทั้ง 2 ตัวแบบ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด เพื่อทดสอบว่ากำหนดการสร้างเซตของ



คำแนะนำดังกล่าวสามารถเพิ่มค่าความถูกต้อง (Precision) หรือค่าเรียกคืน (Recall) ในสถานการณ์ทั้ง 2 ได้หรือไม่ การทดสอบดังกล่าวแสดงในหัวข้อต่อไป

4.4.3 การเปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์ของทั้ง 2 ตัวแบบโดยเปลี่ยนข้อกำหนดของการสร้างเซตของคำแนะนำ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

ผลการทดสอบประสิทธิภาพการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์ของทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของกฎความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรกหรือการสร้างเซตของคำแนะนำแบบใหม่ในการทดสอบที่แล้วนั้นเป็นการเปรียบเทียบประสิทธิภาพจากการใช้ค่าเอฟเมเชอร์ (F-measure) เท่านั้น ผู้วิจัยจึงมีความต้องการศึกษาเพิ่มเติมโดยการทดสอบเปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์ของทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-13 แสดงตารางค่าความถูกต้องและค่าเรียกคืนของการทดสอบเพิ่มเติมโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

สถานการณ์การนำทาง				สถานการณ์การป้องกันการเกิดข้อผิดพลาด			
ตัวแบบที่ 1		ตัวแบบที่ 2		ตัวแบบที่ 1		ตัวแบบที่ 2	
P_{μ}	R_{μ}	P_{μ}	R_{μ}	P_{μ}	R_{μ}	P_{μ}	R_{μ}
0.4034	0.3856	0.4191	0.3985	0.4621	0.5477	0.4568	0.4945

จากตารางสรุปข้างต้นแสดงให้เห็นว่าค่าเฉลี่ยของค่าความถูกต้อง (Precision) หรือ P_{μ} และค่าเฉลี่ยของค่าเรียกคืน (Recall) หรือ R_{μ} ของการทดสอบข้างต้นนั้นมีค่าแตกต่างกันทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด แต่ไม่สามารถสรุปได้ว่าค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหากฎความสัมพันธ์ของตัว

แบบทั้ง 2 ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้นแตกต่างกันอย่างมีนัยสำคัญ ผู้วิจัยจึงจำเป็นต้องวิเคราะห์ผลการทดสอบความแตกต่างกันอย่างมีนัยสำคัญ ดังรายละเอียดต่อไปนี้

การวิเคราะห์การแจกแจงปกติ

เนื่องจากผู้วิจัยต้องการเปรียบเทียบค่าความถูกต้อง (Precision) หรือค่าเรียกคืน (Recall) ของการค้นหากฎความสัมพันธ์ของ 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ สำหรับสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ดังนั้นผู้วิจัยจะตรวจสอบว่าค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ทั้งหมดของทั้ง 2 สถานการณ์มีการแจกแจงแบบปกติหรือไม่ โดยตั้งสมมติฐานของการทดสอบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) แต่ละสถานการณ์มีการแจกแจงแบบปกติหรือไม่ภายใต้สมมติฐานทางสถิติ ดังนี้

- 1) ทดสอบการแจกแจงของข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง
 - H_0 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ
 - H_1 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง ไม่แจกแจงแบบปกติ
- 2) ทดสอบการแจกแจงของข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง
 - H_0 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ
 - H_1 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง ไม่แจกแจงแบบปกติ
- 3) ทดสอบการแจกแจงของข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด
 - H_0 : ข้อมูลค่าความถูกต้องของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าเรียกคืนของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ไม่แจกแจงแบบปกติ

8) ทดสอบการแจกแจงของข้อมูลค่าเรียกคืนของการค้นกฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

H_0 : ข้อมูลค่าเรียกคืนของการค้นกฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าเรียกคืนของการค้นกฎความสัมพันธ์ตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ไม่แจกแจงแบบปกติ

ในการตรวจสอบการแจกแจงของข้อมูลว่าเป็นแบบปกติโดยใช้สถิติทดสอบนั้น มีสถิติทดสอบที่ใช้คือ Kolmogorov-Sminov สำหรับหน่วยทดลองมากกว่า 50 หน่วย และ Shapiro-Wilk สำหรับหน่วยทดลองน้อยกว่า 50 หน่วย โดยจะยอมรับสมมติฐาน H_0 เมื่อค่า Sig. มีค่ามากกว่าค่า α ซึ่งกำหนดให้เท่ากับ 0.05 ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-14 แสดงค่าสถิติทดสอบการแจกแจงปกติ (Normality Test) ของค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นกฎความสัมพันธ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

	ตัวแบบ ที่	Kolmogorov-Smirnov ^a			Shapiro-Wilk			
		Statistic	df	Sig.	Statistic	df	Sig.	
สถานการณ์ การนำทาง	P_μ	1	0.112	451	0.000	0.955	451	0.000
		2	0.111	451	0.000	0.949	451	0.000
	R_μ	1	0.113	451	0.000	0.906	451	0.000
		2	0.09	451	0.000	0.927	451	0.000
สถานการณ์ การป้องกัน การเกิดข้อผิดพลาด	P_μ	1	0.339	451	0.000	0.673	451	0.000
		2	0.333	451	0.000	0.683	451	0.000
	R_μ	1	0.366	451	0.000	0.633	451	0.000
		2	0.344	451	0.000	0.636	451	0.000

ผลการทดสอบในตารางที่ 4-14 ข้างต้นพบว่าค่า Sig. ของตัวแปรค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาภูควมสัมพันธ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดเป็นดังนี้

- 1) สำหรับสถานการณ์การนำทาง ค่าความถูกต้องของการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 2) สำหรับสถานการณ์การนำทาง ค่าความถูกต้องของการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 3) สำหรับสถานการณ์การนำทาง ค่าเรียกคืนของการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 4) สำหรับสถานการณ์การนำทาง ค่าเรียกคืนของการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 5) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าความถูกต้องของการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 6) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าความถูกต้องของการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่ามากกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0

- 7) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าเรียกคืนของการค้นหา ความสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธ สมมติฐาน H_0
- 8) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าเรียกคืนของการค้นหา ความสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่ามากกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธ สมมติฐาน H_0

ดังนั้นสรุปได้ว่าการแจกแจงของตัวแปรค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ใน ทั้ง 2 สถานการณ์นั้นไม่แจกแจงแบบปกติ

ผลการทดสอบ

จากการวิเคราะห์การแจกแจงข้อมูลข้างต้นพบว่าค่าความถูกต้องของการค้นหา ความสัมพันธ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง และสถานการณ์การป้องกันการเกิดข้อผิดพลาดไม่มีการแจกแจงแบบปกติ ดังนั้นผู้วิจัยจึงเลือกใช้ การทดสอบสมมติฐานแบบไม่อิงพารามิเตอร์ (Non Parametric Test) นั่นคือสถิติทดสอบ เครื่องหมายลำดับที่ของวิลคอกซันสำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) กับการทดสอบต่อไปนี้ โดยที่จะปฏิเสธ สมมติฐาน H_0 ได้เมื่อถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05 และค่าสถิติ Z มากกว่า 0 โดยที่ผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks)

การวิเคราะห์เปรียบเทียบนี้เป็นการวิเคราะห์เปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ระหว่างการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 กับการค้นหา ความสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง และสถานการณ์การป้องกันการเกิดข้อผิดพลาด ซึ่งสามารถตั้งสมมติฐานได้ดังนี้

- 1) วิเคราะห์เปรียบเทียบค่าความถูกต้องของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหา ความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของ คำแนะนำแบบใหม่ ในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

2) วิเคราะห์เปรียบเทียบค่าเรียกคืนของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์โคฟว์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

3) วิเคราะห์เปรียบเทียบค่าความถูกต้องของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์โคฟว์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

M_2 คือ ค่ามัธยฐานของค่าความถูกต้องของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : \mu_2 \leq \mu_1$$

$$H_1 : \mu_2 > \mu_1$$

- 4) วิเคราะห์เปรียบเทียบค่าเรียกคืนของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

M_2 คือ ค่ามัธยฐานของค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

ผู้วิจัยจึงใช้การทดสอบสมมติฐานแบบไม่อิงพารามิเตอร์ (Non Parametric Test) นั่นคือ สถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) กับการทดสอบต่อไปนี้ โดยที่ผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-15 แสดงสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ของค่าความถูกต้องและค่าเรียกคืนของการค้นหาภูควมสัมพันธ์ตัวแบบที่ 2 เทียบกับตัวแบบที่ 1 โดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

	ค่าความถูกต้องของการค้นหาภูควมสัมพันธ์ตัวแบบที่ 2 - ค่าความถูกต้องของการค้นหาภูควมสัมพันธ์ตัวแบบที่ 1		ค่าเรียกคืนของการค้นหาภูควมสัมพันธ์ตัวแบบที่ 2 - ค่าเรียกคืนของการค้นหาภูควมสัมพันธ์ตัวแบบที่ 1	
	สถานการณ์การนำทาง	สถานการณ์การป้องกันการเกิดข้อผิดพลาด	สถานการณ์การนำทาง	สถานการณ์การป้องกันการเกิดข้อผิดพลาด
Z	-3.342 ^a	-3.878 ^b	-3.290 ^a	-4.116 ^b
Asymp. Sig. (2-tailed)	0.001	0.000	0.001	0.000

a. Based on negative ranks.

b. Based on positive ranks.

จากตารางที่ 4-15 ได้ผลทดสอบดังนี้

1) การเปรียบเทียบค่าความถูกต้องในสถานการณ์การนำทาง ได้สถิติทดสอบค่า Z เท่ากับ -3.342 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.001 / 2 = 0.0005$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางลบ (Based on negative ranks) ดังนั้นจึงสามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าความถูกต้องของการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 2 มากกว่าการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05

2) การเปรียบเทียบค่าความถูกต้องในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ได้สถิติทดสอบค่า Z เท่ากับ -3.878 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าความถูกต้องของการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าการค้นหาภูควมสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

3) การเปรียบเทียบค่าเรียกคืนในสถานการณ์การนำทาง ได้สถิติทดสอบค่า Z เท่ากับ -3.290 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.001 / 2 = 0.0005$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางลบ (Based on negative ranks) ดังนั้นจึงสามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มากกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05

4) การเปรียบเทียบค่าเรียกคืนในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ได้สถิติทดสอบค่า Z เท่ากับ -4.116 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเรียกคืนของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

สรุปผลการทดสอบ

จากการวิเคราะห์ผลการเปรียบเทียบค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบโดยการสร้างเซตของคำแนะนำแบบใหม่ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ผู้วิจัยสามารถสรุปได้ว่า ในสถานการณ์การนำทางการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ให้ค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) มากกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาดการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ให้ค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1

จากผลการทดสอบนี้และผลการทดสอบในหัวข้อ 4.4.1 แสดงให้เห็นว่าการกำหนดการสร้างเซตของคำแนะนำแบบใหม่หรือการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรกนั้น ทำให้ค่าความถูกต้อง (Precision) ของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ในสถานการณ์การนำทางเพิ่มขึ้นได้ตามที่ผู้วิจัยได้ตั้งข้อสังเกตไว้ แต่สำหรับใน

สถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้นการกำหนดการสร้างเซตของคำแนะนำแบบใหม่ไม่ได้ทำให้ค่าความถูกต้อง (Precision) ของการค้นหาจากความสัมพันธ์ด้วยตัวแบบที่ 2 ดีขึ้น ด้วยเหตุนี้ผู้วิจัยจึงตั้งข้อสังเกตต่อไปว่าสาเหตุที่ทำให้ผลการทดสอบออกมาเป็นเช่นนั้นเพราะในสถานการณ์การนำทางจากความสัมพันธ์ที่ได้รับการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในอันดับต้นๆนั้นมีความสัมพันธ์ที่เป็นผลบวกวงปะปนอยู่มากในขณะที่ความสัมพันธ์ที่ได้รับการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ในอันดับต้นๆนั้นมีความสัมพันธ์ที่เป็นผลบวกวงน้อยกว่าแต่ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้นความสัมพันธ์ที่ได้รับการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในอันดับต้นๆนั้นมีความสัมพันธ์ที่เป็นผลบวกวงปะปนอยู่น้อย ในขณะที่ความสัมพันธ์ที่ได้รับการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ในอันดับต้นๆนั้นมีความสัมพันธ์ที่เป็นผลบวกวงมากกว่า

ด้วยเหตุนี้ผู้วิจัยจึงมีความต้องการที่จะทดสอบเพิ่มเติมตามข้อสังเกตข้างต้นเพื่ออธิบายสาเหตุของผลการทดสอบหลักของงานวิจัยนี้ ผู้วิจัยกำหนดการทดสอบเพิ่มเติมโดยการเปรียบเทียบประสิทธิภาพการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์ของทั้ง 2 ตัวแบบโดยปรับจำนวนของความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็น 7 อันดับแรก 5 อันดับแรก และ 3 อันดับแรก ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ตามลำดับ

4.4.4 การเปรียบเทียบประสิทธิภาพการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์ของทั้ง 2 ตัวแบบโดยปรับจำนวนของความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำ

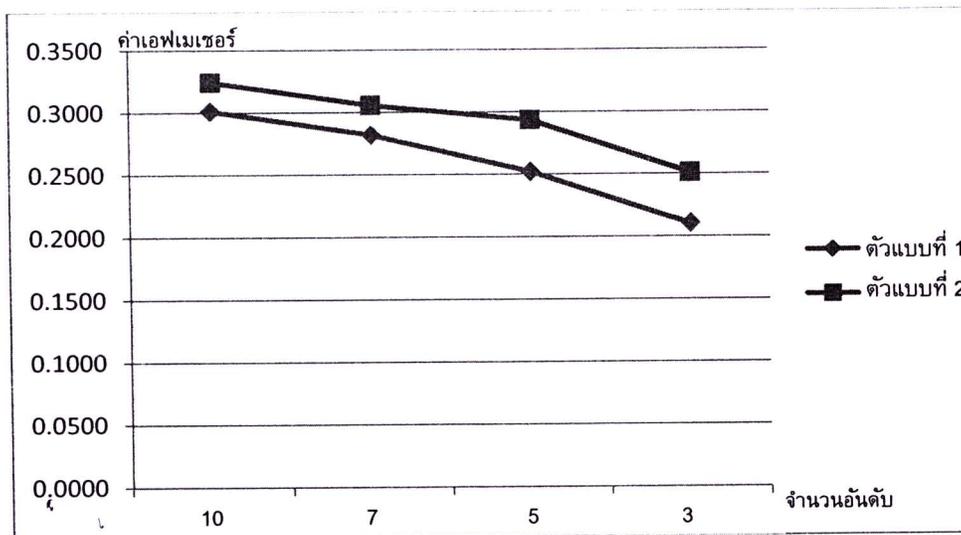
ผู้วิจัยยังตั้งข้อสังเกตอีกว่าความสัมพันธ์ทั้ง 10 อันดับแรกที่ได้จากการค้นหาความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นนั้น ความสัมพันธ์ที่ตรงกับเซตของผลลัพธ์ที่คาดไว้มักจะไม่ใช่ความสัมพันธ์ที่อยู่ในอันดับแรกๆ แต่สำหรับการค้นหาความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่นั้น ความสัมพันธ์ที่ตรงกับเซตของผลลัพธ์ที่คาดไว้มักจะเป็นความสัมพันธ์ที่อยู่ในอันดับแรกๆ เหตุผลที่เป็นเช่นนั้นผู้วิจัยคิดว่าอาจเป็นไปได้ว่าความสัมพันธ์ที่ได้จากการค้นหาความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นในอันดับแรกๆมักเป็นผลบวกวง ผู้วิจัยจึงสนใจวิเคราะห์เปรียบเทียบประสิทธิภาพ

ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์ของทั้ง 2 ตัวแบบในสถานการณ์นำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด โดยทดลองปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ คือ 1) สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก 2) สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก และ 3) สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 3 อันดับแรก ผลการทดสอบแสดงดังตารางต่อไปนี้

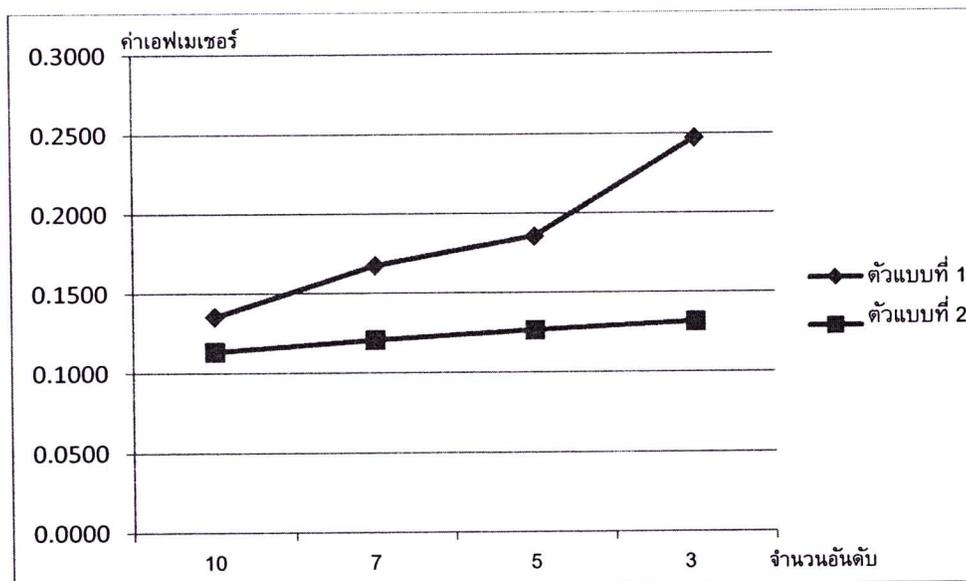
ตารางที่ 4-16 แสดงตารางค่าเฉลี่ยของค่าเอฟเมสเซอร์ของการทดสอบการปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

	ตัวแบบที่	ค่าเฉลี่ยของค่าเอฟเมสเซอร์			
		10 อันดับแรก	7 อันดับแรก	5 อันดับแรก	3 อันดับแรก
สถานการณ์การนำทาง	1	0.3013	0.2820	0.2522	0.2105
	2	0.3245	0.3058	0.2936	0.2511
สถานการณ์การป้องกันการเกิดข้อผิดพลาด	1	0.1353	0.1673	0.1857	0.2472
	2	0.1135	0.1207	0.1267	0.1316

จากตารางที่ 4-16 ข้างบนสามารถนำมาแสดงในรูปแบบข้างกราฟเส้นเพื่อให้เห็นแนวโน้มค่าเฉลี่ยของค่าเอฟเมสเซอร์ของการค้นหากฎความสัมพันธ์ทั้ง 2 ตัวแบบ เมื่อปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ได้ดังนี้



รูปที่ 4-3 แสดงกราฟเส้นค่าเฉลี่ยของค่าเอฟเมเซอร์ของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบ เมื่อปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ ในสถานการณ์การนำทาง



รูปที่ 4-4 แสดงกราฟเส้นค่าเฉลี่ยของค่าเอฟเมเซอร์ของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบ เมื่อปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

จากรูปที่ 4-3 จะเห็นว่าค่าเฉลี่ยของค่าเอฟเมเซอร์ของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 มีค่ามากกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในทุกๆ การปรับจำนวนของกฎ

ความสัมพันธ์ที่นำมาสร้างเป็นเซตของค่าแนะนำเป็นค่าต่างๆ ในสถานการณ์การนำทาง นอกจากนั้นจะเห็นว่าค่าเฉลี่ยของค่าเอฟเมสเซอร์ของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบมีแนวโน้มลดลงเรื่อยๆ เมื่อมีการปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของค่าแนะนำเป็นค่าที่น้อยลง แสดงให้เห็นว่าการค้นหาความสัมพันธ์ในสถานการณ์การนำทางนั้น กฎความสัมพันธ์ที่ถูกต้องจริงๆ นั้นมันจะเป็นกฎความสัมพันธ์ที่อยู่ในอันดับท้ายๆ และกฎความสัมพันธ์ที่อยู่ในอันดับต้นๆ นั้นเป็นผลบวกลงในทั้ง 2 ตัวแบบ

จากรูปที่ 4-4 จะเห็นว่าค่าเฉลี่ยของค่าเอฟเมสเซอร์ของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 มีค่ามากกว่าการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ในทุกๆ การปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของค่าแนะนำเป็นค่าต่างๆ นอกจากนั้นจะเห็นว่าค่าเฉลี่ยของค่าเอฟเมสเซอร์ของการค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบมีแนวโน้มสูงขึ้นเรื่อยๆ เมื่อมีการปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของค่าแนะนำเป็นค่าที่น้อยลง สาเหตุมาจากค่าเรียกคืนที่ยังคงที่เสมอในขณะที่ค่าความถูกต้องมีค่ามากขึ้นเรื่อยๆ (ขนาดเซตของค่าแนะนำที่ลดลงทำให้ค่าความถูกต้องสูงขึ้น ในขณะที่มีกฎความสัมพันธ์ที่ถูกต้องเท่าเดิมหรือไม่ลดลงจึงทำให้ค่าเรียกคืนเท่าเดิม) แสดงให้เห็นว่าการค้นหาความสัมพันธ์ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้น กฎความสัมพันธ์ที่ถูกต้องจริงๆ นั้นมันจะเป็นกฎความสัมพันธ์ที่อยู่ในอันดับต้นๆ ในทั้ง 2 ตัวแบบ

จากตารางและกราฟข้างต้นแสดงให้เห็นว่าค่าเอฟเมสเซอร์หรือค่าประสิทธิภาพของการทดสอบข้างต้นนั้นมีค่าแตกต่างกันทั้ง 2 ตัวแบบเมื่อปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของค่าแนะนำเป็นค่าต่างๆ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด แต่ไม่สามารถสรุปได้ว่าประสิทธิภาพของการค้นหาความสัมพันธ์ของตัวแบบทั้ง 2 เมื่อปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของค่าแนะนำเป็นค่าต่างๆ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้นแตกต่างกันอย่างมีนัยสำคัญ ผู้วิจัยจึงจำเป็นต้องวิเคราะห์ผลการทดสอบความแตกต่างกันอย่างมีนัยสำคัญ ดังรายละเอียดต่อไปนี้

การวิเคราะห์การแจกแจงปกติ

เนื่องจากผู้วิจัยต้องการเปรียบเทียบประสิทธิภาพของการค้นหากฎความสัมพันธ์ของ 2 ตัวแบบสำหรับสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด โดยทดลองปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ ดังนั้นผู้วิจัยจะตรวจสอบว่าค่าเอฟเมเชอร์ทั้งหมดมีการแจกแจงแบบปกติหรือไม่ โดยตั้งสมมติฐานของการทดสอบค่าเอฟเมเชอร์แต่ละการทดสอบมีการแจกแจงแบบปกติหรือไม่ภายใต้สมมติฐานทางสถิติ ดังนี้

- 1) ทดสอบการแจกแจงของข้อมูลค่าเอฟเมเชอร์ของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง

H_0 : ข้อมูลค่าเอฟเมเชอร์ของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าเอฟเมเชอร์ของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง ไม่มีการแจกแจงแบบปกติ
- 2) ทดสอบการแจกแจงของข้อมูลค่าเอฟเมเชอร์ของการค้นกฎความสัมพันธ์ตัวแบบที่ 2 โดยสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง

H_0 : ข้อมูลค่าเอฟเมเชอร์ของการค้นกฎความสัมพันธ์ตัวแบบที่ 2 โดยสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ

H_1 : ข้อมูลค่าเอฟเมเชอร์ของการค้นกฎความสัมพันธ์ตัวแบบที่ 2 โดยสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง ไม่มีการแจกแจงแบบปกติ
- 3) ทดสอบการแจกแจงของข้อมูลค่าเอฟเมเชอร์ของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก ในสถานการณ์การนำทาง

H_0 : ข้อมูลค่าเอฟเมเชอร์ของการค้นกฎความสัมพันธ์ตัวแบบที่ 1 โดยสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก ในสถานการณ์การนำทาง มีการแจกแจงแบบปกติ

ในการตรวจสอบการแจกแจงของข้อมูลว่าเป็นแบบปกติโดยใช้สถิติทดสอบนั้น มีสถิติทดสอบที่ใช้คือ Kolmogorov-Sminov สำหรับหน่วยทดลองมากกว่า 50 หน่วย และ Shapiro-Wilk สำหรับหน่วยทดลองน้อยกว่า 50 หน่วย โดยจะยอมรับสมมติฐาน H_0 เมื่อค่า Sig. มีค่ามากกว่าค่า α ซึ่งกำหนดให้เท่ากับ 0.05 ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-17 แสดงค่าสถิติทดสอบการแจกแจงปกติ (Normality Test) ของค่าเอฟเมเซอร์ของการค้นกฎความสัมพันธ์ทั้ง 2 ตัวแบบ โดยปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

สถานการณ์	จ.กฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำ	ตัวแบบที่	Kolmogorov-Smirnov			Shapiro-Wilk		
			Statistic	df	Sig.	Statistic	df	Sig.
สถานการณ์การนำทาง	3	1	.127	451	.000	.886	451	.000
		2	.130	451	.000	.929	451	.000
	5	1	.087	451	.000	.952	451	.000
		2	.085	451	.000	.957	451	.000
	7	1	.073	451	.000	.960	451	.000
		2	.081	451	.000	.960	451	.000
สถานการณ์การป้องกันการเกิดข้อผิดพลาด	3	1	.372	451	.000	.700	451	.000
		2	.409	451	.000	.674	451	.000
	5	1	.350	451	.000	.732	451	.000
		2	.361	451	.000	.752	451	.000
	7	1	.311	451	.000	.798	451	.000
		2	.340	451	.000	.779	451	.000

ผลการทดสอบในตารางที่ 4-17 ชี้ให้เห็นพบว่าค่า Sig. ของตัวแปรค่าเอฟเมเซอร์ของการค้นหากฎความสัมพันธ์ทั้ง 2 ตัวแบบ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดเป็นดังนี้



เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0

- 10) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด การค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 ที่สร้างเขตของค่าแนะนำจากกฎความสัมพัทธ์ 5 อันดับแรก มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่ามากกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 11) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด การค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ที่สร้างเขตของค่าแนะนำจากกฎความสัมพัทธ์ 3 อันดับแรก มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0
- 12) สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด การค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 ที่สร้างเขตของค่าแนะนำจากกฎความสัมพัทธ์ 3 อันดับแรก มีค่า Sig. เท่ากับ 0.000 ซึ่งมีค่ามากกว่าค่าระดับนัยสำคัญ $\alpha = 0.05$ ดังนั้นจึงปฏิเสธสมมติฐาน H_0

ดังนั้นสรุปได้ว่าการแจกแจงของตัวแปรค่าประสิทธิภาพของการค้นหาความสัมพัทธ์ทั้ง 2 ตัวแบบโดยปรับจำนวนของกฎความสัมพัทธ์ที่นำมาสร้างเป็นเขตของค่าแนะนำเป็นค่าต่างๆ ในทั้ง 2 สถานการณ์นั้นไม่เป็นแบบปกติ

ผลการทดสอบ

จากการวิเคราะห์การแจกแจงข้อมูลข้างต้นพบว่าค่าเอฟเมเชอร์ของการค้นหาความสัมพัทธ์ทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดไม่มีการแจกแจงแบบปกติ ดังนั้นผู้วิจัยจึงเลือกใช้การทดสอบสมมติฐานแบบไม่อิงพารามิเตอร์ (Non Parametric Test) นั่นคือสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซันสำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) กับการทดสอบต่อไปนี้ โดยที่จะปฏิเสธสมมติฐาน H_0 ได้เมื่อถ้าค่า Sig. (Significance) ที่คำนวณได้น้อยกว่า 0.05 และค่าสถิติ Z มากกว่า 0 โดยที่ผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks)

การวิเคราะห์เปรียบเทียบนี้เป็นการวิเคราะห์เปรียบเทียบค่าเอฟเมเชอร์ระหว่างการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 กับการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 โดยปรับจำนวนของกฎความสัมพัทธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ซึ่งสามารถตั้งสมมติฐานได้ดังนี้

- 1) วิเคราะห์เปรียบเทียบค่าเอฟเมเชอร์ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพัทธ์กับข้อมูลซอฟต์แวร์อาร์โคฟว์ทั้ง 2 ตัวแบบ ที่สร้างเซตของคำแนะนำจากกฎความสัมพัทธ์ 7 อันดับแรก ในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของคำแนะนำจากกฎความสัมพัทธ์ 7 อันดับแรก ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 ที่สร้างเซตของคำแนะนำจากกฎความสัมพัทธ์ 7 อันดับแรก ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 2) วิเคราะห์เปรียบเทียบค่าเอฟเมเชอร์ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบ ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก ในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 3) วิเคราะห์เปรียบเทียบค่าเอฟเมเชอร์ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบ ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก ในสถานการณ์การนำทาง ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง

M_2 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การนำทาง

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 4) วิเคราะห์เปรียบเทียบค่าเอฟเมเชอร์ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบ ที่สร้างเซตของคำแนะนำ

จากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

M_2 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 5) วิเคราะห์เปรียบเทียบค่าเอฟเมเชอร์ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบ ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้ M_1 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

M_2 คือ ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 5 อันดับแรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

- 6) วิเคราะห์เปรียบเทียบค่าเอฟเมเชอร์ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหาความสัมพันธ์กับข้อมูลซอฟต์แวร์อาร์ไคฟ์ทั้ง 2 ตัวแบบ ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 3 อันดับแรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ว่ามีความแตกต่างกันหรือไม่

กำหนดให้	M_1 คือ	ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ด้วย ตัวแบบที่ 1 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 3 อันดับ แรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด
	M_2 คือ	ค่ามัธยฐานของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ด้วย ตัวแบบที่ 2 ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 3 อันดับ แรก ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด

$$H_0 : M_2 \leq M_1$$

$$H_1 : M_2 > M_1$$

ผู้วิจัยใช้การทดสอบสมมติฐานโดยสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ (The Wilcoxon Signed Rank Sum Test for the Matched Paired Difference) กับการทดสอบต่อไปนี้ ผลการทดสอบแสดงดังตารางต่อไปนี้

ตารางที่ 4-18 แสดงสถิติทดสอบเครื่องหมายลำดับที่ของวิลคอกซ์สำหรับการทดสอบแบบจับคู่ของค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ทั้ง 2 ตัวแบบ โดยปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ

	ค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ตัวแบบที่ 2 - ค่าเอฟเมเชอร์ของการค้นหากฎความสัมพันธ์ตัวแบบที่ 1					
	สถานการณ์การนำทาง			สถานการณ์การป้องกันข้อผิดพลาด		
จน.กฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำ	3	5	7	3	5	7
Z	-4.795 ^a	-5.815 ^a	-3.973 ^a	-8.766 ^b	-8.280 ^b	-9.852 ^b
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000

a. Based on negative ranks.

b. Based on positive ranks.

จากตารางที่ 4-18 ได้ผลทดสอบดังนี้

1) ในสถานการณ์การนำทาง ที่สร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 7 อันดับแรก ได้สถิติทดสอบค่า Z เท่ากับ -4.795 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียว ดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางลบ (Based on negative ranks) ดังนั้นจึงสามารถปฏิเสธ

สมมติ H_0 ได้ นั่นคือค่าเอฟเมเชอร์ของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 มากกว่าการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของค่าแนะนำจากกฎความสัมพัทธ์ 7 อันดับแรกในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05

2) ในสถานการณ์การนำทาง ที่สร้างเซตของค่าแนะนำจากกฎความสัมพัทธ์ 5 อันดับแรก ได้สถิติทดสอบค่า Z เท่ากับ -5.815 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียว ดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางลบ (Based on negative ranks) ดังนั้นจึงสามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเอฟเมเชอร์ของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 มากกว่าการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของค่าแนะนำจากกฎความสัมพัทธ์ 5 อันดับแรกในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05

3) ในสถานการณ์การนำทาง ที่สร้างเซตของค่าแนะนำจากกฎความสัมพัทธ์ 3 อันดับแรก ได้สถิติทดสอบค่า Z เท่ากับ -3.973 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียว ดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางลบ (Based on negative ranks) ดังนั้นจึงสามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเอฟเมเชอร์ของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 มากกว่าการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของค่าแนะนำจากกฎความสัมพัทธ์ 3 อันดับแรกในสถานการณ์การนำทาง ที่ระดับนัยสำคัญ 0.05

4) ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ที่สร้างเซตของค่าแนะนำจากกฎความสัมพัทธ์ 7 อันดับแรก ได้สถิติทดสอบค่า Z เท่ากับ -8.766 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียว ดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเอฟเมเชอร์ของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของค่าแนะนำจากกฎความสัมพัทธ์ 7 อันดับแรกในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

5) ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ที่สร้างเซตของค่าแนะนำจากกฎความสัมพัทธ์ 5 อันดับแรก ได้สถิติทดสอบค่า Z เท่ากับ -8.280 ซึ่งน้อยกว่า 0 และจากการ

ตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเอฟเมเซอร์ของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของคำแนะนำจากกฎความสัมพัทธ์ 5 อันดับแรกในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

6) ในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ที่สร้างเซตของคำแนะนำจากกฎความสัมพัทธ์ 3 อันดับแรก ได้สถิติทดสอบค่า Z เท่ากับ -9.852 ซึ่งน้อยกว่า 0 และจากการตั้งสมมติฐานเป็นแบบทางเดียวดังนั้นค่า Sig. จึงเท่ากับ $0.000 / 2 = 0.000$ ซึ่งน้อยกว่าค่า $\alpha = 0.05$ และเนื่องจากผลการวิเคราะห์ที่ออกมาตั้งบนพื้นฐานทางบวก (Based on positive ranks) ดังนั้นจึงไม่สามารถปฏิเสธสมมติ H_0 ได้ นั่นคือค่าเอฟเมเซอร์ของการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 ไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ที่สร้างเซตของคำแนะนำจากกฎความสัมพัทธ์ 3 อันดับแรกในสถานการณ์การป้องกันการเกิดข้อผิดพลาดที่ระดับนัยสำคัญ 0.05

สรุปผลการทดสอบ

จากการวิเคราะห์ผลการเปรียบเทียบค่าเอฟเมเซอร์หรือประสิทธิภาพของการค้นหาความสัมพัทธ์ทั้ง 2 ตัวแบบ ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด ของผลการทดสอบนี้และผลการทดสอบหลักในหัวข้อ 4.3 ผู้วิจัยสามารถสรุปได้ว่า ในสถานการณ์การนำทาง การค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 มีค่าเอฟเมเซอร์เฉลี่ยสูงกว่าการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ไม่ว่าจะกำหนดจำนวนของกฎความสัมพัทธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเท่าใดก็ตาม โดยการสร้างเซตของคำแนะนำจากกฎความสัมพัทธ์จำนวนต่าง ๆ กันที่ให้ผลต่างค่าเอฟเมเซอร์หรือประสิทธิภาพของการค้นหาความสัมพัทธ์ทั้ง 2 ตัวแบบสูงที่สุดในสถานการณ์การนำทางคือ การสร้างเซตของคำแนะนำจากกฎความสัมพัทธ์ 5 อันดับแรก 3 อันดับแรก 10 อันดับแรก และ 7 อันดับแรก ตามลำดับ สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาด การค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 2 มีค่าเอฟเมเซอร์เฉลี่ยไม่ต่างกันหรือน้อยกว่าการค้นหาความสัมพัทธ์ด้วยตัวแบบที่ 1 ไม่ว่าจะกำหนดจำนวนของกฎความสัมพัทธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเท่าไรก็ตาม โดยการสร้างเซตของคำแนะนำจากกฎความสัมพัทธ์จำนวนต่าง ๆ กันที่ให้ผลต่างค่าเอฟเมเซอร์หรือประสิทธิภาพของการค้นหา

ความสัมพันธ์ทั้ง 2 ตัวแบบมากที่สุด ในสถานการณ์การป้องกันการเกิดข้อผิดพลาดคือ การสร้างเซตของคำแนะนำจากกฎความสัมพันธ์ 3 อันดับแรก 5 อันดับแรก 7 อันดับแรก และ 10 อันดับแรก ตามลำดับ

สรุปผลการทดลองข้างต้นสามารถตอบข้อสังเกตของผู้วิจัยที่ตั้งไว้ได้คือ ในสถานการณ์การนำทาง การค้นหากฎความสัมพันธ์ทั้ง 2 ตัวแบบมีกฎความสัมพันธ์ที่เป็นผลบวกวงปะปนอยู่ในอันดับต้นๆทั้งคู่ แต่กฎความสัมพันธ์ที่ได้การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในอันดับต้นๆนั้นมีกฎความสัมพันธ์ที่เป็นผลบวกวงปะปนอยู่มากกว่ากฎความสัมพันธ์ที่ได้การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ในอันดับต้นๆ แต่สำหรับในสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้น การค้นหาความสัมพันธ์ทั้ง 2 ตัวแบบมีกฎความสัมพันธ์ที่เป็นผลบวกวงปะปนอยู่ในอันดับต้นๆในปริมาณที่น้อยทั้งคู่ โดยที่กฎความสัมพันธ์ที่ได้การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 1 ในอันดับต้นๆนั้นมีกฎความสัมพันธ์ที่เป็นผลบวกวงปะปนอยู่น้อยกว่ากฎความสัมพันธ์ที่ได้การค้นหาความสัมพันธ์ด้วยตัวแบบที่ 2 ในอันดับต้นๆนั้นมีกฎความสัมพันธ์ที่เป็นผลบวกวง

ผู้วิจัยสังเกตเห็นว่าในสถานการณ์การนำทางนั้น มักจะเกิดกฎความสัมพันธ์ที่เป็นผลบวกวงขึ้นในอันดับต้นๆเสมอไม่ว่าใช้ตัวแบบที่ 1 หรือตัวแบบที่ 2 แต่สำหรับในสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้น กฎความสัมพันธ์ในอันดับต้นๆนั้นค่อนข้างถูกต้องแม่นยำและมีกฎความสัมพันธ์ที่เป็นผลบวกวงปะปนอยู่น้อยไม่ว่าใช้ตัวแบบที่ 1 หรือตัวแบบที่ 2 ด้วยเหตุนี้ผู้วิจัยจึงสนใจที่จะวิเคราะห์ข้อสังเกตดังกล่าวในหัวข้อถัดไป

4.4.5 การวิเคราะห์ค่าประเมินระดับความน่าสนใจของกฎความสัมพันธ์ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

จากการทดสอบเพิ่มเติมในหัวข้อที่ 4.4.4 ทำให้ผู้วิจัยทราบว่าในสถานการณ์การนำทางนั้น มักจะเกิดกฎความสัมพันธ์ที่เป็นผลบวกวงขึ้นในอันดับต้นๆเสมอไม่ว่าใช้ตัวแบบที่ 1 หรือตัวแบบที่ 2 แต่สำหรับในสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้น กฎความสัมพันธ์ในอันดับต้นๆนั้นค่อนข้างถูกต้องแม่นยำและมีกฎความสัมพันธ์ที่เป็นผลบวกวงปะปนอยู่ในอันดับต้นๆน้อยไม่ว่าจะใช้ตัวแบบที่ 1 หรือตัวแบบที่ 2 ผู้วิจัยเห็นว่าสาเหตุที่เป็นเช่นนั้นเพราะการคำนวณค่าความเชื่อมั่นใหม่นั้นแตกต่างจากการคำนวณค่าความเชื่อมั่นใหม่ตรงที่ค่าความเชื่อมั่นใหม่มีการ

นำเอาค่าความน่าจะเป็นในการไม่พบเซตรายการที่ตามมาเข้ามาพิจารณาด้วย นอกจากนั้นผู้วิจัยเห็นว่าสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดมีลักษณะเฉพาะตัวที่แตกต่างกันที่ทำให้ผลของการนำเอาค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนแต่ไม่มีเซตรายการที่ตามมาเข้ามาพิจารณาามีผลแตกต่างกัน ผู้วิจัยจึงสนใจที่จะวิเคราะห์การคำนวณค่าประเมินระดับความน่าเชื่อถือของกฎความสัมพันธ์ทั้ง 2 ค่า รวมถึงการวิเคราะห์ลักษณะเฉพาะของสถานการณ์ทั้ง 2 สถานการณ์ด้วย

ตารางที่ 4-19 แสดงการเปรียบเทียบสูตรคำนวณและพิสัยของค่าความเชื่อมั่นและค่าความเชื่อมั่นใหม่

	ค่าความเชื่อมั่น	ค่าความเชื่อมั่นใหม่
สูตรคำนวณ	$\text{Conf}(X \rightarrow Y) = \frac{P(X \text{ and } Y)}{P(X)}$	$\text{NConf}(X \rightarrow Y) = \frac{P(X \text{ and } Y)}{P(Y)} - \frac{P(X \text{ and } \bar{Y})}{P(\bar{Y})}$
พิสัย	[0,1]	[-1,1]

จากตารางข้างบนแสดงให้เห็นว่าความแตกต่างของค่าความเชื่อมั่นกับค่าความเชื่อมั่นใหม่คือการนำเอาค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนแต่ไม่มีเซตรายการที่ตามมาเข้ามาพิจารณาด้วย ในสูตรคำนวณของค่าความเชื่อมั่นใหม่จะเห็นว่าถ้าพจน์หลังของสูตรมีค่ามากจะทำให้ค่าความเชื่อมั่นใหม่มีค่าน้อย นั่นคือถ้าค่าความน่าจะเป็นในการพบเซตรายการที่มาก่อนในทรานแซคชันที่ไม่มีเซตรายการที่ตามมาอยู่มากค่ามากแล้วกฎความสัมพันธ์นั้นจะมีค่าความเชื่อมั่นใหม่น้อย ในขณะที่สูตรคำนวณของค่าความเชื่อมั่นไม่ได้นำเอาค่าความน่าจะเป็นในการไม่พบเซตรายการที่ตามมาเข้ามาพิจารณาด้วย แต่กฎความสัมพันธ์ที่มีค่าความเชื่อมั่นมากอาจจะมีค่าความน่าจะเป็นในการพบเซตรายการที่มาก่อนในทรานแซคชันที่ไม่มีเซตรายการที่ตามมาอยู่ที่มากก็เป็นไปได้ ซึ่งถ้าเป็นเช่นนั้นกฎความสัมพันธ์ดังกล่าวก็คือกฎความสัมพันธ์ที่ถูกแนะนำเป็นอันดับต้นๆแต่จริงๆแล้วเป็นผลบวกลงนั่นเอง (Liu et al., 2008)

ตารางที่ 4-20 แสดงการเปรียบเทียบขนาดของเซตรายการที่มาก่อนและเซตรายการที่ตามมาของกฎความสัมพันธ์ขนาด n ในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาด

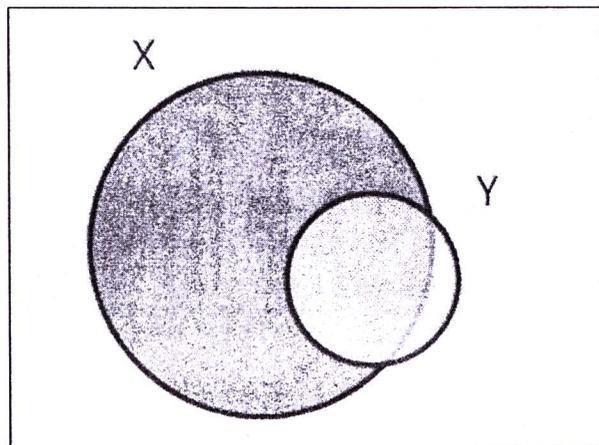
	สถานการณ์การนำทาง	สถานการณ์การป้องกันการเกิดข้อผิดพลาด
ขนาดของเซตรายการที่มาก่อน	1	$n-1$
ขนาดของเซตรายการที่ตามมา	$n-1$	1

จากตารางข้างบนแสดงให้เห็นว่าความแตกต่างของกฎความสัมพันธ์สถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดก็คือ ในสถานการณ์การนำทางขนาดของเซตรายการที่ตามมามีขนาดใหญ่ขึ้นเป็นเชิงเส้นเมื่อขนาดของกฎความสัมพันธ์ใหญ่ขึ้นโดยที่มีขนาดของเซตรายการที่มาก่อนเป็น 1 เสมอ สำหรับสถานการณ์การป้องกันการเกิดข้อผิดพลาดขนาดของเซตรายการที่ตามมาเป็น 1 เสมอโดยที่มีขนาดของเซตรายการที่มาก่อนจะมีขนาดใหญ่ขึ้นเป็นเชิงเส้นเมื่อขนาดของกฎความสัมพันธ์ใหญ่ขึ้น

ในปีค.ศ. 2004 Lenca และคณะ (Lenca et al, 2004) ก็เสนอคุณสมบัติ 5 ข้อที่ค่าประเมินความน่าสนใจของกฎความสัมพันธ์ควรจะมี หลังจากนั้นในปี ค.ศ. 2007 Lenca และคณะได้ปรับปรุงคุณสมบัติทั้ง 5 ข้อนั้นเล็กน้อยเพื่อให้มีความยืดหยุ่นมากขึ้น (Lenca et al, 2007) คุณสมบัติทั้ง 5 ข้อของ Lenca และคณะนั้นได้แก่คุณสมบัติ Q1 Q2 Q3 Q4 และ Q5 ที่อธิบายไว้ในบทที่ 2 คุณสมบัติ Q1 นั้นกล่าวไว้ว่า กฎความสัมพันธ์นั้นมีความน่าสนใจสูงสุด เมื่อค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนแต่ไม่มีเซตรายการที่ตามมาในฐานข้อมูลเท่ากับ 0 และค่าประเมินความน่าสนใจของกฎความสัมพันธ์ควรจะเป็นค่าคงที่ค่าใดค่าหนึ่งหรือเป็นค่าอนันต์เพื่อสื่อความหมายอย่างชัดเจนว่ากฎความสัมพันธ์นั้นน่าสนใจสูงสุดด้วย

จากคุณสมบัตินี้จะเห็นว่าการประเมินว่ากฎความสัมพันธ์ใดมีความน่าสนใจสูงนั้น ไม่สามารถพิจารณาเพียงแค่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนและเซตรายการที่ตามมาเท่านั้น แต่จะต้องพิจารณาที่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนแต่ไม่มีเซตรายการที่ตามมาควบคู่ด้วย หรือเพียงพิจารณาที่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนแต่ไม่มีเซตรายการที่ตามมาเป็นหลัก สาเหตุที่

เป็นเช่นนั้นก็เพราะในบางกรณีที่กฎความสัมพันธ์มีค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนและเซตรายการที่ตามมามีค่าสูง แต่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนแต่ไม่มีเซตรายการที่ตามมามีค่าสูงกว่า หรือกล่าวคือในทรานแซคชันที่มีเซตรายการที่มาก่อนอยู่มากจะไม่มีเซตรายการที่ตามมาเป็นส่วนใหญ่ แต่ในทรานแซคชันที่มีเซตรายการที่ตามมาอยู่มากจะมีเซตรายการที่มาก่อนเป็นส่วนใหญ่ ตามรูปด้านล่างนี้



รูปที่ 4-5 แสดงที่กฎความสัมพันธ์มีค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนและเซตรายการที่ตามมามีค่าสูง แต่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนแต่ไม่มีเซตรายการที่ตามมามีค่าสูงกว่า

กฎความสัมพันธ์ที่พบว่า ทรานแซคชันส่วนใหญ่่นั้นที่มีเซตรายการที่มาก่อนอยู่มากจะไม่มีเซตรายการที่ตามมา สามารถกล่าวได้ว่าเป็นกฎความสัมพันธ์เซตรายการที่มาก่อนกับเซตรายการที่ตามมาที่มีความสัมพันธ์เชิงลบต่อกัน ถ้ากฎความสัมพันธ์นี้ถูกจัดอันดับว่ามีความน่าสนใจในอันดับต้นๆ กฎความสัมพันธ์ก็คือผลบวกลงนั่นเอง (Liu et al., 2008) ดังนั้นการใช้ค่าประเมินความน่าสนใจที่ขึ้นกับค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนและเซตรายการที่ตามมาเป็นหลักเพียงอย่างเดียว อย่างเช่น ค่าสนับสนุน ค่าความเชื่อมั่น ค่าลิฟท์ เป็นต้น จึงมีโอกาสที่จะทำให้เกิดกฎความสัมพันธ์ที่เป็นผลบวกลงได้ง่าย

กรณีที่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนและเซตรายการที่ตามมามีค่าสูง แต่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนแต่ไม่มีเซตรายการที่ตามมามีค่าสูงกว่าดังรูปด้านบนนั้น จะมีโอกาสเกิดขึ้นได้ง่ายเมื่อพบว่าค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนมีสูงมากๆ ในขณะที่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่ตามามีน้อยกว่ามากๆ เมื่อพิจารณาที่รูปแบบกฎความสัมพันธ์

ในสถานการณ์การนำทาง จะเห็นว่าเซตรายการที่มาก่อนที่มีขนาดเท่ากับ 1 เสมอนั้นมีโอกาสสูงที่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนมีค่าสูงมากๆ ในขณะที่เซตรายการที่ตามมาที่มีขนาดเพิ่มขึ้นเป็นเชิงเส้นตามขนาดของกฎความสัมพันธ์นั้นมีโอกาสสูงที่ค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่ตามมาจะมีน้อยมากๆ ด้วยเหตุนี้กฎความสัมพันธ์ในสถานการณ์การนำทางที่ใช้ค่าประเมินความน่าสนใจของกฎความสัมพันธ์ที่ขึ้นกับค่าความน่าจะเป็นในการพบทรานแซคชันที่มีเซตรายการที่มาก่อนและเซตรายการที่ตามมาเป็นหลักเพียงอย่างเดียวจึงมีโอกาสเกิดผลบวกลงได้สูง

เมื่อโอกาสเกิดกฎความสัมพันธ์ที่เป็นผลบวกลงในสถานการณ์การนำทางมีมาก การนำเอาค่าความน่าจะเป็นในการไม่พบเซตรายการที่ตามมาเข้ามาพิจารณาด้วยของค่าความเชื่อมั่นใหม่จึงสามารถลดระดับความน่าสนใจของกฎความสัมพันธ์ที่เป็นผลบวกลงได้มากเช่นกัน สำหรับในสถานการณ์การป้องกันการเกิดข้อผิดพลาด ค่าความน่าจะเป็นในการไม่พบเซตรายการที่ตามมาที่น้อยส่งผลกฎความสัมพันธ์ที่มีค่าความเชื่อมั่นหรือค่าเชื่อมั่นใหม่สูงมีกฎความสัมพันธ์ที่เป็นผลบวกลงปะปนอยู่น้อย การนำเอาค่าความน่าจะเป็นในการไม่พบเซตรายการที่ตามมาเข้ามาพิจารณาด้วยของค่าความเชื่อมั่นใหม่จึงสามารถช่วยลดระดับความน่าสนใจของกฎความสัมพันธ์ที่เป็นผลบวกลงได้น้อยด้วย

4.4.6 สรุปผลการศึกษาเพิ่มเติม

การทดสอบการเปรียบเทียบค่าประสิทธิภาพหรือค่าเอฟเมเชอร์ (F-measure) ค่าความถูกต้อง (Precision) และค่าเรียกคืน (Recall) ของการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์ของทั้ง 2 ตัวแบบในสถานการณ์การนำทางและสถานการณ์การป้องกันการเกิดข้อผิดพลาดโดยการสร้างเซตของคำแนะนำแบบปกติกับการสร้างเซตของคำแนะนำจากการยูเนียน (Union) ของรายการการเปลี่ยนแปลงแก้ไขที่อยู่ในเซตรายการที่ตามมาของกฎความสัมพันธ์อันดับสูงสุดจำนวน 10 รายการแรก ทำให้ผู้วิจัยทราบว่าการค้นหากฎความสัมพันธ์ด้วยตัวแบบที่ 2 มีประสิทธิภาพเพิ่มขึ้นเมื่อพิจารณาเพียงกฎความสัมพันธ์ที่อยู่ในอันดับต้นๆ ผู้วิจัยจึงศึกษาเพิ่มเติมโดยการทดสอบการเปรียบเทียบประสิทธิภาพการทำเหมืองข้อมูลด้วยเทคนิคการค้นหากฎความสัมพันธ์ของทั้ง 2 ตัวแบบโดยปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆคือ 7 อันดับแรก 5 อันดับแรกและ 3 อันดับแรก จากการศึกษาทดสอบนี้ทำให้ผู้วิจัยทราบถึงแนวโน้มการลดลงหรือเพิ่มของประสิทธิภาพของการค้นหากฎความสัมพันธ์ของตัวแบบทั้ง 2 ที่แตกต่างกันในสถานการณ์การนำทางและสถานการณ์การ

ป้องกันการเกิดข้อผิดพลาดเมื่อมีการปรับจำนวนของกฎความสัมพันธ์ที่นำมาสร้างเป็นเซตของคำแนะนำเป็นค่าต่างๆ ผู้วิจัยจึงวิเคราะห์สาเหตุของแนวโน้มที่แตกต่างกันใน 2 สถานการณ์นั้น

การทดสอบเพิ่มทั้งหมดนี้เพื่อศึกษาสาเหตุของผลการทดสอบหลังของงานวิจัยนี้ นั่นคือการเปรียบเทียบประสิทธิภาพของการประยุกต์เทคนิคการค้นหากฎความสัมพันธ์บนข้อมูลซอฟต์แวร์อาร์ไคฟ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นและตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่ ผลการทดสอบแสดงให้เห็นว่าตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่ช่วยเพิ่มประสิทธิภาพให้กับการประยุกต์เทคนิคการค้นหากฎความสัมพันธ์บนข้อมูลซอฟต์แวร์อาร์ไคฟ์ได้เฉพาะในสถานการณ์การนำทางเท่านั้น และผลการทดสอบเพิ่มเติมแสดงให้เห็นว่าสาเหตุที่ตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่ช่วยเพิ่มประสิทธิภาพในสถานการณ์การนำทางเนื่องมาจากการใช้ตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นนั้นสามารถทำให้เกิดกฎความสัมพันธ์ที่เป็นผลบวกลงได้มาก แต่การใช้ตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่สามารถลดการเกิดผลบวกลงเหล่านั้นได้ แต่สำหรับในสถานการณ์การป้องกันการเกิดข้อผิดพลาดนั้นการใช้ตัวแบบทั้ง 2 ให้กฎความสัมพันธ์ที่เป็นผลบวกลงค่อนข้างน้อย และการใช้ตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นให้ค่าความถูกต้องและค่าเรียกคืนที่ดีกว่าการใช้ตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่จึงทำให้ประสิทธิภาพของการค้นหากฎความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นดีกว่าการค้นหากฎความสัมพันธ์ด้วยตัวแบบค่าสนับสนุน-ค่าความเชื่อมั่นใหม่