

การวิเคราะห์ปัจจัยที่ส่งผลต่อการลาออกกลางคันของนักศึกษามหาวิทยาลัยราชภัฏภูเก็ต  
โดยใช้กฎความสัมพันธ์

Analysing Factors Influencing Student Dropout at Phuket Rajabhat University  
Using Association Rules

สุพาณี สลีสองสม\* [1], แสงธรรม ตันติธาดิ [2], วิภาวรรณ บัวทอง [3], ณสิทธิ์ เหล่าเส็น [4],  
พิทา จารุพูนผล [5]

[1, 2, 3, 4, 5] สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏภูเก็ต

\* E-mail : s6011423153@pkru.ac.th .....

### บทคัดย่อ

งานวิจัยนี้วิเคราะห์ปัจจัยที่ส่งผลต่อการลาออกกลางคันของนักศึกษาจากข้อมูลนักศึกษามหาวิทยาลัยราชภัฏภูเก็ตทั้งหมด 7,904 คน ระหว่างปี 2559 – 2563 ข้อมูลนักศึกษาถูกนำมาวิเคราะห์หากฎความสัมพันธ์ (Association Rules) ด้วยอัลกอริทึม Apriori โดยกำหนด minimum confidence เท่ากับ 0.9 เพื่อให้ได้กฎที่มีค่าความมั่นใจสูง ซึ่งค่า minimum support จะถูกปรับลดครั้งละ 0.1 ตั้งแต่ 0.5 ถึง 0.3 เพื่อให้ได้เห็นถึงปัจจัยที่มีความสำคัญลดหลั่นกันลงมา เพื่อวิเคราะห์ความสัมพันธ์จากการกำหนดค่าความเชื่อมั่นที่เหมาะสมที่สุด และสอดคล้องกับปัจจัยที่มีผลต่อการลาออกกลางคัน ผลการทดลองแสดงให้เห็นว่าค่าที่เหมาะสมที่สุดอยู่บนพื้นฐานของกฎความสัมพันธ์ที่มีค่านับสนับสนุนต่ำสุด (Minimum Support) เท่ากับ 0.3 และค่าความเชื่อมั่นต่ำสุด (Minimum Confidence) เท่ากับ 0.9 จากผลการวิเคราะห์สรุปได้ว่าปัจจัยที่มีผลต่อการลาออกกลางคันของนักศึกษาเรียงตามลำดับความสำคัญได้แก่ รายได้ของบิดามารดา อาชีพของบิดามารดา และ เกรดเฉลี่ย

**คำสำคัญ :** กฎความสัมพันธ์, การจำแนกประเภท, ค่าความเชื่อมั่น, เหมืองข้อมูล

### Abstract

This research analysed factors affecting student dropout from a total of 7,904 Phuket Rajabhat University (PKRU) students between 2016 and 2020. Student data were analyzed for Association Rules using the Apriori algorithm with a minimum confidence of 0.9 to obtain high-confidence rules. The minimum support was reduced by 0.1 from 0.5 to 0.3 to observe the descending importance of factors to analyze the correlation based

on the optimal confidence determination and following the factors affecting the halfway dropout. The results indicated that the optimal value was based on the correlation rule with minimum support of 0.3 and minimum confidence of 0.9. The analysis concluded that the factors affecting the student's halfway dropout in order of importance were: parents' income, parent's occupation, and GPA.

**Keyword** : Association Rules, Classification, Confidence, Data Mining

## บทนำ

ปัจจุบันการทำเหมืองข้อมูล (Data Mining) เป็นการวิเคราะห์และสืบค้นองค์ความรู้ หรือสิ่งที่สำคัญออกมาจากข้อมูลจำนวนมาก โดยได้มีการนำเทคนิคการทำเหมืองข้อมูลไปประยุกต์ใช้ในงานด้านต่าง ๆ มากขึ้น เช่น การพัฒนากฎการทำนายผลการเรียนของนักศึกษาชั้นปีที่ 1 โดยใช้เทคนิคการทำเหมืองข้อมูล (พรพมิล ชัยวุฒิศักดิ์ และ ยุวดี กล่อมวิเศษ, 2562) การวิเคราะห์การทำนายการลาออกกลางคันของนักศึกษาระดับปริญญาตรี โดยใช้เทคนิควิธีการทำเหมืองข้อมูล (ชณิดาภา บุญประสม และ จริญญา แสนราช, 2561) การวิเคราะห์การซื้อสินค้ากลุ่มผลิตภัณฑ์อาหารเสริม เครื่องสำอางและอุปโภคบริโภค สำหรับธุรกิจขายตรง โดยวิธีกฎความสัมพันธ์ (สิทธิชัย วรโชติกำจร และ จริญญา แสนราช, 2557) การใช้เทคนิค Association Rule Discovery เพื่อการจัดสรรกฎหมายในการพิจารณาคดีความ (กฤษณะ ไวยมัย และ วีระวัฒน์ พงษ์ศิริปริดา, 2544) จากข้อมูล สถาบันการศึกษามีบทบาทต่อการพัฒนาประเทศและเป็นรากฐานที่สำคัญในการสร้างบุคคลให้ มีความรู้ความสามารถปฏิบัติหน้าที่และสามารถดำรงชีวิตอยู่ในสังคมได้ การยกระดับคุณภาพการศึกษาในระดับอุดมศึกษาเป็นสิ่งที่สำคัญยิ่ง ตัวบ่งชี้ในการประเมินคุณภาพการศึกษาของมหาวิทยาลัยอย่างหนึ่งคือจำนวนนักศึกษาที่สำเร็จการศึกษาในระยะเวลาที่กำหนดไว้ในหลักสูตรในการจัดทำแผนพัฒนาการศึกษาระดับอุดมศึกษาฉบับที่ 11 (พ.ศ.2555-2559) ได้กำหนดเป้าหมายให้มีผู้สำเร็จการศึกษาภายใน 4 ปี คิดเป็นร้อยละ 70 (ณัฐธิดา สุวรรณโณ และ อันธิกา สิงห์เอี่ยม, 2554) จากผลการศึกษาพบว่าสาเหตุการลาออกกลางคันของนักศึกษาวิทยาลัยเทคโนโลยีวิทยุวิทยุการบริหารธุรกิจรามอินทรา มีสาเหตุการลาออกกลางคันที่เป็นสาเหตุหลักคือสาเหตุด้านเศรษฐกิจของครอบครัวของนักศึกษาอยู่ในระดับมากที่สุด (ค่าเฉลี่ย = 4.31) และสาเหตุรองลงมาคือสาเหตุด้านสิ่งเสพติด การทะเลาะวิวาทก้าวร้าวหนีเรียน ชู้สาว อยู่ในระดับมากที่สุด (ค่าเฉลี่ย = 4.23) สาเหตุด้านครอบครัว ของนักเรียนนักศึกษา อยู่ในระดับมาก (ค่าเฉลี่ย =

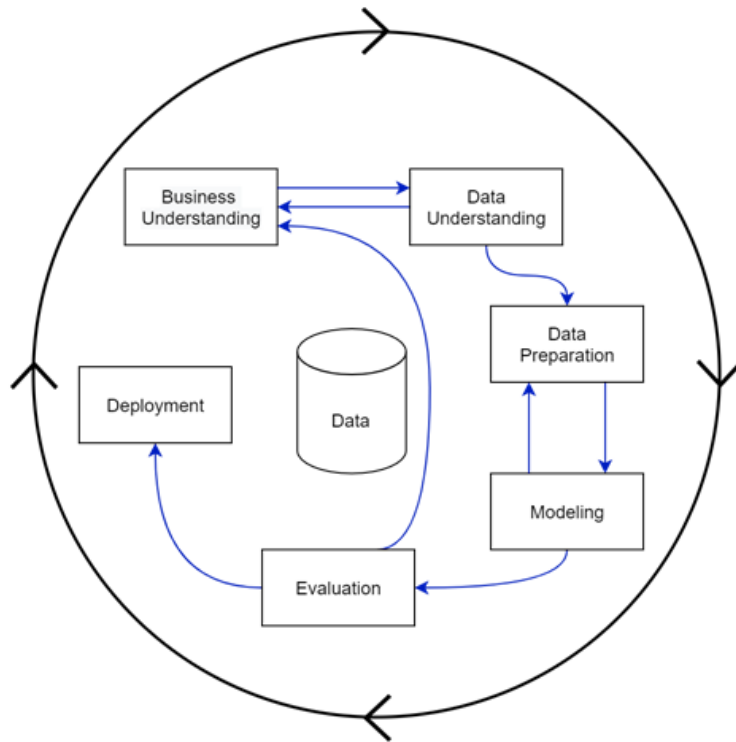
4.16) สาเหตุด้านการเรียนของนักศึกษา อยู่ในระดับมาก (เฉลี่ย = 4.03) และสาเหตุด้านสุขภาพ อยู่ในระดับมาก (เฉลี่ย = 3.93) (สุदारัตน์ พิมลรัตนกานต์, 2558)

การทำเหมืองกฎความสัมพันธ์ (Association Rule Mining) เป็นเทคนิคสำหรับค้นหาความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่เพื่อนำไปหารูปแบบที่เกิดขึ้นบ่อยครั้ง (Frequent Pattern) และใช้ในการวิเคราะห์ความสัมพันธ์หรือทำนายปรากฏการณ์ต่าง ๆ โดยที่กฎความสัมพันธ์จะถูกเขียนในรูปเซตของรายการที่เป็นเหตุ (Premise) ไปสู่เซตของรายการที่เป็นผล (Conclusion) ฐานข้อมูลที่ใช้ในการทำเหมืองกฎความสัมพันธ์ส่วนมากจะเป็นฐานข้อมูลธุรกรรม (Transaction Database) และข้อมูลที่น่านำมาใช้จะอยู่ในรูปแบบ Nominal หรือ Ordinal เท่านั้น อัลกอริทึมการทำเหมืองกฎความสัมพันธ์ที่มีประสิทธิภาพตัวหนึ่งคืออัลกอริทึม Apriori ซึ่งนำค่าสนับสนุนขั้นต่ำ (Minimum Support) และค่าความเชื่อมั่นขั้นต่ำ (Minimum Confidence) มาใช้ในการค้นหากฎความสัมพันธ์ โดยกฎความสัมพันธ์ที่ได้นั้นจะต้องมีค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) ไม่น้อยกว่าค่าขั้นต่ำที่กำหนดเอาไว้ข้างต้น ในขณะที่ค่าสนับสนุน คือ เพอร์เซ็นต์ของจำนวน Item sets ทั้งหมดที่เกิดขึ้นในฐานข้อมูล ค่าความเชื่อมั่น คือ เพอร์เซ็นต์ของจำนวน Item sets ทั้งหมดที่เกิดขึ้นในฐานข้อมูลต่อจำนวน Item sets ที่เกิดขึ้นทางด้านซ้ายมือของกฎ (ปรัชญา นวนแก้ว, 2562)

จากที่กล่าวมาข้างต้น งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์หาปัจจัยที่ส่งผลต่อสาเหตุการลาออกกลางคันของนักศึกษาด้วยการทำเหมืองข้อมูล ซึ่งเป็นวิธีการในการค้นหารูปแบบและแนวโน้มที่มีประโยชน์จากแหล่งข้อมูลขนาดใหญ่ (บุษราภรณ์ มหัทธนะชัยและคณะ, 2559) โดยจะใช้เทคนิคการทำเหมืองความสัมพันธ์ด้วยอัลกอริทึม Apriori เพื่อหากฎความสัมพันธ์ของสาเหตุการลาออกกลางคันของนักศึกษา

### อุปกรณ์และวิธีการวิจัย

อุปกรณ์และวิธีการวิจัยเพื่อศึกษาหาลักษณะสำคัญของปัจจัยในการลาออกกลางคันของนักศึกษาตามหลัก Cross-Industry Standard Process for Data Mining (CRISP-DM) ที่เป็นกระบวนการมาตรฐานสำหรับการทำเหมืองข้อมูล ซึ่งเป็นแบบจำลองกระบวนการมาตรฐานแบบเปิด (Open Standard) ที่อธิบายวิธีการทั่วไปสำหรับการทำเหมืองข้อมูลโดยผู้เชี่ยวชาญ และเป็นรูปแบบการวิเคราะห์ที่ใช้กันอย่างแพร่หลายมากที่สุด (Shearer, 2000) ประกอบด้วย 6 ขั้นตอนหลักที่มีการทำงานตามภาพที่ 1 ดังนี้



ภาพที่ 1: วิธีการวิจัยเพื่อศึกษาหาลักษณะสำคัญของปัจจัยในการลาออกกลางคันของนักศึกษาตามหลัก CRISP-DM (Shearer, 2000)

### ขั้นตอนที่ 1 การทำความเข้าใจธุรกิจ (Business Understanding)

การทำความเข้าใจคือขั้นตอนแรกที่มีงเน้นการทำความเข้าใจวัตถุประสงค์และข้อกำหนดของโครงการ เมื่อทราบถึงวัตถุประสงค์ที่ต้องการทราบถึงปัจจัยที่ส่งผลต่อสาเหตุการลาออกกลางคันของนักศึกษา ผู้วิจัยจึงทำการศึกษาถึงความเป็นไปได้ในการได้มาซึ่งข้อมูลที่เกี่ยวข้อง และความพร้อมของทรัพยากรเพื่อใช้ในการวิเคราะห์ข้อมูล ซึ่งโปรแกรม Weka ที่ถูกพัฒนาโดยมหาวิทยาลัยไวคาโต (University of Waikato) ประเทศนิวซีแลนด์ คือเครื่องมือหลักที่ใช้สำหรับการวิเคราะห์ข้อมูล และรองรับเทคนิคการทำเหมืองข้อมูลหลายรูปแบบ เช่น การประมวลผลข้อมูลล่วงหน้า (Data Preprocessing) การจัดกลุ่ม (Clustering) การจัดประเภท (Classification) และการเลือกคุณสมบัติ (Feature Selection) เป็นต้น

### ขั้นตอนที่ 2 การทำความเข้าใจข้อมูล (Data Understanding)

หลังจากการทำความเข้าใจธุรกิจ ขั้นตอนต่อไปคือการทำความเข้าใจข้อมูล โดยเริ่มต้นจากการรวบรวมข้อมูลจากฐานข้อมูลนักศึกษามหาวิทยาลัยราชภัฏภูเก็ตในช่วงระยะเวลาระหว่างปี 2559 - 2563 ซึ่งทางผู้วิจัยได้ดำเนินการประสานงานผ่านคณบดีคณะวิทยาศาสตร์และเทคโนโลยีขอ

ความอนุเคราะห์ข้อมูลจากผู้อำนวยการสำนักส่งเสริมวิชาการและงานทะเบียน ซึ่งเป็นหน่วยงานที่รับผิดชอบข้อมูลนักศึกษาของมหาวิทยาลัยราชภัฏภูเก็ต หลังจากนั้น ผู้วิจัยจึงทำการคัดเลือกข้อมูลที่เกี่ยวข้องมาเพื่อศึกษาความสอดคล้องของปัจจัยที่เกี่ยวข้องกับการลาออกกลางคันของนักศึกษา โดยเริ่มต้นจากการศึกษาข้อมูลที่เกี่ยวข้องของนักศึกษาเพื่อหาลักษณะสำคัญของปัจจัยในการลาออก หลังจากนั้นจึงทำการคัดกรองข้อมูลในส่วนที่เกี่ยวข้องต่อการลาออก ได้แก่ GPA (เกรดเฉลี่ยสะสม) อาชีพบิดา รายได้บิดา/เดือน อาชีพมารดา รายได้มารดา/เดือน จำนวนพี่น้อง และสถานะนักศึกษา ซึ่งข้อมูลทั้งหมดเป็นข้อมูลระดับนามบัญญัติ (Nominal Scale) ที่ไม่สามารถนำมาคำนวณหรือจัดลำดับแต่สามารถนำมาใช้ในการสร้างกฎความสัมพันธ์ของข้อมูลได้

### ขั้นตอนที่ 3 การจัดเตรียมข้อมูล (Data Preparation)

เนื่องจากข้อมูลที่ได้รับรวบรวมมานั้นยังไม่สามารถใช้งานผ่านกระบวนการทำเหมืองข้อมูลได้ การจัดเตรียมข้อมูลคือขั้นตอนสำคัญในการกรองหรือทำความสะอาดข้อมูล (Data Cleaning) (วิภาวรรณ บัวทอง, 2557) ซึ่งประกอบด้วยกระบวนการตรวจสอบและแก้ไขเพื่อให้รายการข้อมูลที่ไม่มีความสอดคล้องของข้อมูล หรือไม่ถูกต้องออกไปจากชุดข้อมูล เนื่องจากข้อผิดพลาดที่อาจเกิดจากการบันทึกข้อมูล การส่งข้อมูล หรือการให้ความหมายของข้อมูลที่จัดเก็บแตกต่างกัน ดังนั้น จึงต้องมีการจัดการข้อมูลดังนี้

- 1) คัดเลือกเฉพาะคอลัมน์สำคัญที่คาดว่าจะสามารถนำไปใช้ประโยชน์ได้และเป็นคอลัมน์ที่มีข้อมูลค่อนข้างครบถ้วน
- 2) คัดออกคอลัมน์ที่มีค่าทุกแถวเป็นค่าเดียวกัน เพราะเป็นข้อมูลที่ไม่สามารถแยกความแตกต่างของแต่ละแถวได้
- 3) คัดออกคอลัมน์ที่มีข้อมูลไม่ซ้ำกัน เพราะข้อมูลเหล่านี้ไม่สามารถหาแถวที่มีข้อมูลสัมพันธ์กันได้
- 4) แก้ไขข้อมูลให้ถูกต้องสมบูรณ์ เช่น การแก้ไขค่าว่างของข้อมูลโดยจำกัดข้อมูลให้อยู่ในแถวเป็นค่าว่าง (NULL)
- 5) ปรับเปลี่ยนข้อมูลให้มีค่าเหมาะสมในการตัดสินใจ
- 6) แบ่งช่วงข้อมูลหรือจัดกลุ่มข้อมูลเพื่อลดการกระจาย (Data Binning)

### ขั้นตอนที่ 4 การสร้างโมเดล (Modelling)

ข้อมูลที่ถูกคัดกรองจะถูกนำมาวิเคราะห์ด้วยอัลกอริทึม Apriori ซึ่งเป็นอัลกอริทึมพื้นฐานที่ใช้ในการหาความสัมพันธ์ของข้อมูลโดยใช้โปรแกรม Weka อัลกอริทึม Apriori ใช้หลักการค้นหาแบบวง

กว้างก่อนนับธุรกรรม ซึ่งจะทำการสร้างและตรวจสอบเซตไอเท็มที่เกิดขึ้นบ่อยที่ละชั้น โดยเริ่มจากเซตไอเท็มที่มีจำนวนสมาชิกเท่ากับหนึ่ง ถ้าเซตไอเท็มใดมีค่านับสนุนน้อยกว่าค่านับสนุนขั้นต่ำที่กำหนดก็จะตัดเซตไอเท็มนั้นออกและไม่นำไปสร้างเซตไอเท็มในชั้นต่อไป (พินิตา ทรงรัมย์, 2555) และในทำนองเดียวกัน เมื่อสร้างเซตไอเท็มในชั้นใหม่แล้วก็พิจารณาตัดเซตไอเท็มที่มีค่านับสนุนน้อยกว่าค่านับสนุนขั้นต่ำที่กำหนดออกโดยไม่นำไปสร้างเซตไอเท็มในชั้นต่อไป อัลกอริทึมจะทำซ้ำขั้นตอนดังกล่าวไปจนกระทั่งไม่มีเซตไอเท็มใหม่ถูกสร้างขึ้นมาจากนั้นจะนำเซตไอเท็มที่ได้ไปสกัดกฎความสัมพันธ์ต่อไป เพื่อความรวดเร็วจะเก็บเซตไอเท็มในแต่ละระดับชั้นทั้งหมดไว้ในโครงสร้างต้นไม้แฮช (Hash Tree) ซึ่งจุดเด่นของอัลกอริทึมนี้อยู่ที่ความสามารถในความเร็วของการค้นหาไอเท็มเซตที่ปรากฏบ่อยด้วยการละเว้นการพิจารณาไอเท็มเซตที่ปรากฏซ้ำด้วยความถี่ที่ต่ำกว่าเกณฑ์ (วิภาวรรณ บัวทอง, 2557)

#### **ขั้นตอนที่ 5 การประเมินผล (Evaluation)**

หลังจากผ่านสี่ขั้นตอนแรกของรูปแบบกระบวนการมาตรฐานสำหรับการทำเหมืองข้อมูล (CRISP-DM) จะได้รูปแบบโมเดลแล้ว ขั้นตอนต่อไปคือการประเมินผลเพื่อประเมินผลลัพธ์ว่าได้ผลตามวัตถุประสงค์ที่กำหนดไว้หรือไม่ ทบทวนกระบวนการเพื่อหาข้อผิดพลาดและปรับปรุงกระบวนการว่ามีขั้นตอนใดที่ถูกมองข้ามและควรทบทวนหรือไม่ และกำหนดขั้นตอนต่อไป ซึ่งขั้นตอนนี้เริ่มมีการทดสอบแบบจำลองในแอปพลิเคชันเพื่อพิจารณาว่าแบบจำลองนั้นทำงานได้ดีในสถานที่ทำงานเหมือนกับที่ทำการทดสอบหรือไม่

#### **ขั้นตอนที่ 6 การปรับใช้งาน (Deployment)**

การปรับใช้งานคือขั้นตอนหลังจากการประเมินผลและเป็นขั้นตอนสุดท้ายของกระบวนการ CRISP-DM ซึ่งประกอบด้วย การปรับใช้การวางแผน การวางแผนตรวจสอบและบำรุงรักษา การรายงานผล และการทบทวนผลลัพธ์ ซึ่งข้อมูลที่ถูกประเมินจะถูกนำมาปรับใช้งานบนเว็บแอปพลิเคชันที่ถูกพัฒนาขึ้นเพื่อรวบรวมข้อมูลเกี่ยวกับการลาออกของนักศึกษาได้ง่ายโดยผ่านหน้าเว็บที่สร้างขึ้น ภาษาที่ใช้เขียนคือ พีเอชพี (PHP)

### **ผลการทดลองและอภิปรายผล**

#### **ผลการเตรียมข้อมูล**

ชุดข้อมูลนักศึกษาที่ลาออกที่นำมาใช้ในการวิเคราะห์มี 7,904 แถวและมี 12 คอลัมน์ คือ ปีการศึกษา คณะ สาขา สถานะ ชั้นปี GPA อาชีพบิดา อาชีพมารดา จำนวนพี่น้อง รายได้บิดา รายได้มารดา ตัวอย่างข้อมูลดิบแสดงได้ดังตารางที่ 1

ตารางที่ 1: ตัวอย่างข้อมูลดิบที่ยังไม่ได้ผ่านการ Cleaning

ปีการศึกษา	คณะ	สาขา	สถานะ	ชั้นปี	เพศ	GPA	อาชีพบิดา	อาชีพมารดา	จำนวนพี่น้อง	รายได้บิดา	รายได้มารดา
2559	ครุศาสตร์	สาขาวิชาภาษาไทย	10	5	F	3.84	ธุรกิจส่วนตัว	รับราชการ	3	20000	18000
2559	มนุษยศาสตร์และสังคมศาสตร์	สาขาวิชารัฐประศาสนศาสตร์	40	4	M	2.8	ธุรกิจส่วนตัว	ธุรกิจส่วนตัว	3	15000	15000
2559	วิทยาการการจัดการ	สาขาวิชาคอมพิวเตอร์ธุรกิจ	40	4	M	2.55	รับจ้าง	รับจ้าง	2	9000	9000
...	...	...	...	...	...	...	...	...	...	...	...
2563	วิทยาการการจัดการ	สาขาวิชาบริหารธุรกิจ	10	1	F	2.34	รับจ้าง	แม่บ้าน	3	9000	0
2563	วิทยาการการจัดการ	บัญชีบัณฑิต	10	1	F	3	รัฐวิสาหกิจ	แม่บ้าน	2	15000	0
2563	มนุษยศาสตร์และสังคมศาสตร์	สาขาวิชาภาษาอังกฤษเพื่ออุตสาหกรรมบริการ	10	1	F	2.5	ธุรกิจส่วนตัว	รับจ้าง	3	18000	9000

ผู้วิจัยตัดแถวข้อมูลให้เหลือเฉพาะแถวข้อมูลที่มีสถานะเป็น 60 (ลาออก) และเปลี่ยนเป็น "F" ทำให้จำนวนข้อมูลเหลือ 512 แถว ส่วนคอลัมน์ GPA อาชีพบิดา อาชีพมารดา จำนวนพี่น้อง รายได้บิดา และ รายได้มารดา ถูกปรับเปลี่ยนโดยใช้เงื่อนไขที่แสดงในตารางที่ 2 ตัวอย่างผลลัพธ์ของการคัดกรองแสดงได้ดังตารางที่ 3

ตารางที่ 2: การแทนค่าตัวแปรข้อมูลเพื่อการเตรียมข้อมูล

GPA (เกรดเฉลี่ยสะสม)		FW (อาชีพบิดา)		FI (รายได้บิดา/เดือน)	
0-0.99	F	ธุรกิจส่วนตัว	P	น้อยกว่า 9,000	M1
1-1.99	D	รับจ้าง	H	9,000 – 15,000	M2
2-2.99	C	เกษตรกร	F	มากกว่า 15,000	M3
3-3.99	B	พ่อบ้านแม่บ้าน	M		
4	A	ข้าราชการ	O		
		รัฐวิสาหกิจ	S		
Sibling (พี่น้อง)		MW (อาชีพมารดา)		MI (รายได้มารดา/เดือน)	
มีพี่น้อง	yes	ธุรกิจส่วนตัว	P	น้อยกว่า 9,000	M1
ไม่มีพี่น้อง	no	รับจ้าง	H	9,000 – 15,000	M2
		เกษตรกร	F	มากกว่า 15,000	M3
		พ่อบ้านแม่บ้าน	M		
		ข้าราชการ	O		
		รัฐวิสาหกิจ	S		

ตารางที่ 3: ตัวอย่างข้อมูลนักศึกษาที่ผ่านการเตรียมข้อมูลแล้ว

GPA	FW	FI	MW	MI	Sibling	Status
D	H	M1	M	M1	yes	F
F	H	M1	M	M1	yes	F
C	H	M1	F	M2	yes	F
F	P	M2	M	M1	yes	F
C	F	M2	F	M2	yes	F
D	F	M2	F	M2	yes	F
F	P	M2	M	M1	yes	F
D	H	M1	M	M1	yes	F
...	...	...	...	...	...	...
F	O	M2	M	M1	yes	F
F	H	M1	P	M2	yes	F
F	F	M1	P	M1	yes	F

ผลการวิเคราะห์ข้อมูล



ผู้วิจัยใช้โปรแกรม Weka รันอัลกอริทึม Apriori โดยกำหนด Minimum Confidence = 0.9 เพื่อให้ได้กฎที่มีค่าความมั่นใจสูง และเปลี่ยนค่า Minimum Support ลดลงครั้งละ 0.1 ตั้งแต่ 0.5 ถึง 0.3 เพื่อให้ได้เห็นถึงปัจจัยที่มีความสำคัญลดหลั่นกันลงมา

กฎที่ได้จากอัลกอริทึม Apriori มีหลายกฎ ผู้วิจัยเลือกพิจารณาเฉพาะกฎที่มีส่วน Conclusion เป็น ลาออก (Status = F) เท่านั้น เนื่องจากต้องการหาปัจจัยที่เกิดควบคู่กับการลาออก กฎที่เป็นผลลัพธ์แสดงได้ดังตารางที่ 4 ถึงตารางที่ 6

**ตารางที่ 4:** ผลลัพธ์กฎความสัมพันธ์เมื่อกำหนด Minimum Support = 0.5 และ Minimum Confidence = 0.9

Rule No.	Rule
1	Sibling=yes ==> Status=F
2	MI=M1 ==> Status=F
3	MI=M1 Sibling=yes ==> Status=F
4	FI=M1 ==> Status=F
5	FI=M1 Sibling=yes ==> Status=F

**ตารางที่ 5:** ผลลัพธ์กฎความสัมพันธ์เมื่อกำหนด Minimum support = 0.4 และ Minimum Confidence = 0.9

Rule No.	Rule
1	Sibling=yes ==> Status=F
2	MI=M1 ==> Status=F
3	MI=M1 Sibling=yes ==> Status=F
4	FI=M1 ==> Status=F
5	FI=M1 Sibling=yes ==> Status=F
6	FW=H ==> Status=F
7	FW=H FI=M1 ==> Status=F
8	FW=H Sibling=yes ==> Status=F
9	FW=H FI=M1 Sibling=yes ==> Status=F

ตารางที่ 6: ผลลัพธ์กฎความสัมพันธ์เมื่อกำหนด Minimum Support = 0.3 และ Minimum Confidence = 0.9

Rule No.	Rule
1	Sibling=yes ==> Status=F
2	MI=M1 ==> Status=F
3	MI=M1 Sibling=yes ==> Status=F
4	FI=M1 ==> Status=F
5	FI=M1 Sibling=yes ==> Status=F
6	FW=H ==> Status=F
7	FW=H FI=M1 ==> Status=F
8	FW=H Sibling=yes ==> Status=F
9	FW=H FI=M1 Sibling=yes ==> Status=F
10	FI=M2 ==> Status=F
11	MI=M2 ==> Status=F
12	MI=M2 Sibling=yes ==> Status=F
13	FI=M2 Sibling=yes ==> Status=F
14	GPA=F ==> Status=F
15	GPA=F Sibling=yes ==> Status=F
16	FI=M1 MI=M1 ==> Status=F
17	FI=M1 MI=M1 Sibling=yes ==> Status=F
18	FW=H MI=M1 ==> Status=F
19	FW=H FI=M1 MI=M1 ==> Status=F
20	FW=H MI=M1 Sibling=yes ==> Status=F
21	FW=H FI=M1 MI=M1 Sibling=yes ==> Status=F

จากตารางที่ 4 แสดงให้เห็นว่าเมื่อกำหนด Minimum Support เป็น 0.5 จะเห็นว่าปัจจัยที่เกิดขึ้นบ่อยควบคู่กับการลาออก คือ รายได้ของบิดามารดาน้อย (MI = M1 และ FI = M1) เมื่อลดค่า Minimum Support เป็น 0.4 จะเห็นว่า นอกจากรายได้ของบิดามารดาแล้ว ปัจจัยที่เกิดขึ้นบ่อยควบคู่กับการลาออก คือ อาชีพของบิดาเป็นรับจ้าง (FW = H) ดังแสดงในตารางที่ 5 และ เมื่อลด

ค่า Minimum Support เป็น 0.3 จะเห็นว่า นอกจากรายได้และอาชีพของบิดาแล้ว ปัจจัยที่เกิดขึ้น บ่อยควบคู่กับการลาออก คือ เกรดเฉลี่ยน้อย ( $GPA = F$ ) ดังแสดงในตารางที่ 6

จากผลการวิเคราะห์สรุปได้ว่าปัจจัยที่มีผลต่อการลาออกของนักศึกษาเรียงตามลำดับ ความสำคัญได้แก่ รายได้ของบิดามารดา อาชีพของบิดา และ เกรดเฉลี่ย

### พัฒนาเว็บแอปพลิเคชัน

ผู้วิจัยได้พัฒนาเว็บแอปพลิเคชันเพื่อแสดงผลลัพธ์ของการวิเคราะห์ปัจจัยที่มีผลต่อการ ลาออกกลางคันของนักศึกษา เพื่อเผยแพร่แก่ผู้ที่สนใจนำข้อมูลไปใช้ประโยชน์ ตัวอย่างหน้าเว็บ เพจแสดงได้ดังภาพที่ 2



(a)



(b)



(c)

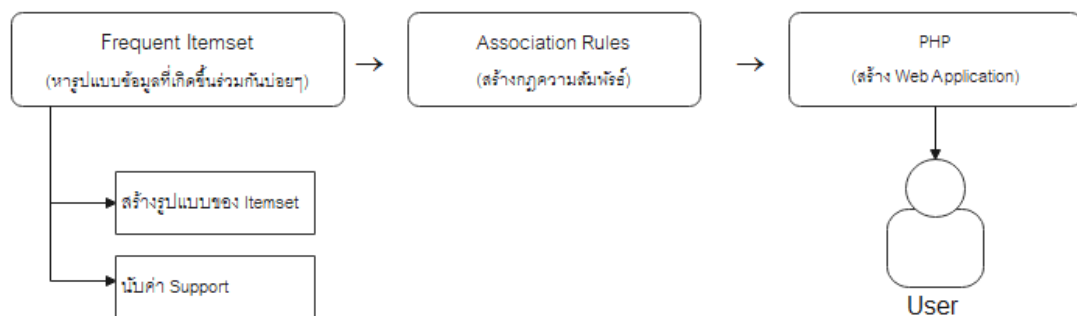


(d)

ภาพที่ 2: เว็บแอปพลิเคชันสำหรับแสดงข้อมูลปัจจัยที่มีผลต่อการลาออกกลางคันของนักศึกษา มหาวิทยาลัยราชภัฏภูเก็ต

### อภิปรายผลการวิจัย

ผู้วิจัยเริ่มต้นศึกษาความเกี่ยวข้องของปัจจัยที่มีผลต่อการลาออกกลางคันของนักศึกษา มหาวิทยาลัยราชภัฏภูเก็ต เพื่อต้องการทราบปัจจัยที่เกี่ยวข้องและค่าของผลการคำนวณที่ถูกต้องแม่นยำที่สุด โดยจากการรวบรวมข้อมูลและวิเคราะห์ศึกษา ทำให้ได้กรอบแนวคิดที่ได้จากการวิจัย ซึ่งเริ่มต้นจากการหารูปแบบข้อมูลที่เกิดขึ้นร่วมกันบ่อยครั้งจากการสร้างรูปแบบของ Itemset และการนับค่า Support หลังจากนั้นจึงสร้างกฎความสัมพันธ์เพื่อหาปัจจัยที่ส่งผลการลาออกของนักศึกษา เมื่อหาค่ากฎความสัมพันธ์ได้แล้ว จึงทำการพัฒนาเว็บแอปพลิเคชันเพื่อรายงานผลจากกฎความสัมพันธ์ในรูปแบบการนำเสนอที่เป็นมิตรกับผู้ใช้ ภาพที่ 3 แสดงกรอบแนวคิดที่ได้จากการวิจัยครั้งนี้



ภาพที่ 3: กรอบแนวคิดที่ได้จากการวิจัย

จากการศึกษาข้อมูลนักศึกษาที่ลาออกและนำเทคนิคการทำเหมืองข้อมูลรูปแบบ Association Rules มาวิเคราะห์ความสัมพันธ์เพื่อหาปัจจัยที่เป็นสาเหตุของการลาออก เพื่อนำผลลัพธ์ที่ได้จากกฎความสัมพันธ์ที่ดีที่สุดมาสรุปผล ผู้วิจัยใช้โปรแกรม Weka รันอัลกอริทึม Apriori 3 ครั้ง โดยกำหนด minimum confidence = 0.9 เพื่อให้ได้กฎที่มีค่าความมั่นใจสูง และเปลี่ยนค่า minimum support ลดลงครั้งละ 0.1 ตั้งแต่ 0.5 ถึง 0.3 เพื่อให้ได้เห็นถึงปัจจัยที่มีความสำคัญลดหลั่นกันลงมา สรุปได้ว่าปัจจัยที่มีผลต่อการลาออกของนักศึกษาโดยมีการเรียงลำดับความสำคัญ ได้แก่ รายได้ของบิดามารดา อาชีพของบิดามารดาและเกรดเฉลี่ย ผลการวิจัยที่มีกฎความสัมพันธ์ที่มีค่าต่ำสุดจะถูกจัดให้เป็นปัจจัยที่ส่งผลต่อการลาออกมากที่สุด

### สรุปผลและเสนอแนะ

กฎความสัมพันธ์ที่มีค่าความเชื่อมั่น (minimum support) และค่าความเชื่อมั่นในกฎความสัมพันธ์ที่เหมาะสมที่สุดในชุดข้อมูลของงานวิจัยนี้คือ minimum support = 0.3 หมายถึง ค่าความเชื่อมั่นในชุดข้อมูลในงานวิจัยนี้เท่ากับ 30% ของจำนวนข้อมูลทั้งหมด และ confidence = 0.9 หมายความว่า ค่าความเชื่อมั่นในกฎความสัมพันธ์ที่ผู้วิจัยต้องการบทความสัมพันธ์มากกว่าหรือเท่ากับ 90% ของกฎความสัมพันธ์ทั้งหมด

ผลการวิจัยที่ใช้เทคนิคการทำเหมืองข้อมูลในการหาความสัมพันธ์ของชุดข้อมูล โดยการนำรูปแบบของข้อมูลที่เกิดขึ้นร่วมกันบ่อย ๆ มาสร้างกฎความสัมพันธ์ สังเกตได้ว่า รายได้ของบิดามารดาน้อย ( $MI = M1$  และ  $FI = M1$ ) ถูกนับเป็น ItemSet ในหลายกฎความสัมพันธ์ สามารถอธิบายได้ว่า ปัจจัยที่ส่งผลต่อการลาออกคือรายได้ของบิดามารดา ควบคู่ไปกับ อาชีพของบิดาที่เป็นอาชีพรับจ้าง ( $FW = H$ ) และเกรดเฉลี่ยน้อย ( $GPA = F$ ) พบว่า จากกฎที่ได้ทำให้ทราบว่าปัจจัยหลักของการลาออกของนักศึกษาคือ รายได้ของบิดามารดา อาชีพของบิดาที่เป็นอาชีพรับจ้างและเกรดเฉลี่ยน้อย ซึ่งสามารถนำมาเป็นแนวทางในการป้องกันการลาออกกลางคันของนักศึกษา

จากการทำเหมืองข้อมูลจากชุดข้อมูลที่ได้นำมาใช้ศึกษาในการวิจัยครั้งนี้ ได้นำข้อมูลลงไปแสดงผลบนเว็บแอปพลิเคชันเบื้องต้นให้ผู้ที่ใช้งานสามารถดูผลการวิเคราะห์ของข้อมูลที่ได้ ่อย่างไรก็ตาม ตัวอย่างเว็บแอปพลิเคชันนี้มีข้อจำกัดที่ขึ้นอยู่กับผู้ดูแลระบบเพียงคนเดียวที่สามารถทำการอัปเดตข้อมูลได้ ดังนั้น การรายงานผลที่เกี่ยวข้องกับปัจจัยของการลาออกกลางคันของนักศึกษาในแอปพลิเคชันนี้อาจยังขาดความสมบูรณ์ เนื่องจากผู้วิจัยยังไม่ได้ศึกษาลงในรายละเอียดถึงความต้องการของผู้ใช้ในอนาคตประกอบทางด้านกรรายงานผลที่อาจมีรายละเอียดปลีกย่อยเพิ่มเติม ซึ่งผู้วิจัยเชื่อมั่นว่างานวิจัยนี้ยังสามารถพัฒนาต่อยอดได้อีกมากในอนาคต

## บรรณานุกรม

- กฤษณะ ไวยมัย และ ธีระวัฒน์ พงษ์ศิริปรีดา. (2544). การใช้เทคนิค Association Rule Discovery เพื่อการจัดสรรกฎหมายในการพิจารณาคดีความ. NECTEC Technical Journal, 3(11), 143-152
- ชนิตาภา บุญประสม และ จรัญ แสนราช. (2561) การวิเคราะห์การทำนายการลาออกกกลางคันของ นักศึกษาระดับปริญญาตรี โดยใช้เทคนิควิธีการทำเหมืองข้อมูล. Technical Education Journal King Mongkut's University of Technology North Bangkok, 9(1), 142-151
- ณัฐธิดา สุวรรณโณ และ อรรณิกา สิงห์เอี่ยม. (2554). การหาปัจจัยที่ส่งผลต่อความเสี่ยงของนักศึกษาเรียนอ่อน ด้วย เทคนิค กฎ ความ สัม พัน ธ์ กรณี คี ก ษ า : มหาวิทยาลัยสงขลานครินทร์. วารสารวิทยาการจัดการ, 28(1), 65-79
- บุษราภรณ์ มัทธนชัย และคณะ. (2559). กฎความสัมพันธ์ของรายวิชาที่มีผลต่อการผัน สภาพนักศึกษาโดยใช้อัลกอริทึมอพริโอริ. การประชุมวิชาการ ระดับชาติมหาวิทยาลัยราชภัฏ กำแพงเพชร. การประชุมวิชาการระดับชาติ มหาวิทยาลัยราชภัฏกำแพงเพชร, 3(1), 456-469
- ปรัชญา นวนแก้ว. (2562). เทคโนโลยีเพื่องานประยุกต์ทางธุรกิจ 3 [เอกสารไม่ได้ตีพิมพ์]. คณะเทคโนโลยีสารสนเทศและการสื่อสาร, มหาวิทยาลัยพะเยา
- พนิดา ทรงรัมย์. (2555). การสร้างกฎความสัมพันธ์จากเซต รายการความถี่และเซตรายการความถี่แบบปิด. วารสารเทคโนโลยีสารสนเทศ. 8(2), 22-28
- พรพมิล ชัยวุฒศิริศักดิ์ และ ยุวดี กล่อมวิเศษ. (2562) การพัฒนานาฏการทำนายผลการเรียนของนักศึกษาชั้นปีที่ 1 โดยใช้เทคนิคการทำเหมืองข้อมูล. วารสารวิจัยรวมคำแห่ง (วิทยาศาสตร์และเทคโนโลยี), 22(2), 43-52
- วิภาวรรณ บัวทอง. (2557). Data Mining for Classification [เอกสารไม่ได้ตีพิมพ์]. คณะวิทยาศาสตร์และเทคโนโลยี, มหาวิทยาลัยราชภัฏภูเก็ต
- สุदारัตน์ พิมลรัตนกานต. (2558). การศึกษาสาเหตุการลาออกกลางคันของนักศึกษาหลักสูตรประกาศนียบัตรวิชาชีพและหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง กรณีศึกษา : วิทยาลัยเทคโนโลยีวิบูลย์บริหารธุรกิจ รามอินทรา เขตบางเขน กรุงเทพมหานคร. วารสาร

วิทยาลัยโลจิสติกส์และซัพพลายเชน Journal of Logistics and Supply Chain College,  
1(1), 67-79

Shearer, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. Journal of  
Data Warehousing, 5, 13-22.