

บทที่ 2

ทฤษฎีและรายงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 การหารูปแบบความสัมพันธ์ (Larose, D. T., 2005)

รูปแบบทั่วไปของการหาความสัมพันธ์คือ $A \rightarrow B$ โดยที่ A เป็นเงื่อนไข และ B เป็นผลลัพธ์ที่เกิดขึ้น

การประเมินค่าของกฎจะใช้ค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) โดยที่

ค่าสนับสนุนคือร้อยละของข้อมูลที่มีเงื่อนไขและผลลัพธ์สอดคล้องตามกฎต่อจำนวนข้อมูลทั้งหมด สามารถเขียนเป็นสมการดังนี้

$$\text{ค่าสนับสนุน (A, B)} = \frac{\text{จำนวนของรายการ (A, B)}}{\text{จำนวนรายการทั้งหมด}}$$

โดยที่ A หมายถึง เหตุการณ์ที่ใช้เป็นเงื่อนไขในการหาผลลัพธ์

B หมายถึง เหตุการณ์ที่เป็นผลลัพธ์

รายการ (A, B) หมายถึง เหตุการณ์ที่ประกอบด้วยเหตุการณ์ A และ B

ค่าความเชื่อมั่นคือร้อยละของข้อมูลที่มีเงื่อนไขและผลลัพธ์สอดคล้องตามกฎต่อจำนวนข้อมูลทั้งหมดที่เป็นเงื่อนไข สามารถเขียนเป็นสมการดังนี้

$$\text{ค่าความเชื่อมั่น (A, B)} = \frac{\text{จำนวนของรายการ (A, B)}}{\text{จำนวนรายการ (A)}}$$

โดยที่ รายการ (A) หมายถึง เหตุการณ์ที่ประกอบด้วยเหตุการณ์ A เพียงอย่างเดียว

ในการเลือกจะใช้กฎใดนั้นจะต้องพิจารณาค่าสนับสนุนและค่าความเชื่อมั่นที่มีค่าสูงกว่าค่า Threshold ที่ตั้งไว้ ในที่นี้กำหนดค่าสนับสนุนต่ำสุด (Minimum Support) เป็น 0.5 และค่าความเชื่อมั่นต่ำสุด (Minimum Confidence) เป็น 0.5

2.1.2 การเปรียบเทียบสัดส่วนการยืมหนังสือเรียนของนักศึกษาจำแนกตามเพศ ขณะที่นักศึกษาสังกัด ชั้นปี วันที่ยืมหนังสือเรียน และช่วงเวลาที่ยืมหนังสือเรียน

ในการเปรียบเทียบสัดส่วนการยืมหนังสือเรียนของนักศึกษา จะทำการเปรียบเทียบว่าเพศของนักศึกษา คณะที่นักศึกษาสังกัด ชั้นปี วันที่ยืมหนังสือเรียน และช่วงเวลาที่ยืมหนังสือเรียน มีสัดส่วนที่แตกต่างกันหรือไม่ โดยใช้ตัวสถิติทดสอบไค-สแควร์จำแนกทางเดียว (การทดสอบความแตกต่างระหว่างค่าสัดส่วนของประชากร k กลุ่ม) ดังนี้

2.1.2.1 การทดสอบความแตกต่างระหว่างค่าสัดส่วนของประชากร k กลุ่ม ($k \geq 3$) (สายชล สนิทบุรณท์ทอง, 2555)

ในการทดสอบความแตกต่างระหว่างค่าสัดส่วนของประชากรตั้งแต่ 3 กลุ่ม ขึ้นไป จะใช้
ตัวสถิติทดสอบ χ^2 ซึ่งเป็นการทดสอบสมมติฐานสำหรับข้อมูลจำแนกทางเดียว

สมมติฐานที่ใช้ในการทดสอบ

$$H_0 : p_1 = p_2 = \dots = p_k = \frac{1}{k}$$

$$H_1 : \text{มี } p_i \neq p_j \text{ อย่างน้อย 1 ค่า} ; \quad i \neq j ; i, j = 1, \dots, k$$

ตัวสถิติทดสอบ

$$\chi_{\text{cal}}^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

โดยที่ $E_i = np_i = \frac{n}{k}$

อาณาเขตวิกฤต

$$\text{จะปฏิเสธ } H_0 \text{ ถ้า } \chi_{\text{cal}}^2 > \chi_{\alpha, k-1}^2$$

ข้อจำกัดของการทดสอบแบบไค-สแควร์จำแนกทางเดียว

- 1) ค่าสังเกตทุกค่าจะต้องเป็นอิสระกัน
- 2) ข้อมูลในตารางที่จะนำมาวิเคราะห์ควรเป็นข้อมูลที่เป็นความถี่ ไม่ควรอยู่ในรูปสัดส่วนหรือร้อยละ

- 3) จำนวนความถี่ทั้งหมดหรือขนาดตัวอย่างทั้งหมดควรมีขนาดใหญ่ คือ ไม่ต่ำกว่า 50

- 4) ความถี่คาดหวังในแต่ละกลุ่มจะต้องมีอย่างน้อย 1 และมีจำนวนความถี่คาดหวังที่น้อยกว่า 5 ได้ไม่เกิน 20 % ของจำนวนกลุ่มทั้งหมด แต่ถ้ากลุ่มใดมีค่าความถี่คาดหวังน้อยกว่า 5 อาจแก้ไขได้ดังนี้

4.1 เพิ่มขนาดของตัวอย่าง (n) ให้มากขึ้น หรือเพิ่มค่าสังเกตให้มากขึ้น

4.2 รวมกลุ่มที่อยู่ติดกันเข้าด้วยกัน จนกว่าค่าความถี่คาดหวังไม่น้อยกว่า 5 วิธีนี้จะทำได้เมื่อการรวมกลุ่มเข้าด้วยกันแล้วไม่ทำให้ความหมายของกลุ่มที่นำมารวมกันเปลี่ยนแปลงไป

- 5) จำนวนองศาความเป็นอิสระเท่ากับ 1 และขนาดตัวอย่าง $n < 50$ ตัวสถิติไค-สแควร์จะไม่นับความต่อเนื่อง จำเป็นต้องปรับค่าต่อเนื่องด้วยวิธีการของเยทส์ (Yate) ดังนี้

$$\chi^2 = \sum_{i=1}^k \frac{\left(|O_i - E_i| - \frac{1}{2}\right)^2}{E_i}$$

แต่ถ้า $n \geq 50$ ยังคงใช้สูตร

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

2.1.3 การศึกษาความสัมพันธ์ระหว่างเพศ คณะที่นักศึกษาสังกัด และชั้นปี กับวันที่ยืมหนังสือเรียนและช่วงเวลาที่ยืมหนังสือเรียน

ในการศึกษาความสัมพันธ์ระหว่างเพศ คณะที่นักศึกษาสังกัด และชั้นปี กับวันที่ยืมหนังสือเรียนและช่วงเวลาที่ยืมหนังสือเรียน จะทำการศึกษาว่าเพศของนักศึกษา คณะที่นักศึกษาสังกัด และชั้นปี กับวันที่ยืมหนังสือเรียนและช่วงเวลาที่ยืมหนังสือเรียนมีความสัมพันธ์กันหรือไม่ โดยใช้ตัวสถิติทดสอบไค-สแควร์จำแนกสองทาง (การทดสอบความเป็นอิสระ) และมีความสัมพันธ์กันมากน้อยเพียงไร โดยใช้ตัวสถิติทดสอบคราเมอร์วีดังนี้

2.1.3.1 การทดสอบความเป็นอิสระ (Test for Independence) (สายชล สินสมบูรณ์ทอง, 2555)

ถ้าข้อมูลมีการจำแนกตามลักษณะอย่างใดอย่างหนึ่งตั้งแต่ 2 ลักษณะขึ้นไป และต้องการทดสอบว่าลักษณะอย่างหนึ่งเป็นอิสระจากลักษณะอีกอย่างหนึ่งหรือไม่ หรือทดสอบความเป็นอิสระกันของตัวแปรที่ใช้จัดจำแนก 2 ตัว โดยที่ตัวแปรทั้ง 2 ตัว เป็นตัวแปรเชิงคุณภาพ เช่น ต้องการทดสอบว่าอายุของคนขับรถเป็นอิสระจากจำนวนอุบัติเหตุที่เกิดขึ้นหรือไม่ หรือต้องการทดสอบว่าอุปนิสัยการสูบบุหรี่เป็นอิสระจากกลุ่มนักศึกษาหรือไม่ เป็นต้น การทดสอบแบบนี้จะสุ่มตัวอย่างเพียง 1 กลุ่ม ที่มีขนาด n จากประชากรหนึ่ง แล้วนำมาจัดลงในตารางการจรณ์ 2 ทาง ชนิด $r \times c$ โดยจัดจำแนกว่าเป็นแต่ละลักษณะด้วยจำนวนเท่าไร

สมมติฐานที่ใช้ในการทดสอบ

- H_0 : ลักษณะที่ 1 และลักษณะที่ 2 เป็นอิสระกัน
 H_1 : ลักษณะที่ 1 และลักษณะที่ 2 ไม่เป็นอิสระกัน
 หรือ H_0 : ลักษณะที่ 1 และลักษณะที่ 2 ไม่มีความสัมพันธ์กัน
 H_1 : ลักษณะที่ 1 และลักษณะที่ 2 มีความสัมพันธ์กัน
 หรือ H_0 : ลักษณะที่ 1 และลักษณะที่ 2 ไม่ขึ้นอยู่กับกัน
 H_1 : ลักษณะที่ 1 และลักษณะที่ 2 ขึ้นอยู่กับกัน

ตัวสถิติทดสอบ

$$\chi^2_{\text{cal}} = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

การหาความถี่คาดหวัง

$$E_{ij} = \frac{(\text{ผลรวมของความถี่ในแถวตอนที่ } i) (\text{ผลรวมของความถี่ในแถวตั้งที่ } j)}{\text{ผลรวมของความถี่ทั้งหมด}}$$

$$= \frac{R_i C_j}{n}$$

อาณาเขตวิกฤต

$$\text{จะปฏิเสธ } H_0 \text{ ถ้า } \chi^2_{\text{cal}} > \chi^2_{\alpha; (r-1)(c-1)}$$

ข้อจำกัดของการทดสอบแบบไค-สแควร์จำแนกสองทาง

- 1) ค่าสังเกตทุกค่าจะต้องเป็นอิสระกัน
- 2) ข้อมูลในตารางที่จะนำมาวิเคราะห์ควรเป็นข้อมูลที่เป็นความถี่ ไม่ควรอยู่ในรูปสัดส่วนหรือร้อยละ

- 3) จำนวนความถี่ทั้งหมดหรือขนาดตัวอย่างทั้งหมดควรมีขนาดใหญ่ คือ ไม่ต่ำกว่า 50

- 4) ความถี่คาดหวังในแต่ละเซลล์จะต้องมีอย่างน้อย 1 และมีจำนวนความถี่คาดหวังที่น้อยกว่า 5 ได้ไม่เกิน 20% ของจำนวนเซลล์ทั้งหมด แต่ถ้าเซลล์ใดมีความถี่คาดหวังน้อยกว่า 5 อาจแก้ไขได้ดังนี้

4.1 เพิ่มขนาดของตัวอย่าง (n) ให้มากขึ้น หรือเพิ่มค่าสังเกตให้มากขึ้น

4.2 รวมเซลล์ที่อยู่ติดกันเข้าด้วยกัน จนกว่าค่าความถี่คาดหวังไม่น้อยกว่า 5 วิธีนี้จะทำได้เมื่อการรวมเซลล์แล้วไม่ทำให้ความหมายของเซลล์ที่นำมารวมกันเปลี่ยนแปลงไป

- 5) ตารางการจรรยาบรรณขนาด 2 x 2 จะทำให้จำนวนองศาความเป็นอิสระเท่ากับ 1 และขนาดตัวอย่าง $n < 50$ ตัวสถิติไค-สแควร์จะไม่มีค่าต่อเนื่อง จำเป็นต้องปรับค่าต่อเนื่องด้วยวิธีการของเยทส์ (Yate) ดังนี้

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{\left(\left| O_{ij} - E_{ij} \right| - \frac{1}{2} \right)^2}{E_{ij}}$$

ถ้าความถี่คาดหวังมีค่ามาก การปรับค่าและไม่ปรับค่าจะให้ผลเหมือนกัน เมื่อความถี่คาดหวังอยู่ระหว่าง 5 และ 10 ควรจะปรับค่าด้วยวิธีการของเยทส์ สำหรับความถี่คาดหวังที่น้อยกว่า 5 ควรใช้วิธีการของฟิชเชอร์-เออร์วิน (Fisher-Irwin exact test)

แต่ถ้า $n \geq 50$ ยังคงใช้สูตร

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

2.1.3.2 การวัดระดับความสัมพันธ์ (ศุภมาส ไวกุลเพชร และสมฤทัย รัชมิธรรม, 2545)

คราเมอร์วี (Cramer's V) เป็นเครื่องมือที่ใช้วัดระดับความสัมพันธ์ของตัวแปร 2 กลุ่มที่มีความสัมพันธ์กันโดยมีค่าตั้งแต่ 0 ถึง 1 วิธีการนี้เหมาะสมสำหรับข้อมูลที่มีลักษณะเป็นแบบนามบัญญัติ และเมื่อมีข้อมูลสองทางหรือมากกว่าสองทางจะสามารถใช้คราเมอร์วีได้โดยไม่ต้องคำนึงถึงข้อจำกัดเกี่ยวกับการแจกแจงของตัวแปร

ในการนำสัมประสิทธิ์คราเมอร์วีมาใช้ไม่จำเป็นต้องใช้ตัวแปรที่ต่อเนื่องเพื่อนำมาใช้วัดค่า โดยที่สัมประสิทธิ์คราเมอร์วีจะคำนวณได้จากตารางการจรณซึ่งเราสามารถที่จะดูค่าจากแถวและคอลัมน์ และนำค่าต่าง ๆ แทนลงในสูตรดังนี้

$$v = \sqrt{\frac{\chi^2}{n \cdot \text{Min}(r-1, c-1)}}$$

เมื่อ v คือ ค่าคราเมอร์วี

n คือ จำนวนค่าสังเกตทั้งหมด

$\text{Min}(r-1, c-1)$ คือ ค่าที่น้อยกว่าระหว่าง $r-1$ และ $c-1$

การแปลความหมายทำได้ดังนี้

0.00 – 0.25	สัมพันธ์กันน้อย
0.26 – 0.50	สัมพันธ์กันปานกลาง
0.51 – 0.75	สัมพันธ์กันค่อนข้างมาก
0.76 – 1.00	สัมพันธ์กันมาก

2.2 รายงานวิจัยที่เกี่ยวข้อง

การหาความสัมพันธ์สามารถนำไปประยุกต์ใช้ได้หลายสาขาได้แก่ การซื้อสินค้าของลูกค้า การศึกษา กฎหมาย การประกันสุขภาพ และการเงินการธนาคาร เป็นต้น โดยมีการทบทวนวรรณกรรมที่เกี่ยวข้องดังนี้

ด้านการซื้อสินค้าของลูกค้า ลักษณ์สุภา พรหมคนตรี (2550) ศึกษาเรื่องการประยุกต์ใช้เทคนิคเหมืองข้อมูลเพื่อศึกษาพฤติกรรมของลูกค้าศูนย์หนังสือจุฬาลงกรณ์มหาวิทยาลัย พบว่าการทำเหมืองข้อมูลมีข้อจำกัด หากต้องการทำเหมืองข้อมูลให้มีประสิทธิภาพและได้ผลลัพธ์ที่น่าเชื่อถือ นั้นขึ้นอยู่กับความถูกต้องของข้อมูล การเลือกข้อมูลที่เหมาะสมในการทำเหมืองข้อมูล การเลือกใช้ซอฟต์แวร์เหมืองข้อมูลที่เหมาะสม การเลือกใช้โมเดลที่เหมาะสมกับข้อมูล และการกำหนดค่าของ

พารามิเตอร์ต่าง ๆ ในการทำงาน ถึงแม้ว่าวิธีการที่ใช้ในการทำเหมืองข้อมูลที่มีอยู่ในปัจจุบันยังสามารถใช้แก้ปัญหาได้ในระดับหนึ่ง แต่ยังมีข้อจำกัดในด้านความสามารถของระบบและความแน่นอนของผลลัพธ์ที่ได้ ดังนั้นการนำหลักการที่ได้จากงานวิจัยมาทดลอง ประยุกต์ใช้กับข้อมูลที่เกิดขึ้นจริงจะมีส่วนช่วยให้การนำเหมืองข้อมูลมาใช้ในองค์กรมีความน่าเชื่อถือมากขึ้น และอาจกลายเป็นเครื่องมือที่สำคัญต่อการสนับสนุนการตัดสินใจดำเนินงานในองค์กรต่อไปในอนาคต

ด้านการศึกษา กฤษณะ ไวยมัย ชิดชนก ส่งศิริ และธนาวิทย์ รักธรรมานนท์ (2001) ได้นำมาประยุกต์ใช้ในการช่วยทำนายเกรดวิชาต่างๆ ในภาคเรียนต่อไปของนิสิต คณะวิศวกรรมศาสตร์ โดยหาความสัมพันธ์ของผลการเรียนในแต่ละวิชาที่ส่งผลต่อกัน ซึ่งทำให้ได้ว่าวิชาใดบ้างที่มีผลต่อวิชาที่ต้องการจะทำนายเกรดล่วงหน้า โมเดลจะทำนายโดยอ้างอิงจากข้อมูลของนิสิตที่เคยเรียนมา และได้ผลการเรียนเช่นเดียวกับนิสิตคนนั้น

บุญมา เฟ่งชวน (2548) ศึกษาเรื่องการใช้เทคนิคเหมืองข้อมูลเพื่อพัฒนาระบบสนับสนุนการตัดสินใจด้านการผลิตบัณฑิตระดับปริญญาตรี กล่าวว่าความต้องการของตลาดมีการเปลี่ยนแปลงไปตามสภาพการเจริญเติบโตของเศรษฐกิจ ณ ช่วงเวลาหนึ่ง ตลาดแรงงานอาจต้องการบุคคลที่มีความรู้ความสามารถด้านวิศวกรรมศาสตร์ คอมพิวเตอร์ แต่เมื่อเวลาผ่านไป ความต้องการของตลาดก็อาจเปลี่ยนไป ดังนั้นสถาบันการศึกษาซึ่งมีหน้าที่ผลิตบัณฑิตจะต้องมีการกำหนดเป้าหมายการผลิตบัณฑิตเพื่อรองรับการขยายตัวของเศรษฐกิจ การวิจัยครั้งนี้จึงได้นำข้อมูลภาวะการมีงานทำของบัณฑิตมาวิเคราะห์โดยใช้เทคนิคเหมืองข้อมูลคือการค้นหากฎความสัมพันธ์ ซึ่งมีวัตถุประสงค์เพื่อศึกษาและพัฒนาตัวแบบ เพื่อใช้ในการทำนายแนวโน้มเลือกอาชีพแรกหลังสำเร็จการศึกษาของนักศึกษาระดับปริญญาตรี และนำข้อมูลรายบุคคลนักศึกษาของสถาบันอุดมศึกษาของรัฐปีการศึกษา 2548 มาวิเคราะห์กับตัวแบบที่สร้างได้และนำเสนอในรูปแบบของรายงานแบบตารางและกราฟเพื่อนำไปสนับสนุนการตัดสินใจในการผลิตบัณฑิต ผลการวิจัยพบว่าตัวแบบที่พัฒนาขึ้นมีค่าเฉลี่ยร้อยละความถูกต้องเป็น 65.39 ถ้าแบ่งสัดส่วนข้อมูล Training set เป็นร้อยละ 60 และค่าเฉลี่ยร้อยละความถูกต้องเป็น 74.72 ถ้าแบ่งสัดส่วนข้อมูล Training set เป็นร้อยละ 70 และค่าเฉลี่ยร้อยละความถูกต้องเป็น 81.89 ถ้าแบ่งสัดส่วนข้อมูล Training set เป็นร้อยละ 80 จะเห็นได้ว่าเมื่อสัดส่วนของ Training set มากขึ้น ทำให้มีความถูกต้องของผลลัพธ์มากขึ้น ซึ่งสอดคล้องกับการใช้เทคนิคการสุ่มตัวอย่างที่เป็นตัวแทนของประชากรมากขึ้น ซึ่งไม่เพียงแต่ให้สารสนเทศมากขึ้นเท่านั้น แต่ยังให้ความถูกต้องมากขึ้นด้วย ดังนั้นในงานวิจัยนี้จึงสามารถนำตัวแบบมาหาแนวโน้มการเลือกอาชีพแรกหลังสำเร็จการศึกษาของนักศึกษาระดับปริญญาตรีได้และนำข้อมูลรายบุคคลนักศึกษามาประมวลผลกับข้อมูลตัวแบบเพื่อนำไปสนับสนุนการตัดสินใจด้านการผลิตบัณฑิตระดับปริญญาตรีต่อไปได้

ด้านกฎหมาย กฤษณะ ไวยมัย และธีระวัฒน์ พงษ์ศิริปรีดา (2001) ได้ใช้เทคนิคการค้นหากฎความสัมพันธ์และเทคนิคการจำแนกประเภทข้อมูล (Data Classification) มาประยุกต์ใช้เพื่อ

การจัดสรรกฎหมายที่เหมาะสมกับการพิจารณาคดีความ โดยนำเทคนิคการจำแนกประเภทข้อมูลมาสร้างตัวจำแนกข้อมูลจากกฎเกณฑ์ที่ได้จากเทคนิคการค้นหากฎความสัมพันธ์ ตัวจำแนกข้อมูลสามารถนำไปใช้ทำนายคดีความแต่ละคดีว่าควรใช้กฎหมายฉบับใดในการพิจารณาโดยใช้ข้อมูลคดีความของศาลฎีกาซึ่งประกอบด้วยคดีอาญาและคดีแพ่งและใช้การตัดคำด้วยพจนานุกรมภาษาไทยเพื่อแบ่งคดีความให้เป็นวลีสั้น ๆ โดยใช้เทคนิค Suffix array และการหากฎเกณฑ์และชุดทดสอบจะต้องแบ่งเป็น 2 ชุด คือ ระดับกฎหมายและระดับมาตรา ผลการทดลองได้ผลลัพธ์ในการทำนายที่ดีกว่าการใช้การจำแนกประเภทข้อมูลแบบทั่วไปซึ่งอยู่ในรูปของต้นไม้เพื่อการตัดสินใจ (Decision Tree)

ด้านระบบประกันสุขภาพ Viveros, Nearhos and Rothman (1996) ได้ใช้เทคนิคการค้นหากฎความสัมพันธ์และ Neural segmentation มาประยุกต์ใช้กับระบบการประกันสุขภาพข้อมูลที่น่ามาวิเคราะห์คือข้อมูลการเรียกใช้สิทธิ (Claim) ในการรักษาจำนวน 6,800,000 แถว 120 แอทริบิวต์ และข้อมูลการรักษาพยาบาลของแพทย์จำนวน 17,000 แถว 1051 แอทริบิวต์ ย้อนหลัง 5 ปี โดยที่เทคนิคการค้นหากฎความสัมพันธ์นั้นจะใช้ข้อมูลการเรียกใช้สิทธิในการรักษาเพื่อค้นหารูปแบบพฤติกรรม การรักษาโดยคัดเลือกแอทริบิวต์ที่สนใจคือการให้บริการรักษา ผลจากการวิจัยพบว่ามีข้อสงสัยจากผลลัพธ์ที่ได้คือพบข้อผิดพลาดจากการเรียกใช้สิทธิในสัดส่วนที่มาก ทำให้เสียค่าใช้จ่ายมากถึง 550,000 ดอลลาร์ สำหรับรายการนี้ในระยะเวลาระหว่าง 1 ปี และถ้ามีเหตุการณ์แบบนี้เกิดขึ้นอีกโดยที่ไม่สามารถตรวจพบได้ก็จะเกิดการสูญเสียเป็นจำนวนมาก ส่วนเทคนิค Neural segmentation จะใช้ข้อมูลทั้ง 2 ฐานข้อมูลเพื่อจัดกลุ่มของการรักษาที่เกิดจากธรรมชาติหรือเกิดจากการปฏิบัติงานโดยดูจากลักษณะการรักษา

ด้านการเงินการธนาคาร อรุมา นองเนื่อง และณัฐวิ อุดกฤษฎ์ (2553) ศึกษาแบบช่วยวิเคราะห์บริการทางการเงินเพื่อกลุ่มลูกค้านิติบุคคลกรณีศึกษาธนาคารกสิกรไทย โดยมีวัตถุประสงค์เพื่อพัฒนาระบบช่วยวิเคราะห์บริการทางการเงินเพื่อกลุ่มลูกค้านิติบุคคล กรณีศึกษาธนาคารกสิกรไทย เพื่อใช้วางแผนกลยุทธ์การขาย โดยใช้เทคนิคกฎความสัมพันธ์ (Association Rules) ค้นหากฎความสัมพันธ์ระหว่างพฤติกรรมทางการเงินของกลุ่มลูกค้านิติบุคคลกับผลิตภัณฑ์บริการทางการเงินที่ใช้และนำเอาเทคนิคเหมืองข้อมูล (Data mining) มาใช้ในการเตรียมข้อมูลเพื่อการค้นหากฎความสัมพันธ์ เมื่อได้กฎความสัมพันธ์แล้ว ทำให้ทราบถึงแนวทางการใช้ผลิตภัณฑ์ของกลุ่มลูกค้านิติบุคคล ซึ่งเป็นการเพิ่มโอกาสให้เจ้าหน้าที่ธนาคารเสนอขายผลิตภัณฑ์บริการทางการเงินที่มีความเหมาะสมให้กับลูกค้าได้ ผลการประเมินความพึงพอใจระบบโดยรวม โดยมีผู้ประเมินสองกลุ่มคือ กลุ่มผู้เชี่ยวชาญมีค่าเฉลี่ยเท่ากับ 3.96 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.65 และส่วนของผู้ใช้งานมีค่าเฉลี่ยเท่ากับ 4.16 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.52 สรุปได้ว่าผู้ใช้งานระบบมีความพึงพอใจอยู่ในเกณฑ์ดี ดังนั้นระบบที่พัฒนาขึ้นนี้สามารถนำไปใช้งานเพื่อช่วยพิจารณาผลิตภัณฑ์ทางการเงินให้แก่กลุ่มลูกค้านิติบุคคลได้