

การวิเคราะห์แบบจำลองกฎความสัมพันธ์สำหรับข้อมูลการสั่งซื้อ

สินค้าผ่านช่องทางแอปพลิเคชันของบริษัทแห่งหนึ่ง

โดยใช้อัลกอริทึม Apriori

ANALYSIS OF THE ASSOCIATED MODEL FOR
PURCHASING ONLINE PRODUCT FOR A COMPANY

USING THE APRIORI ALGORITHM



สหกิจศึกษานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต (สถิติประยุกต์)

ภาควิชาสถิติ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ปีการศึกษา 2561

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ANALYSIS OF THE ASSOCIATED MODEL FOR
PURCHASING PRODUCT ONLINE FOR A COMPANY
USING THE APRIORI ALGORITHM



A CO-OPERATIVE EDUCATION SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENT FOR
THE DEGREE OF BACHELOR OF SCIENCE (APPLIED STATISTICS)
DEPARTMENT OF STATISTICS, FACULTY OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษา **ACADEMIC YEAR 2018** อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

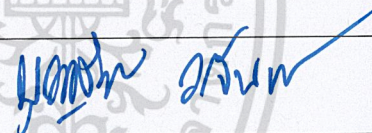

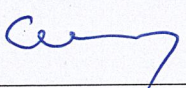
หัวข้อสหกิจศึกษา การวิเคราะห์แบบจำลองกฎความสัมพันธ์สำหรับข้อมูลการสั่งซื้อสินค้าผ่านช่องทางแอปพลิเคชันของบริษัทแห่งหนึ่งโดยใช้วิธีกฎความสัมพันธ์
Analysis of the associated model for purchasing online product for a company using the apriori algorithm

ชื่อนักศึกษา นางสาวรตพรธม ดียิ่ง รหัสนักศึกษา 58051302
นายศิริชัช มัจฉาเขียว รหัสนักศึกษา 58051319

ปริญญา วิทยาศาสตรบัณฑิต (สถิติประยุกต์)
ภาควิชา สถิติ
ปีการศึกษา 2561

อาจารย์ที่ปรึกษา อ.พรชัย หลายพสุ
พี่เลี้ยง คุณสุนิศา พูลสวัสดิ์
คุณพิษณุพันธ์ ออมสิน
คุณพัชรนันต์ คร้ามไพบุลย์

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.) อนุมัติให้สหกิจศึกษานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา วิทยาศาสตรบัณฑิต (สถิติประยุกต์) ประจำปีการศึกษา 2561

คณะกรรมการ	ลายมือชื่อ
ดร.บุญญสิทธิ์ วรจันทร์ ประธานกรรมการ	
ดร.สุกญา ศรีอินมัย กรรมการ	
คุณสุนิศา พูลสวัสดิ์ กรรมการและพี่เลี้ยง	คุณสุนิศา พูลสวัสดิ์
คุณพิษณุพันธ์ ออมสิน กรรมการและพี่เลี้ยง	คุณพิษณุพันธ์ ออมสิน
คุณพัชรนันต์ คร้ามไพบุลย์ กรรมการและพี่เลี้ยง	คุณพัชรนันต์ คร้ามไพบุลย์
อาจารย์พรชัย หลายพสุ กรรมการและอาจารย์ที่ปรึกษา	

ลิขสิทธิ์ของคณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อสหกิจศึกษา	การวิเคราะห์แบบจำลองกฎความสัมพันธ์สำหรับข้อมูลการสั่งซื้อสินค้าผ่านช่องทางแอปพลิเคชันของบริษัทแห่งหนึ่งโดยใช้วิธีกฎความสัมพันธ์	
ชื่อนักศึกษา	นางสาวรตพรรณ ดียิ่ง รหัสนักศึกษา 58051302 นายศิริชัช มัจฉาเขียว รหัสนักศึกษา 58051319	
อาจารย์ที่ปรึกษา	อ.พรชัย หลายพสุ	
ปริญญา	วิทยาศาสตร์บัณฑิต (สถิติประยุกต์)	
ภาควิชา	สถิติ	
ปีการศึกษา	2561	
อาจารย์ที่ปรึกษา พี่เลี้ยง	อ.พรชัย	หลายพสุ
	คุณสุนิศา	พลสวัสดิ์
	คุณพิษณุพันธ์	ออมสิน
	คุณพัชรนันต์	ครามไพบูลย์

บทคัดย่อ

สหกิจครั้งนี้มีวัตถุประสงค์เพื่อหาความสัมพันธ์ของการสั่งซื้อสินค้าผ่านช่องทางแอปพลิเคชันของบริษัทแห่งหนึ่ง ซึ่งทำการค้นหาสินค้าที่มีการสั่งซื้อคู่กัน โดยใช้วิธีกฎความสัมพันธ์ (Association Rules) ด้วยอัลกอริทึมเอปไรอริ (Apriori) ในโปรแกรมเวก้า (Weka) โดยวิเคราะห์จากการสั่งซื้อสินค้าของลูกค้าในแต่ละรายบุคคล ซึ่งผลลัพธ์ที่ได้จากงานวิจัยนี้ สามารถพยากรณ์ความต้องการของลูกค้าออกมาเป็นกฎความสัมพันธ์ได้ 2 รูปแบบประกอบด้วย ค่าสนับสนุนขั้นต่ำที่ 0.01, ค่าความเชื่อมั่นขั้นต่ำที่ 0.5 ได้กฎความสัมพันธ์ของสินค้าทั้งหมด 13 กฎและค่าสนับสนุนขั้นต่ำที่ 0.02, ค่าความเชื่อมั่นที่ 0.5 ได้กฎความสัมพันธ์ของสินค้าทั้งหมด 5 กฎ จากผลการวิจัยพบว่า กฎที่เหมาะสมที่สุดคือถ้ามีการสั่งซื้อน้ำดื่มขวดแก้วแล้วจะมีการสั่งซื้อน้ำดื่มขวดพลาสติก (กลาง)

ผลการศึกษาครั้งนี้สามารถนำไปใช้เป็นข้อมูลในการสนับสนุนการตัดสินใจ ในการนำเสนอสินค้าของบริษัทได้ตามความต้องการของลูกค้าในแต่ละราย เป็นการสร้างความพึงพอใจให้กับลูกค้า ซึ่งเป็นกลยุทธ์หนึ่งของการวางแผนการตลาดได้อย่างเหมาะสม และมีประสิทธิภาพ

คำสำคัญ : กฎความสัมพันธ์ เหมือนข้อมูล เอปไรอริ

Title	Analysis of the associated model for purchasing online product for a company using the apriori algorithm
Students	Miss Ratapan Deeying 58051302 Mr. Sirachat Muchachiew 58051319
Degree	Bachelor of Science (Applied Statistic)
Department	Statistic
Faculty	Science
University	King Mongkut's Institute of Technology Ladkrabang (KMITL)
Academic Year	2018
Advisor	Pornchai Laipasu
Job advisor	Sunisa Poolsawad Pisanupun Oomsin Patcharanun Krampaiboon

Abstract

The objective of this cooperative study was to analyze sales data of online purchases by individuals of a beverage company to find out which product was bought together frequently with which product. We used the association rules with Apriori algorithm in Weka software to find out which product was most frequently bought with which product. Using the support and confidence parameter thresholds of 0.01 and 0.5, respectively, the result shows 13 association rules. Meanwhile, when using the support and confidence parameter thresholds of 0.02 and 0.5, respectively, the result shows 5 association rules. One of the main results was that drinking water in glass bottles was frequently bought with drinking water in plastic bottles (M)

All of those findings could support the executives of the company in their decision to offer what kinds of products to their customers and in their formulation of good marketing strategies for the products.

Keywords : Apriori, Association Rules, Data mining

กิตติกรรมประกาศ

การทำเล่มสหกิจศึกษานี้ประสบความสำเร็จลุล่วงไปได้ด้วยความกรุณาเป็นอย่างยิ่งจาก อาจารย์พรชัย หลายพล อาจารย์ที่ปรึกษาสหกิจศึกษาที่กรุณาสละเวลาให้คำปรึกษาช่วยชี้แนะแนวทางในการทำเล่มสหกิจศึกษาตลอดจนแนวทางในการแก้ไขปัญหาและอุปสรรคด้านต่างๆที่เกิดขึ้นระหว่างการทำงานให้ผ่านไปได้และคำแนะนำที่เป็นประโยชน์ต่อการทำเล่มสหกิจศึกษานี้เสมอมา

ขอขอบพระคุณ ดร.บุญญสิทธิ์ วรรณจันทร์ และ ดร.สฤณา ศรีอโนมัย ที่ให้เกียรติเป็น คณะกรรมการสอบสหกิจ ให้คำปรึกษาและคำแนะนำที่เป็นประโยชน์ต่องานวิจัยเสมอมา และกราบขอบพระคุณ คุณสุนิศา พูลสวัสดิ์ คุณพิชญพันธ์ ออมสิน และคุณพัชรนันต์ คร้ามไพบุลย์ ที่ให้ความเมตตาในการตรวจสอบรวมทั้งให้คำปรึกษาเพื่อแก้ไขเครื่องมือที่ใช้ในการทำสหกิจให้มีคุณภาพ และขอกราบขอบพระคุณคณาจารย์ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ทุกท่านที่ได้มอบวิชาความรู้และประสบการณ์การเรียนการสอนกับผู้จัดทำ รวมทั้งให้คำแนะนำที่เป็นประโยชน์ต่องานวิจัยและขอบคุณเจ้าหน้าที่ทุกท่านที่ช่วยดำเนินการเรื่องต่างๆ ให้เป็นอย่างดี

ขอขอบคุณครอบครัวที่คอยให้กำลังใจ และสนับสนุนผู้วิจัยในทุกๆด้านตลอดระยะเวลา การศึกษาจนจนสำเร็จการศึกษา รวมถึงขอขอบคุณเพื่อนนักศึกษาและพี่ร่วมงานที่ให้คำปรึกษาให้กำลังใจในการทำสหกิจศึกษาครั้งนี้

สุดท้ายนี้ขอขอบคุณบริษัทกรณีศึกษาที่ได้อนุญาตให้นำข้อมูลมาใช้งานในการทำเล่มสหกิจศึกษานี้

รตพรรณ ดียิ่ง

ศิริชัช มัจฉาเขียว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญตาราง	ช
สารบัญรูป	ซ
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์	2
1.3 ขอบเขตการศึกษา	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	2
1.5 นิยามศัพท์เฉพาะ	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	4
2.1 นิยามเกี่ยวกับเซต	5
2.2 กฎความสัมพันธ์	6
2.2.1 Association Discovery and Sequence Detection	6
2.2.2 Market Basket Analysis	7
2.2.3 Link Analysis	7
2.2.4 Product Association	7
2.2.5 มาตรวัด	9
2.2.5.1 Support	9
2.2.5.2 Confidence	10
2.2.5.3 Lift	10
2.3 คลังข้อมูล	11
2.3.1 การทำเหมืองข้อมูล	12
2.3.1.1 กระบวนการค้นหาความรู้	13
2.3.1.2 ขั้นตอนการทำเหมืองข้อมูลแบบ CRISP-DM	14
2.3.1.3 รูปแบบการทำเหมืองข้อมูล	15
2.4 การบริหารลูกค้าสัมพันธ์	17
2.4.1 หลักการสำคัญในการบริหารลูกค้าสัมพันธ์	18

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น ยกเว้นกรณีที่มีขออนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.4.2 การบริหารจัดการแคมเปญ.....	18
2.4.3 ประโยชน์ของการบริหารลูกค้าสัมพันธ์ (CRM)	19
2.5 โปรแกรม Weka.....	20
2.5.1 การติดตั้งโปรแกรม Weka.....	21
2.6 สถิติที่เกี่ยวข้อง.....	22
2.6.1 ลักษณะข้อมูลที่ใช้ในงานวิจัย.....	23
2.6.1.1 ข้อมูลเชิงปริมาณ.....	23
2.6.1.2 ข้อมูลเชิงคุณภาพ.....	24
2.6.2 สถิติพื้นฐาน.....	24
2.7 งานวิจัยที่เกี่ยวข้อง.....	25
บทที่ 3 วิธีการดำเนินงานวิจัย.....	28
3.1 ทำความเข้าใจและศึกษาข้อมูลที่เกี่ยวข้อง.....	28
3.1.1 ทบทวนวรรณกรรมที่เกี่ยวข้อง.....	28
3.1.2 การวิเคราะห์ข้อมูลด้วยอัลกอริทึมเอโพอริ.....	29
3.2 การเก็บรวบรวมและจัดเตรียมข้อมูล.....	30
3.2.1 คัดเลือกข้อมูล (Data Preprocessing).....	30
3.2.2 จัดรูปแบบข้อมูล (Data Transformation).....	30
3.2.3 การเตรียมข้อมูล.....	30
3.3 การสร้างแบบจำลอง (Modeling).....	33
บทที่ 4 ผลการวิจัย.....	36
4.1 ขั้นตอนการนำเข้าโปรแกรม.....	36
4.2 ผลการดำเนินงาน.....	41
4.3 ผลการทดสอบเพื่อวัดประสิทธิภาพ.....	55
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	59
5.1 สรุปผลการวิจัย.....	59
5.1.1 สรุปผลตามวัตถุประสงค์.....	59
5.1.2 สรุปผลการทดลองงานวิจัย.....	59
5.2 ข้อเสนอแนะ.....	62

สารบัญ (ต่อ)

	หน้า
ภาคผนวก ก ข้อมูลการสั่งซื้อ.....	66
ภาคผนวก ข ผลการวิเคราะห์ทัศนคติความสัมพันธ์	67



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 ตัวอย่างรายการการสั่งซื้อสินค้า.....	7
2.2 ตัวอย่างรายการการสั่งซื้อสินค้าร่วมกัน.....	8
3.1 ตัวแปรที่ใช้ในงานวิจัย.....	31
3.2 ข้อมูลการสั่งซื้อสินค้า มกราคม-ธันวาคม 2018.....	32
4.1 รายชื่อสินค้า.....	38
4.2 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.01, Confidence 0.5..	43
4.3 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.02, Confidence 0.5..	44
4.4 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.01, Confidence 0.5..	45
4.5 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.02, Confidence 0.5..	46
4.6 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.01, Confidence 0.5..	47
4.7 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.02, Confidence 0.5..	48
4.8 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.01, Confidence 0.5..	49
4.9 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.02, Confidence 0.5..	50
4.10 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.01, Confidence 0.5..	51
4.11 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.02, Confidence 0.5..	52
4.12 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.01, Confidence 0.5..	53
4.13 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน Support 0.02, Confidence 0.5..	54
4.14 ผลการทดสอบตัวแบบจำลองของกฎความสัมพันธ์โดยใช้ข้อมูลลูกค้าที่ได้รับโปรโมชั่น.....	57
5.1 สรุปผลการทดลองงานวิจัย.....	60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2-1 แสดงกระบวนการค้นหาความรู้ (Knowledge Discovery from Database: KDD).....	14
2-2 กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM	15
2-3 การติดตั้งโปรแกรม Weka.....	22
2-4 การติดตั้งโปรแกรม Weka.....	22
3-1 ชุดข้อมูลวิเคราะห์หากฎความสัมพันธ์	33
3-2 ลักษณะการปรับข้อมูลเชิงปริมาณเป็นข้อมูลเชิงคุณภาพ.....	34
3-4 ลักษณะข้อมูลทั้งหมดที่ใช้ในการวิเคราะห์	34
3-5 อัลกอริทึมที่ใช้ในการทดสอบ.....	35
4-1 การทำ ETL ในโปรแกรม Power BI.....	36
4-2 การ Mapping.....	37
4-3 การแปลงข้อมูลในการเข้าโปรแกรม Weka ด้วย Association Rule	37
4-4 การนำข้อมูลเข้าโปรแกรม Weka	39
4-5 การนำเข้าข้อมูลในโปรแกรม Weka ด้วยเทคนิค Association Rule โดยใช้อัลกอริทึม Apriori.....	40
4-6 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.01, Confidence 0.5.....	43
4-7 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.02, Confidence 0.5...	44
4-8 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.01, Confidence 0.5.....	45
4-9 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.02, Confidence 0.5	46
4-10 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.01, Confidence 0.5.....	47
4-11 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.02, Confidence 0.5.....	49
4-12 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.01, Confidence 0.5.....	50
4-13 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.02, Confidence 0.5.....	52
4-14 กฎความสัมพันธ์ของสินค้าที่พบว่ามี การซื้อพร้อมกัน Support 0.01, Confidence 0.5.....	53
4-15 นำข้อมูลที่ใช้ในการทดสอบตัวแบบจำลองความสัมพันธ์เข้าสู่โปรแกรม Excel.....	56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ปัจจุบันธุรกิจการขายสินค้ามีช่องทางการจำหน่ายที่หลากหลาย ทั้งการค้าปลีก การค้าส่ง และยังมีช่องทางการค้าออนไลน์ ไม่ว่าจะเป็นผ่านช่องทางเว็บไซต์หรือแอปพลิเคชัน โดยการนำเทคโนโลยีสารสนเทศเข้ามามีบทบาทในการจัดการข้อมูล ทั้งในด้านการส่งเสริมการตลาด และการตัดสินใจดำเนินงานขององค์กรเพื่อให้เกิดประสิทธิผล อีกทั้งยังสามารถนำเทคโนโลยีสารสนเทศเข้ามาเป็นเครื่องมือในการช่วยวิเคราะห์ เพื่อหากระบวนการเพิ่มส่วนแบ่งทางการตลาดให้เหนือกว่าคู่แข่งอื่น ทำให้ธุรกิจมีการเติบโตอย่างต่อเนื่องและมั่นคงในสภาพเศรษฐกิจที่มีการเปลี่ยนแปลงอยู่ตลอดเวลา แม้ว่าเทคโนโลยีสารสนเทศมีความก้าวหน้าเพียงใด แต่อัตราการใช้ประโยชน์จากข้อมูลจำนวนมากนั้นยังน้อยมาก และยังคงมีข้อมูลที่เป็นประโยชน์ซ่อนอยู่ซึ่งยังไม่ถูกนำออกมาใช้

การนำข้อมูลที่มีปริมาณมหาศาลในเชิงธุรกิจมาวิเคราะห์และทำนายเพื่อประกอบการตัดสินใจหรือกำหนดทิศทางขององค์กร สามารถสร้างรายได้ให้เกิดประโยชน์สูงสุดจากปริมาณข้อมูลมหาศาลที่มีอยู่ในบริษัทหลากหลายรูปแบบ ทั้งข้อมูลที่มีโครงสร้างไม่ชัดเจนและข้อมูลที่มีโครงสร้างชัดเจน ตัวอย่างเช่น ข้อมูลลูกค้าที่มีการสั่งซื้อสินค้า สามารถนำข้อมูลมาระบุพฤติกรรมหรือคาดการณ์โอกาสทางการค้าได้เปรียบในทางการตลาด การทำเหมืองข้อมูล (Data Mining) หรือเรียกอีกอย่างว่า กระบวนการค้นหาความรู้ในฐานข้อมูล (Knowledge Discovery in Database) เป็นกระบวนการที่ใช้ในการค้นหารูปแบบจากข้อมูลจำนวนมากโดยอัตโนมัติ ด้วยวิธีการทางสถิติการเรียนรู้ของเครื่องมือในปัจจุบัน แต่ละองค์กรพยายามใช้กระบวนการทำเหมืองข้อมูลในการจัดการระบบสารสนเทศเพื่อสร้างความได้เปรียบให้กับองค์กร

การค้นหากฎความสัมพันธ์เป็นเทคนิคหนึ่งที่สำคัญของกระบวนการทำเหมืองข้อมูล เพื่อหาความสัมพันธ์ของข้อมูลที่เกิดขึ้นบ่อยครั้งจากข้อมูลขนาดใหญ่ ซึ่งข้อมูลที่น่ามาใช้มักเป็นประเภทข้อมูลทางธุรกรรม (Transaction Database) และทำการสกัดข้อมูลที่น่าสนใจให้ออกมาอยู่ในรูปของกฎความสัมพันธ์ ถ้าซื้อสินค้า A แล้วจะซื้อสินค้า B เขียนได้ในรูป เซตของรายการที่เป็นเหตุไปสู่เซตของรายการที่เป็นผล ซึ่งมีรากฐานมาจากการวิเคราะห์ของตะกร้าตลาด (Market Basket Analysis)

ตัวอย่างเช่น ถ้าลูกค้าที่ซื้อผ้าอ้อมส่วนใหญ่จะซื้อเบียร์ด้วย เป็นต้น (สายชล สีนสมบูรณ์ทอง, 2558)

เอกสารนี้เป็นเอกสารที่ลิขสิทธิ์และสิทธิการใช้งานเป็นของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ไม่อนุญาตให้เผยแพร่หรือใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในขั้นตอนการค้นหากฎความสัมพันธ์แบ่งออกเป็น 2 ขั้นตอน (พนิตา ทรงรัมย์, 2555) ขั้นตอนแรก คือ การสืบค้นข้อมูลของเซตรายการด้วยความถี่ (Frequent Itemset Mining) ซึ่งค้นหาเซตรายการที่มีความถี่หรือค่าสนับสนุน (Support) มากกว่าหรือเท่ากับค่าสนับสนุนขั้นต่ำ (Minimum Support) จากนั้นเซตรายการแต่ละเซตจะถูกนำไปสร้างกฎความสัมพันธ์ในขั้นตอนที่สอง ซึ่งกฎความสัมพันธ์ที่ได้จะขึ้นอยู่กับเซตของรายการความถี่ของขั้นตอนแรก

ผู้วิจัยจึงมีแนวคิดที่จะนำการค้นหากฎความสัมพันธ์มาช่วยในการตัดสินใจ เพื่อส่งเสริมการตลาดของบริษัทแห่งหนึ่ง โดยออกแบบและพัฒนาแบบจำลองข้อมูล เพื่อหาความสัมพันธ์ของการซื้อสินค้าประเภทอุปโภคบริโภคของ 6 หน่วยขาย โดยใช้วิธีกฎความสัมพันธ์ด้วยอัลกอริทึม Apriori จากข้อมูลการสั่งซื้อสินค้าของลูกค้า และนำกฎความสัมพันธ์ที่ได้มาจัดโปรโมชั่นหรือออกแคมเปญให้กับลูกค้า และสามารถนำเสนอสินค้าได้ตามความต้องการของลูกค้า เพื่อตอบสนองความพึงพอใจของลูกค้า อีกทั้งยังสามารถคาดการณ์ความต้องการสินค้าของแต่ละหน่วยขายได้อย่างเหมาะสม เพื่อวางแผนการตลาดได้อย่างมีประสิทธิภาพ

1.2 วัตถุประสงค์

1. เพื่อวิเคราะห์ความสัมพันธ์ของรายการสั่งซื้อสินค้าของลูกค้า
2. เพื่อหาแบบจำลองความสัมพันธ์ของสินค้าแต่ละชนิด

1.3 ขอบเขตของงานวิจัย

งานวิจัยนี้ได้ใช้ข้อมูลจากการสั่งซื้อสินค้าผ่านช่องทาง Application ของบริษัทแห่งหนึ่งเพื่อใช้ในการวิเคราะห์หากฎความสัมพันธ์ โดยใช้รายการสั่งซื้อสินค้าตั้งแต่เดือนมกราคม 2018 ถึง เดือนธันวาคม 2018

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ผลจากการวิเคราะห์ความสัมพันธ์ของรายการสินค้า ใช้เป็นข้อมูลสนับสนุนการตัดสินใจในการวางแผนการตลาดของบริษัท
2. ผลจากแบบจำลองความสัมพันธ์ของสินค้าแต่ละชนิดสำหรับทุกหน่วยขาย ในการวางแผนการจัดการโปรโมชั่นทางการตลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 นิยามคำศัพท์

ในงานวิจัยนี้มีการใช้คำศัพท์เฉพาะทางการศึกษาภูมิความสัมพันธ์เพื่อให้เข้าใจตรงกัน ผู้วิจัยจึงได้นิยามคำศัพท์ที่สำคัญและนิยมในงานวิจัยไว้ดังนี้

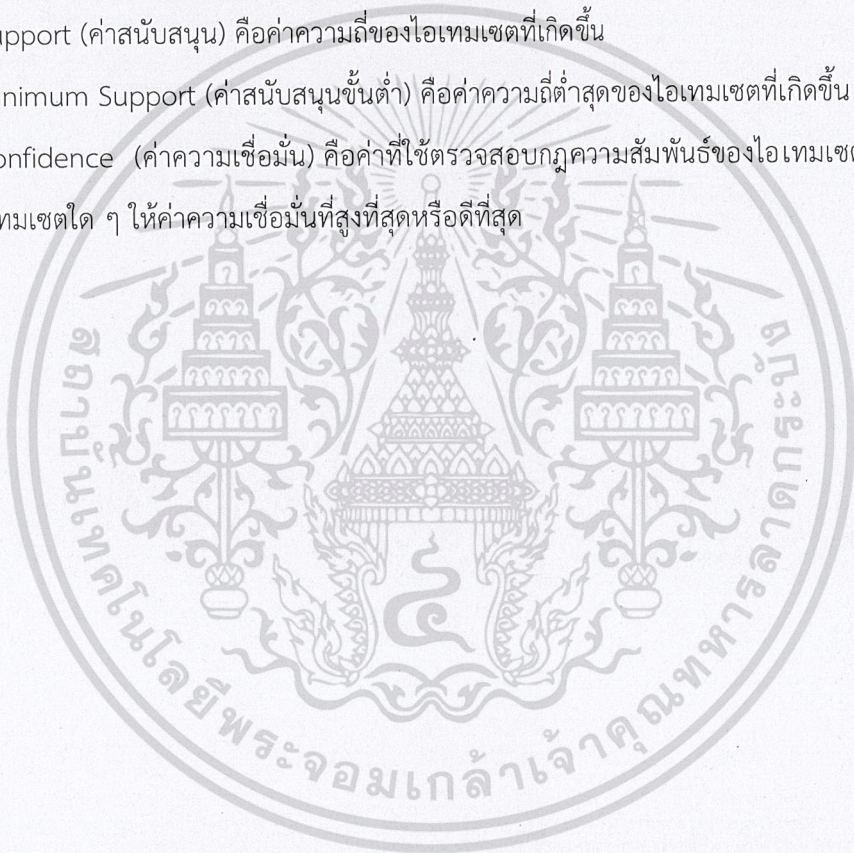
Item frequency คือ เซตของไอเทมที่เราสนใจ โดยเซตของไอเทมใด ๆ นั้นจะถูกเรียกว่า Item frequency เมื่อค่าสนับสนุนของ ไอเทมเซตมีค่ามากกว่าหรือเท่ากับค่าสนับสนุนขั้นต่ำ

Association Rules (กฎความสัมพันธ์) คือกระบวนการหนึ่งในการทำ Data Mining ใช้ในการค้นหาความสัมพันธ์ของข้อมูลตั้งแต่สองชนิดขึ้นไป

Support (ค่าสนับสนุน) คือค่าความถี่ของไอเทมเซตที่เกิดขึ้น

Minimum Support (ค่าสนับสนุนขั้นต่ำ) คือค่าความถี่ต่ำสุดของไอเทมเซตที่เกิดขึ้น

Confidence (ค่าความเชื่อมั่น) คือค่าที่ใช้ตรวจสอบกฎความสัมพันธ์ของไอเทมเซตที่เกิดขึ้น โดยที่ไอเทมเซตใด ๆ ให้ค่าความเชื่อมั่นที่สูงที่สุดหรือดีที่สุด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ธุรกิจพาณิชย์อิเล็กทรอนิกส์ (e-commerce) เป็นธุรกิจที่สามารถทำการเข้าถึงลูกค้าได้โดยตรงและรวดเร็ว ผู้ประกอบการในประเทศไทยได้เล็งเห็นถึงประโยชน์ของ e-commerce และหลายธุรกิจของไทยได้หันมาใช้ช่องทางนี้ ในการจัดจำหน่ายสินค้าและบริการ ทำให้เกิดการแข่งขันทางด้านการให้บริการ เพื่อตอบสนองความต้องการของผู้บริโภคให้ได้มากที่สุด เช่น การชำระเงินออนไลน์ การขนส่งสินค้าตามระยะเวลาที่กำหนด

สำหรับกรณีศึกษานี้เป็นหนึ่งในบริษัทของผู้ผลิตสินค้าอุปโภคบริโภค โดยผลิตสินค้าประเภทน้ำดื่ม โซดา ชาเขียว เป็นต้น ซึ่งมีสำนักงานใหญ่ตั้งอยู่ในประเทศไทย และมีสาขาย่อยในต่างประเทศทั่วโลก ทางบริษัทได้พัฒนาแอปพลิเคชันสำหรับให้บริการ เพื่อตอบสนองนโยบายและความต้องการของลูกค้า โดยที่ลูกค้าต้องมีการสั่งซื้อสินค้า เมื่อลูกค้าสั่งซื้อสินค้าและทำการยืนยันการสั่งซื้อ ระบบจะทำการตัดเงินผ่านบัตรเครดิต จากนั้นจึงทำการขนส่งสินค้าตามที่อยู่ของลูกค้า แต่เนื่องจากผู้บริหารต้องการเพิ่มยอดขาย จากสินค้าที่เหมาะสมกับหน่วยขายนั้น ดังนั้นงานวิจัยฉบับนี้จึงมุ่งเน้นไปที่การสร้างแบบจำลองกฎความสัมพันธ์ ในการค้นหาตัวแบบของของสินค้า เพื่อสนับสนุนการวางแผนทางการตลาดและเพิ่มยอดขาย รวมถึงส่วนแบ่งทางการตลาด และการจัดการสินค้าคงคลังแต่ละหน่วยให้เพียงพอต่อความต้องการของลูกค้า

ในการศึกษาวิจัยเรื่องการทำเหมืองข้อมูลด้วยวิธีกฎความสัมพันธ์ผ่านแอปพลิเคชันสั่งซื้อสินค้าของบริษัทแห่งหนึ่ง ผู้วิจัยได้ศึกษาแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง เพื่อเป็นแนวทางในการวิจัยดังต่อไปนี้

1. นิยามเกี่ยวกับเซต (Set)
2. กฎความสัมพันธ์ (Association Rule)
3. คลังข้อมูล (Data Warehouse)
4. การบริหารลูกค้าสัมพันธ์ (Customer Relation Management)
5. โปรแกรม Weka
6. สถิติที่เกี่ยวข้อง
7. งานวิจัยที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1 นิยามเกี่ยวกับเซต (Set)

เซตเป็นแนวคิดที่เป็นรากฐานสำคัญอย่างหนึ่งของคณิตศาสตร์สมัยใหม่ ซึ่งสามารถนำไปประยุกต์ใช้ได้ (อัชมา อระวีพร, 2557) และช่วยในการสร้างขั้นตอนวิธีสำหรับการค้นหาข้อมูล ที่ปรากฏร่วมกันบ่อย จึงขอกล่าวบทนิยามต่าง ๆ ที่เกี่ยวข้องพอสังเขปดังนี้

บทนิยามที่ 2.1.1 เซต (Set) คือ การรวบรวมกลุ่มวัตถุต่าง ๆ ไว้รวมกันทั้งชุดหรือลักษณะนามใช้เรียกกลุ่มของสิ่งต่าง ๆ เช่น กลุ่มของคน สัตว์ สิ่งของ ตัวเลข เป็นต้น และสิ่งต่าง ๆ ที่อยู่ในเซตเรียกว่า "สมาชิก" ของเซต แทนด้วยสัญลักษณ์ \in แทนการเป็นสมาชิกของเซตและใช้สัญลักษณ์ \notin แทนการไม่เป็นสมาชิกของเซต

บทนิยามที่ 2.1.2 เซตว่าง (Empty Set) คือ เซตเพียงหนึ่งเดียวที่ไม่มีสมาชิก หรือเรียกได้ว่าสมาชิก 0 ตัว เขียนแทนด้วยสัญลักษณ์ \emptyset หรือ $\{ \}$

บทนิยามที่ 2.1.3 กำหนดให้เซต A และ B เป็นเซตใด ๆ เซต A เป็นเซตย่อย (Subset) ของเซต B ก็ต่อเมื่อสมาชิกทุกตัวของเซต A เป็นสมาชิกอยู่ในเซต B เขียนแทนด้วยสัญลักษณ์ $A \subseteq B$ และเซต B จะเป็นซูเปอร์เซต (Superset) ของเซต A ก็ต่อเมื่อ A เป็นเซตย่อยของ B เขียนแทนด้วยสัญลักษณ์ $B \supseteq A$

บทนิยามที่ 2.1.4 กำหนดให้เซต A เป็นเซตใด ๆ เพาเวอร์เซต (Power Set) ของ A คือ เซตของเซตย่อยทั้งหมดของ A เขียนแทนด้วย $P(A)$ ถ้า A มีสมาชิก n ตัว $P(A)$ จะมีสมาชิก 2^n ตัว ยกตัวอย่างเช่น $A = \{2,3,5\}$ จะได้ $P(A) = 2^3$ ซึ่งมีสมาชิกทั้งหมด 8 ตัวได้ $\{\{2\},\{3\},\{5\},\{2,3\},\{2,5\},\{3,5\},\{2,3,5\},\emptyset\}$

บทนิยามที่ 2.1.5 กำหนดให้เซต A และ B เป็นเซตใด ๆ ผลรวมของเซต A กับเซต B คือ เซตซึ่งประกอบไปด้วยสมาชิกที่อยู่ในเซต A หรืออยู่ในเซต B เรียกว่า ยูเนียน (Union) และเขียนแทนด้วยสัญลักษณ์ $A \cup B$ โดยที่ $A \cup B = \{x / x \in A \text{ หรือ } x \in B\}$

บทนิยามที่ 2.1.6 กำหนดให้เซต A และ B เป็นเซตใด ๆ ผลร่วมของเซต A และเซต B คือเซตประกอบไปด้วยสมาชิกที่อยู่ในเซต A และอยู่ในเซต B เรียกว่า อินเตอร์เซกชัน (Intersection) เขียนแทนด้วยสัญลักษณ์ $A \cap B$ โดยที่ $A \cap B = \{x / x \in A \text{ และ } x \in B\}$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 กฎความสัมพันธ์ (Association Rule)

การค้นหากฎความสัมพันธ์ เป็นการค้นหากฎความสัมพันธ์ของคุณลักษณะ 2 คุณลักษณะ หรือมากกว่า 2 คุณลักษณะขึ้นไป (Rygielski et al., 2002; Agrawal et al., 1993; Berberidis et al., 2005) ได้มีการพัฒนาขึ้นครั้งแรกโดยนักวิจัยจากศูนย์วิจัย IBM (International Business Machines Corporation) ในประเทศสหรัฐอเมริกา และมีจุดประสงค์เพื่อการค้นหากฎความสัมพันธ์ที่น่าสนใจซึ่งซ่อนอยู่ในข้อมูล โดยการนำไปใช้ที่เห็นได้ชัดเจนสำหรับการค้นหากฎความสัมพันธ์ คือ การค้นหากฎความสัมพันธ์ของการซื้อสินค้าของลูกค้า มีการซื้อสินค้าใดบ้างร่วมกันในรถเข็นของซูเปอร์มาเกต (Market Basket Analysis) เพื่อทำความเข้าใจถึงพฤติกรรมการซื้อสินค้าของลูกค้า ตัวอย่างเช่น เมื่อลูกค้าซื้อนมแล้วจะซื้อขนมปังด้วย เป็นต้น

การค้นหากฎความสัมพันธ์ของข้อมูลในฐานข้อมูล (อนุพงศ์ สุขประเสริฐ, 2017) เป็นเทคนิคหนึ่งของการทำเหมืองข้อมูล (Data Mining) คือการค้นหากฎความสัมพันธ์ของข้อมูล จากข้อมูลขนาดใหญ่ที่มีอยู่ เพื่อนำไปหารูปแบบที่เกิดขึ้นบ่อย ใช้ในการวิเคราะห์ความสัมพันธ์หรือทำนายสิ่งที่จะเกิดขึ้น เป็นต้น

ฐานข้อมูลที่ใช้ในการทำเหมืองความสัมพันธ์ (Association Mining) เป็นฐานข้อมูลของประเภทธุรกรรม (Transaction Database) ซึ่งผลลัพธ์ที่ได้มานั้นเป็นรูปแบบของกฎความสัมพันธ์ (Association Rule) สามารถเขียนได้ในรูป เซตของรายการที่เป็นเหตุไปสู่เซตของรายการที่เป็นผล โดยมีรากฐานมาจากการวิเคราะห์ตะกร้าตลาด (Market Basket Analysis) เช่น ลูกค้าที่ซื้อผ้าอ้อมส่วนใหญ่จะซื้อเปียร์ด้วย ในส่วนของข้อมูลที่น่ามาใช้ในการวิเคราะห์นั้นจะอยู่ในรูปแบบ Nominal หรือ Ordinal เท่านั้น

2.2.1 Association Discovery and Sequence Detection

Association Discovery ใช้ในการหากฎความสัมพันธ์ที่เกิดขึ้นระหว่างสินค้า 2 ชิ้นขึ้นไป เช่น การใช้ Market Basket Analysis เพื่อวิเคราะห์ข้อมูลการสั่งซื้อสินค้า Sequence Detection ซึ่งเป็นวิธีการในทำนองเดียวกัน แต่จะใช้ลำดับของเหตุการณ์ที่เกิดขึ้นเข้ามาเกี่ยวข้องด้วย และสามารถเขียนความสัมพันธ์ของสินค้าออกมาในรูปของ $A \rightarrow B$ เรียกว่า เหตุการณ์ที่เกิดขึ้นก่อน A (Antecedent) หรือ LHS (Left-Hand Side) และเรียกเหตุการณ์ B ว่าผลของเหตุการณ์ที่เกิดขึ้นตามมา (Consequent) หรือ RHS (Right-Hand Side) ตัวอย่างเช่น “ถ้าลูกค้าซื้อค้อน แล้วจะซื้อตะปูด้วย” เหตุการณ์ที่เกิดขึ้นก่อนคือ “คนซื้อค้อน” เหตุการณ์ที่เกิดขึ้นตามมาเป็น “ซื้อตะปู”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2	01-14-2014 11:30	Eggs
3	01-15-2014 14:30	Apple
3	01-15-2014 14:30	Beer
3	01-15-2014 14:30	Cereal
3	01-15-2014 14:30	Eggs
4	01-16-2014 14:15	Beer
4	01-16-2014 14:15	Eggs

จากตารางที่ 2.1 Order Detail จะเห็นได้ว่าในหนึ่งระเบียบ (Record) เป็นการซื้อสินค้าหนึ่งชนิด ดังนั้นถ้าในหนึ่งครั้งมีการซื้อสินค้ามากกว่าหนึ่งชนิด แล้วจะมีจำนวนระเบียบมากกว่าหนึ่งระเบียบ และใช้ตัวแปร TID เป็นตัวบ่งบอกว่าเป็นการซื้อสินค้าในครั้งเดียว ซึ่งข้อมูลที่เก็บอยู่ในลักษณะเช่นนี้ ไม่เหมาะสมที่จะนำไปหาความสัมพันธ์ของสินค้า ดังนั้นจึงจำเป็นต้องทำการแปลงข้อมูลในตาราง Order Detail ให้เป็นตาราง Transaction ดังนี้

ตารางที่ 2.2 ตัวอย่างรายการการสั่งซื้อสินค้าร่วมกัน

TID	Products
1	Apple, Cereal, Diapers
2	Beer, Cereal, Eggs
3	Apple, Beer, Cereal, Eggs
4	Beer, Eggs

จากตารางที่ 2.2 Transaction มีการรวมการสั่งซื้อสินค้าในเลข TID ที่เหมือนกันให้จัดอยู่ในกลุ่มเดียวกัน เช่น TID 1 ประกอบไปด้วย Apple, Cereal, และ Diapers

ขั้นตอนการหาความสัมพันธ์ แบ่งเป็น 2 ขั้นตอน ได้แก่

ขั้นตอนที่ 1 หาความถี่ของกลุ่มสินค้า (Frequent Itemset)

เป็นการนับความถี่ของสินค้าแต่ละชนิดที่เกิดขึ้นทั้งหมด ซึ่งใช้เวลานานกว่าขั้นตอนที่ 2 และมี 2 วิธีที่นิยมใช้กันมากคือ วิธี Apriori (Agrawal and Srikant, 1994) และวิธี FP-Growth

เอกสาร (Han, Pei and Yin, 2000) ระบุการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 2 สร้างกฎความสัมพันธ์

นำความถี่ไอเทมที่ประมวลผลได้จากขั้นตอนแรก ใช้หาความสัมพันธ์ของสินค้าที่เกิดขึ้น เราจะสนใจสินค้าที่มีการซื้อบ่อย หรือสินค้าที่มีความถี่เกิดขึ้นมากกว่าค่าที่เราสนใจ (Threshold)

ความถี่ในการซื้อสินค้าแต่ละประเภทเหล่านี้เรียกว่า “ค่าสนับสนุน” (Support) เช่น Apple มีการซื้อเกิดขึ้นในครั้งที่ 1 และ 3 ดังนั้นค่าสนับสนุนของ Apple คือ 2 ครั้ง จากการซื้อทั้งหมด 4 ครั้ง ซึ่งมีค่าเท่า 50% หรือเขียนได้ว่า $Support(Apple) = 50\%$ ดังตารางที่ 2.2 และแสดงค่าสนับสนุนของสินค้าทั้งหมด

ความถี่ในการซื้อสินค้าขั้นต่ำที่เราสนใจจะเรียกว่า “ค่าสนับสนุนขั้นต่ำ” หรือ ค่าซัพพอร์ตขั้นต่ำ (Minimum Support) กล่าวคือถ้าค่าสนับสนุนมีค่ามากกว่าค่าสนับสนุนขั้นสูงสุด เราจะเรียกสินค้านั้นว่า “สินค้าที่เกิดขึ้นบ่อยครั้ง” (Frequent Itemset) ซึ่งไอเทมเซตเป็นเซตของการซื้อสินค้า เช่น {Beer, Cereal} หมายความว่าซื้อสินค้าพร้อมกัน 2 ชนิดคือ Beer และ Cereal โดยไม่สนใจลำดับว่าซื้อขึ้นไหนก่อนหรือหลัง

2.2.5 มาตรวัด (Measure)

2.2.5.1 Support

ค่าที่ใช้วัดความถี่ของการเกิดร่วมกันของเหตุการณ์สองเหตุการณ์ ว่ามีความถี่มากหรือน้อยเพียงใด โดยค่าความถี่ที่นับได้ของกลุ่มข้อมูลที่เหมือนกันในชุดข้อมูลนั้น เรียกว่าค่าสนับสนุน (Support) ซึ่งแสดงการเกิดขึ้นพร้อมกันของเหตุการณ์ A และ B คือ $A \cap B$ และจำนวนเหตุการณ์ที่เกิดขึ้นทั้งหมดคือ N

ค่าสนับสนุนขั้นต่ำ (Minimum Support) ในการกำหนดค่าสนับสนุนขั้นต่ำนี้ จะขึ้นอยู่กับผู้ใช้ระบบเป็นผู้กำหนดเองหรือคำแนะนำจากผู้เชี่ยวชาญ (Expert User) เป็นผู้กำหนดให้ก็ได้ โดยกฎความสัมพันธ์ที่ได้นั้นจะต้องมีค่าสนับสนุนไม่น้อยกว่าค่าขั้นต่ำที่ได้กำหนดไว้ข้างต้น (Sayad, S., 2019)

$$Support(A \rightarrow B) = \frac{Frequency(A,B)}{N} \quad (2.1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.5.2 Confidence

ค่าที่ใช้วัดความถี่ของการเกิดเหตุการณ์หนึ่งก่อน แล้วจะเกิดอีกเหตุการณ์หนึ่งตามมา เช่น เมื่อเกิดเหตุการณ์ A แล้วบ่อยแค่ไหนที่จะเกิดเหตุการณ์ B ซึ่งสามารถคำนวณหาค่าความเชื่อมั่นได้ ดังสมการ (Sayad, S., 2019)

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Frequency}(A, B)}{\text{Frequency}(A)} \quad (2.2)$$

2.2.5.3 Lift

ค่าที่ใช้วัดประสิทธิภาพของกฎความสัมพันธ์ โดยจะวัดอิทธิพลของความสัมพันธ์ที่เกิดขึ้น และค่า Lift ที่ได้ออกมาสามารถบอกได้ว่า ถ้าเกิดเหตุการณ์ A แล้วจะเกิดเหตุการณ์ B ในกรณี que ค่า Lift มีค่าเท่ากับ 1 หมายถึงการเกิดเหตุการณ์ A ไม่มีความสัมพันธ์กับการเกิดเหตุการณ์ B ถ้าค่า Lift มีค่ามากกว่า 1 หมายถึงการเกิดเหตุการณ์ A ส่งเสริมต่อการเกิดเหตุการณ์ B และ ถ้าค่า Lift มีค่าน้อยกว่า 1 แสดงว่าการเกิดเหตุการณ์ A ไม่ส่งเสริมต่อการเกิดเหตุการณ์ B ซึ่ง สามารถคำนวณได้ดังสมการ (Sayad, S., 2019)

$$\text{Lift}(A \rightarrow B) = \frac{\text{Support}(A \rightarrow B)}{\text{Support}(A) \times \text{Support}(B)} \quad (2.3)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 คลังข้อมูล (Data Warehouse)

คลังข้อมูล หมายถึง ฐานข้อมูลขนาดใหญ่ขององค์กรหรือหน่วยงานหนึ่ง ๆ ซึ่งมีการเก็บรวบรวมข้อมูลจากฐานระบบงานประจำวัน หรือฐานข้อมูลปฏิบัติการ หรือเรียกอีกอย่างว่า “Operational database” และฐานข้อมูลอื่นภายนอกองค์กร เรียกว่า “External database” ซึ่งข้อมูลต่าง ๆ ที่จะนำมาจัดเก็บไว้ในคลังข้อมูลจะถูกทำให้อยู่ในรูปแบบที่เป็นมาตรฐานเดียวกัน โดยมีวัตถุประสงค์เพื่อใช้วิเคราะห์ข้อมูลเพื่อทำการตัดสินใจ (Decision Making) เพื่อบริหารงานภายในองค์กร

จากแนวคิดของระบบคลังข้อมูลที่ William H. Inmon (2000) ได้นิยามไว้ดังนี้ว่า “Data Warehouse is a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management’s decision making process” ซึ่งหมายถึงคลังข้อมูลที่มีคุณสมบัติที่สำคัญ 4 ประการมีดังนี้

1. การจัดเก็บข้อมูลตามหัวข้อ (Subject Oriented) คลังข้อมูลที่มีการถูกออกแบบมาเพื่อมุ่งเน้นไปเฉพาะในส่วนของเนื้อหาที่สนใจ ช่วยสร้างสารสนเทศประกอบการตัดสินใจของผู้บริหาร และเจาะจงในแต่ละหัวข้อที่ผู้บริหารสนใจ
2. การรวบรวมข้อมูลเป็นหนึ่งเดียว (Integration) เป็นคุณลักษณะที่สำคัญที่สุดของคลังข้อมูล การ Integration หมายถึง การรวบรวมข้อมูลจากหลาย ๆ แหล่งเข้าด้วยกัน แล้วนำมาจัดเตรียมข้อมูลให้มีมาตรฐานเดียวกัน เช่น การกำหนดรหัสให้กับค่าตัวแปรต่าง ๆ เป็นต้น
3. ความไม่สูญหายของข้อมูล (Non Volatility) ข้อมูลในคลังข้อมูลจะไม่มีเปลี่ยนแปลงบ่อย โดยการแก้ไขปรับปรุงหรือเปลี่ยนแปลงนั้น จะเกิดขึ้นที่ตัวข้อมูลบนฐานข้อมูลระบบงานประจำวัน ข้อมูลที่อยู่บนคลังข้อมูลนั้นจะถูกนำเข้า (Load) จัดเก็บและเข้าถึง (Access) เท่านั้น ซึ่งการเปลี่ยนแปลงที่ข้อมูลจริง จะไม่เกิดขึ้นกับภายในคลังข้อมูลเลย
4. ความสัมพันธ์กับเวลา (Time Variant) ข้อมูลในคลังข้อมูลจะมีการถูกจัดเก็บโดยกำหนดแต่ละช่วงเวลาเอาไว้ ซึ่งสัมพันธ์กับการดำเนินธุรกิจของหน่วยธุรกิจนั้น เพื่อช่วยในการตัดสินใจด้านการบริการแต่ละช่วงเวลา และข้อมูลที่อยู่บนคลังข้อมูลแต่ละจุดนั้นจะเกี่ยวข้องกับจุดของช่วงเวลา ซึ่งข้อมูลแต่ละจุดสามารถเปรียบเทียบกันได้ตามแกนของเวลา

ข้อดีของระบบคลังข้อมูล (William H. Inmon, 2000)

1. ง่ายต่อการนำเสนอสารสนเทศ เพราะมีการจัดข้อมูลให้อยู่ในรูปแบบที่เหมาะสม และ

เอกสารข้อมูลถูกนำมาจากหลายแหล่ง การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ข้อมูลมีคุณภาพ เนื่องจากข้อมูลต่าง ๆ ในคลังข้อมูลมีการทำข้อมูลให้สมบูรณ์ (Data Cleaning) มาเรียบร้อยแล้ว

3. สามารถเข้าถึงข้อมูลได้อย่างรวดเร็ว

4. เป็นการแบ่งแยกงานประเภทสนับสนุนการตัดสินใจ (Decision Support) ออกจากงานประจำวัน

5. สามารถใช้ประโยชน์ เพื่อก่อให้เกิดการได้เปรียบคู่แข่งขั้นได้

6. ลดต้นทุน เช่น ต้นทุนในการจัดทำรายงานแต่ละครั้ง ต้นทุนการขยายโครงสร้างพื้นฐานของการประมวลผล (Computing Infrastructure) เนื่องจากมีข้อมูลส่วนกลาง เป็นต้น

ข้อดีของระบบคลังข้อมูล (William H. Inmon, 2000)

1. มีความยุ่งยากและต้องการความร่วมมือจากผู้ที่เกี่ยวข้องเป็นอย่างมากในการพัฒนา
2. ใช้เวลาในการจัดทำนาน
3. มีค่าใช้จ่ายสูง
4. ต้องมีการอบรมการใช้งานกับผู้ใช้

2.3.1 การทำเหมืองข้อมูล (Data Mining)

Data Mining คือ (Piyanut pintasiri, 2558) การค้นหาความสัมพันธ์และรูปแบบทั้งหมด (Pattern) ซึ่งมีอยู่จริงในฐานข้อมูล แต่ได้ถูกซ่อนไว้ภายในข้อมูลจำนวนมาก Data Mining จะทำการสำรวจ และวิเคราะห์อย่างอัตโนมัติหรือกึ่งอัตโนมัติในปริมาณข้อมูลจำนวนมากให้อยู่ในรูปแบบที่มีความหมาย และอยู่ในรูปของกฎ (Rule) โดยความสัมพันธ์เหล่านี้ แสดงให้เห็นถึงความรู้ต่าง ๆ ที่มีประโยชน์ในฐานข้อมูล เป็นการสังเคราะห์ข้อมูลอย่างละเอียดจากฐานข้อมูลขนาดใหญ่ หรือวิเคราะห์มาจากรายการ Transaction โดยเรียนรู้ข้อมูลจากอดีต หรือปัจจุบันผลลัพธ์ที่ได้จากการสังเคราะห์ของ Data Mining อาจจะเป็นข้อมูลแบบ Unknown, Valid หรือ Actionable ซึ่งความหมายของข้อมูลทั้ง 3 ประเภทนี้ มีดังนี้

1. ข้อมูลแบบ Unknown หมายถึง ข้อมูลที่ผู้ใช้งานไม่เคยรู้มาก่อน ไม่ชัดเจน ไม่สามารถตั้งสมมติฐานล่วงหน้าว่าจะเป็นแบบใด

2. ข้อมูลแบบ Valid เมื่อผู้ใช้งานเริ่มใช้เทคนิคของการทำเหมืองข้อมูลจะค้นพบสิ่งที่น่าสนใจตลอดเวลา แต่จะต้องพิจารณาด้วยว่าสิ่งนั้นถูกต้อง (Valid) หรือไม่ ตัวอย่างเช่น มักจะพบว่ามีความสัมพันธ์ของการซื้อสินค้า 2 อย่างเสมอ เมื่อจำนวนความหลากหลายของสินค้ามากขึ้น แต่

ไม่ได้หมายความว่าจำเป็นต้องให้ห้างสรรพสินค้า เก็บสินค้าในคลังมากขึ้น เพราะข้อมูลที่ได้อาจเกิดจาก
 เอกสารไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

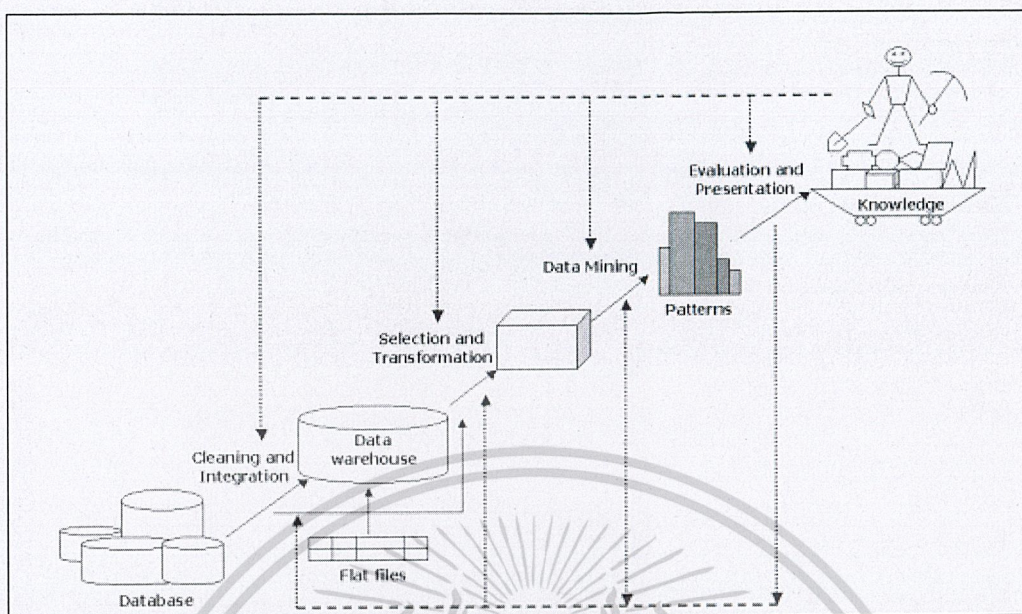
ความคลาดเคลื่อนเพราะฉะนั้นจะต้องทำการตรวจสอบความถูกต้อง (Validation and Checking) ของข้อมูลและวิเคราะห์ความถูกต้องอีกครั้ง

3. ข้อมูลแบบ Actionable ข้อมูลจะต้องถูกแปลงออกมา และนำมาตัดสินใจเพื่อสร้างความได้เปรียบในเชิงธุรกิจ บางครั้งข้อมูลที่เราค้นพบเป็นสิ่งที่คู่แข่งได้ทำไปแล้ว (เราช้าไป) หรืออาจผิดกฎหมาย ซึ่งจะต้องมีวิจรณ์ญาณในการใช้ บางครั้งข้อมูลดังกล่าวอาจไม่มีประโยชน์อะไร

2.3.1.1 กระบวนการค้นหาความรู้ (Knowledge Discovery from Database: KDD)

กระบวนการค้นหาความรู้ นั้นประกอบด้วยขั้นตอนแรก (Jiawei Han and Micheline Kamber, 2007) คือ Data cleaning เป็นขั้นตอนสำหรับการคัดข้อมูลที่ไม่มีสมบูรณ์ หรือข้อมูลที่ไม่เกี่ยวข้องออกไป (Missing Value / Noising data) เพื่อเป็นการเพิ่มคุณภาพให้ข้อมูลก่อนทำการค้นหาความรู้ ข้อมูลนั้นอาจมาจากหลายแหล่งจึงต้องทำการรวมแหล่งข้อมูลมาไว้ที่เดียวกัน เรียกว่า “Data integration” และหลังจากนั้นทำการคัดเลือกข้อมูลที่บันทึกไว้มาทำการวิเคราะห์ (Data Selection) ก่อนจะทำการวิเคราะห์ต้องแปลงข้อมูลในขั้นตอนการคัดเลือกให้เหมาะกับการทำเหมืองข้อมูลก่อน (Data transformation หรือ Data consolidation) และทำการสกัดรูปแบบที่เป็นประโยชน์จากข้อมูลทำการเตรียมไว้ (Data mining) เมื่อได้รูปแบบที่ต้องการแล้ว จึงทำการประเมินรูปแบบ (Pattern evaluation) ที่ได้จากขั้นตอนของการทำเหมืองข้อมูล เพื่อให้ได้รูปแบบที่เป็นตัวแทนความรู้ที่ต้องการ ขั้นตอนสุดท้ายคือการนำเสนอความรู้ที่ค้นหาได้ต่อผู้ใช้ (Knowledge representation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



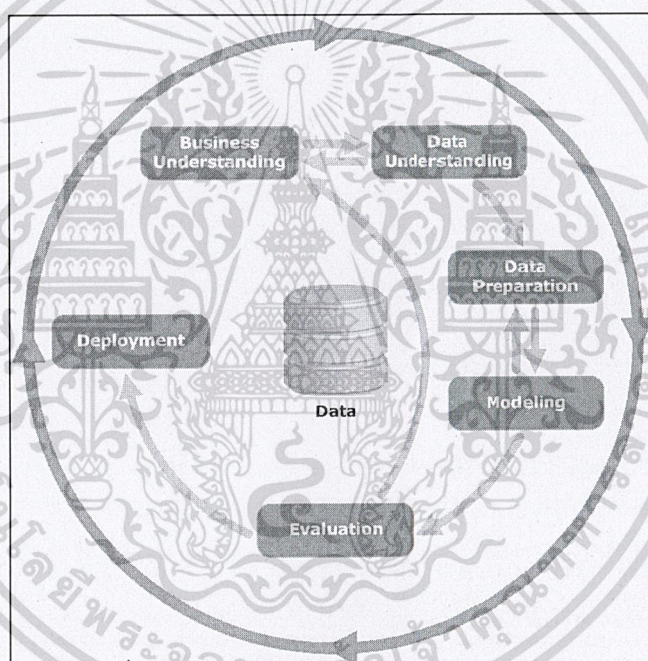
รูปที่ 2-1 แสดงกระบวนการค้นหาความรู้
(Knowledge Discovery from Database: KDD)
ที่มา: Han (2000)

2.3.1.2 ขั้นตอนการทำเหมืองข้อมูลแบบ CRISP-DM ประกอบด้วย

กระบวนการ CRISP-DM ประกอบด้วย 6 ขั้นตอน (สายชล ลินสมบูรณ์ทอง, 2558) ซึ่งต้องทำความเข้าใจ ระบุปัญหาหรือโอกาสเชิงธุรกิจ จากนั้นทำการแปลงโจทย์ที่ได้ให้อยู่ในรูปแบบที่เหมาะสมต่อการนำมาใช้วิเคราะห์ข้อมูล ซึ่งเรียกว่า การทำความเข้าใจเกี่ยวกับธุรกิจ (Business Understanding) และต้องทำความเข้าใจเกี่ยวกับข้อมูล (Data Understanding) โดยทำการรวบรวมข้อมูลที่เกี่ยวข้อง เพื่อใช้ในการวิเคราะห์ด้วยเทคนิคการทำเหมืองข้อมูล ในการรวบรวมข้อมูลนั้น ต้องพิจารณาด้วยว่าข้อมูลที่ได้นั้นต้องมาจากแหล่งข้อมูลที่ถูกต้องและน่าเชื่อถือ มีปริมาณเพียงพอต่อการนำไปใช้ในการวิเคราะห์และเป็นข้อมูลที่เหมาะสม มีรายละเอียดเพียงพอต่อการนำข้อมูลไปใช้ในการวิเคราะห์ จากนั้นเข้าสู่ขั้นตอนการเตรียมข้อมูล (Data Preparation) สามารถแบ่งออกได้เป็น การคัดเลือกข้อมูล (Data Selection) ซึ่งเกี่ยวข้องกับวัตถุประสงค์ในสิ่งที่จะทำการวิเคราะห์ นำข้อมูลที่เลือกมาทำความสะอาด (Data Cleansing) คือ การกำจัดข้อมูลที่ 'ไม่สมบูรณ์' (Missing Value/Noising Data) ออกไป เพื่อเป็นการเพิ่มคุณภาพให้กับข้อมูล ข้อมูลที่ไม่สมบูรณ์หรือข้อมูลที่ไม่ถูกต้อง อาจเกิดมาจากปัญหาในระหว่างการจัดเก็บข้อมูล เช่น

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในวงจำกัดเท่านั้น การคัดลอกข้อมูลที่ไม่ถูกต้องจากการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรือซ้ำซ้อนออกไป หรืออาจทำการซ่อมข้อมูลที่ขาดหายไปด้วยวิธีการบางอย่าง เช่นการพิจารณา ค่าเฉลี่ยจากค่าเฉลี่ยของข้อมูลส่วนใหญ่ เป็นต้น และการแปลงรูปข้อมูล (Data transformation) คือ การเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมนำไปใช้ในการวิเคราะห์ เมื่อเตรียมข้อมูลเสร็จสิ้น จึงทำการสร้างแบบจำลอง (Modeling) เป็นขั้นตอนการวิเคราะห์ข้อมูลด้วยเทคนิคการทำเหมือง ข้อมูลที่ได้เลือกไว้ แล้วทำการทดสอบแบบจำลอง (Model) ที่ได้เพื่อทำการปรับเปลี่ยนหรือหาแบบจำลองที่เหมาะสมและประเมินผล (Evaluation) การประเมินประสิทธิภาพของผลลัพธ์ จากแบบจำลองสามารถครอบคลุม และสามารถตอบสนองต่อเป้าหมายของการทำเหมืองข้อมูล ที่ได้กำหนดไว้หรือไม่ จากนั้นจึงนำแบบจำลองที่ได้ไปประยุกต์ใช้งานจริง (Deployment)



รูปที่ 2-2 กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM

ที่มา: Smart Vision Europe (2015)

2.3.1.3 รูปแบบการทำเหมืองข้อมูล

รูปแบบการทำเหมืองข้อมูลแบ่งเป็น 2 รูปแบบหลัก (สายชล สินสมบูรณ์ทอง, 2558) ดังนี้ รูปแบบที่ 1 แบบจำลองการทำนาย (Predictive Modeling) หรือเรียกอีกอย่างหนึ่งว่า การเรียนรู้แบบมีผู้สอน (Supervised learning) เป็นผลลัพธ์ที่สร้างจากการอนุมาน (Inference)

ซึ่งนำข้อมูลในอดีตมาสร้างตัวแบบ เพื่อใช้ในการทำนายประเภทตัวอย่างในอนาคตโดยข้อมูลฝึกหัด (training data) และข้อมูลทุกตัวจะมีคุณสมบัติที่ใช้ในการทำนาย อัลกอริทึมประเภทนี้จะมุ่งเน้น

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไปที่ การแบ่งแยกข้อมูลออกเป็นกลุ่มตามค่าคุณสมบัติของข้อมูล ซึ่งถ้าค่าคุณสมบัติของข้อมูลมีค่าไม่ต่อเนื่อง จะเรียกกระบวนการที่ใช้แบ่งแยกว่า การจำแนกประเภท (Classification) แต่ถ้าค่าคุณสมบัติของข้อมูลมีค่าต่อเนื่อง จะเรียกกระบวนการที่ใช้แบ่งแยกว่าการถดถอย (Regression) หรือการพยากรณ์ (Forecasting)

รูปแบบที่ 2 แบบจำลองในการบรรยาย (Descriptive Modeling) หรือเรียกว่าการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) อาจเป็นการนำข้อมูลที่มี เพื่อศึกษาหาความสัมพันธ์ต่าง ๆ (Association) หรือหาการจัดกลุ่มข้อมูล (Clustering) ซึ่งไม่ได้มีจุดมุ่งหมายเพื่อการทำนาย โดยกลยุทธ์ในการทำเหมืองข้อมูลนั้น สามารถทำได้หลายรูปแบบ ขึ้นอยู่กับลักษณะของผลลัพธ์ของการทำเหมืองข้อมูล โดยรูปแบบการทำเหมืองข้อมูลมีได้หลายแบบ ดังต่อไปนี้ (สุรพงศ์ เอื้อวัฒนามงคล, 2557)

1. การจำแนกประเภทของข้อมูล (Classification) เป็นการสร้างตัวแบบจากข้อมูลที่มีการจำแนกประเภทแล้ว เพื่อใช้ตัวแบบนั้นในการจำแนกข้อมูลที่ไม่ทราบประเภท ตัวอย่างเช่น การจำแนกความเสี่ยงของลูกค้าในการกู้ยืมเงิน การจำแนกโอกาสที่ลูกค้าจะเลิกใช้บริการ การจำแนกโอกาสที่ผู้ป่วยจะเป็นโรค เป็นต้น ตัวแบบอาจสร้างขึ้นโดยวิธีการเรียนรู้ด้วยเครื่อง (Machine Learning) หรือใช้วิธีการเชิงสถิติ เช่น Regression เป็นต้น

2. การจัดกลุ่มข้อมูล (Clustering) เป็นการแบ่งข้อมูลออกเป็นกลุ่ม โดยข้อมูลที่อยู่ในกลุ่มเดียวกัน จะมีลักษณะที่คล้ายคลึงกัน และข้อมูลที่อยู่ต่างกลุ่มกันจะมีลักษณะที่แตกต่างกัน การจัดกลุ่มข้อมูลต้องอาศัยการวัดความคล้ายคลึง หรือความแตกต่างระหว่างข้อมูล 2 ตัว ซึ่งทำได้หลายแบบ ขึ้นอยู่กับชนิดของข้อมูล และการกำหนดความคล้ายคลึงของข้อมูลตามเป้าหมายในการจัดกลุ่มข้อมูล ตัวอย่างการจัดกลุ่มข้อมูล เช่น การจัดกลุ่มลูกค้า การจัดกลุ่มของเอกสารที่มีเนื้อหาคล้ายคลึงกัน การจัดกลุ่มของยีนส์ในโครโมโซมที่ทำหน้าที่คล้ายคลึงกัน เป็นต้น

3. การสร้างกฎความสัมพันธ์ (Association Rules) เป็นการหากฎที่บอกถึงความสัมพันธ์ระหว่างข้อมูลที่มักเกิดขึ้นพร้อม ๆ กันอยู่เสมอ โดยศึกษาความสัมพันธ์ระหว่างคุณสมบัติต่าง ๆ ในข้อมูล ซึ่งปัญหาหลักขณะนี้เหล่านี้ อาจเรียกว่า การวิเคราะห์ตะกร้าสินค้า (Market Basket Analysis) ผลการวิเคราะห์คือ กลุ่มของกฎความสัมพันธ์ในลักษณะ ถ้า \rightarrow แล้ว และมีมาตรวัดคุณภาพของกฎกำกับ ตัวอย่างข้อมูลอาจเป็นรายการสินค้าที่มีการซื้อขายในแต่ละครั้ง หรือเหตุการณ์ที่เกิดพร้อมกันภายในรายการ (Transaction) กฎความสัมพันธ์ที่สร้างได้จะระบุถึง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความสัมพันธ์ว่า เมื่อพบเหตุการณ์หนึ่งหรือหลายเหตุการณ์เกิดขึ้น จะมีโอกาสสูงที่เหตุการณ์อีกอย่างหนึ่งหรืออีกหลายเหตุการณ์จะเกิดขึ้นด้วย

4. การคาดคะเน (Estimation Regression) มีลักษณะคล้ายกับการจำแนกประเภทข้อมูล คือมีการสร้างตัวแบบ เพื่อคาดคะเนค่าของตัวแปรที่ไม่ทราบค่าจากค่าตัวแปรอื่นที่ทราบค่า แต่มีข้อแตกต่างคือ ตัวแปรที่จะถูกคาดคะเนจะมีชนิดเป็นค่าต่อเนื่อง หรือเป็นตัวเลขแสดงจำนวนเต็ม ต่างกับการจำแนกประเภท ซึ่งผลลัพธ์เป็นประเภท ตัวอย่าง การประมาณปริมาณน้ำฝนที่ตกจากท้องฟ้า สภาพอากาศ การประมาณระดับคะแนนที่ผู้สอบจะทำได้จากคุณสมบัติของผู้สอบ เป็นต้น

5. การค้นหาข้อมูลที่มีความผิดปกติ (Data Anomaly Detection) เป็นการค้นหาข้อมูลที่มีลักษณะผิดปกติ หรือมีความแตกต่างจากข้อมูลส่วนใหญ่ (Outlier Data) ข้อมูลเหล่านี้มีจำนวนน้อยปะปนอยู่ในข้อมูลส่วนใหญ่ ซึ่งอาจส่งผลให้ผลการวิเคราะห์ข้อมูลมีความแม่นยำลดน้อยลง การค้นหาข้อมูลผิดปกติ อาจช่วยในงานบางอย่างที่ต้องการตรวจสอบข้อมูลที่มีความผิดปกติ เช่น การจำแนกธุรกรรมที่มีความผิดปกติ (Fraud Detection) การตรวจจับการส่งผ่านข้อมูลผิดปกติในเครือข่ายคอมพิวเตอร์ (Intruder Detection) เป็นต้น

2.4 การบริหารลูกค้าสัมพันธ์ (Customer Relation Management)

การบริหารลูกค้าสัมพันธ์นับเป็นแนวคิดที่มีความสำคัญสำหรับงานวิจัยในปัจจุบัน เนื่องจากการสร้างความสัมพันธ์กับลูกค้าโดยการใช้เทคโนโลยี และบุคลากรอย่างมีหลักการ ในปัจจุบัน CRM ได้ถูกนำมาใช้มากยิ่งขึ้นเรื่อย ๆ ซึ่งเป็นเหตุมาจากจำนวนคู่แข่งที่เพิ่มสูงขึ้นของธุรกิจ และมีการแข่งขันที่รุนแรงขึ้น ในขณะที่ลูกค้ายังคงเท่าเดิม จึงต้องพยายามหาวิธีที่สร้างความพึงพอใจให้แก่ลูกค้า นำไปสู่ความรู้สึกของลูกค้าที่มีต่อองค์กรหรือสินค้า รวมไปถึงการเข้าใจความต้องการของลูกค้า ซึ่งเป็นสิ่งสำคัญที่เป็นตัวกำหนดว่าองค์กรนั้นจะประสบความสำเร็จหรือล้มเหลว (Kotler, 1997) และ การบริหารลูกค้าสัมพันธ์ เป็นการรวมกันของกระบวนการทางธุรกิจ ที่มีการประยุกต์เพื่อใช้ตอบสนองความต้องการของลูกค้า (Gefen and Ridings, 2002)

2.4.1 หลักการสำคัญในการบริหารลูกค้าสัมพันธ์ (Kotler, 1997)

1) การเก็บรวบรวมข้อมูลของลูกค้าที่ถูกต้องและทันสมัยอยู่เสมอ มีการแยกประเภทลูกค้าจากฐานข้อมูล เนื่องจากลูกค้าแต่ละรายมีข้อมูลการซื้อขาย หรือข้อมูลส่วนบุคคลที่แตกต่างกัน

2) การใช้เทคโนโลยีที่เหมาะสม การนำเทคโนโลยีที่เพิ่มช่องทางให้ลูกค้าสามารถติดต่อกับองค์กรได้ เช่น ระบบ Call Center, Web site, ซอฟต์แวร์ที่ใช้ในการประมวลผล เทคโนโลยีต่าง ๆ เหล่านี้ เป็นสิ่งสำคัญในการผลักดันให้องค์กรสามารถสร้างผลกำไรได้ต่อเนื่อง

3) การประเมินผล เพื่อให้ทราบว่างองค์กรสามารถรักษาลูกค้าได้มากขึ้นหรือไม่ ต้องมีการปรับเปลี่ยนมาอยู่ที่การรักษาลูกค้าในระยะยาว และเพิ่มคุณค่าให้กับลูกค้ามากกว่าคุณค่าที่ลูกค้าคาดหวัง

4) การปฏิบัติเพื่อรักษาลูกค้า เนื่องจากข้อมูลในฐานข้อมูล สามารถทำให้ทราบได้ว่าลูกค้าแต่ละรายอยู่ในกลุ่มใด และลูกค้ากลุ่มใดที่ทำกำไรให้องค์กรสูงสุด หลังจากนั้นองค์กรต้องมาค้นหาวิธีเพื่อรักษาลูกค้ากลุ่มที่ทำกำไรสูงสุดไว้ เพื่อเข้าถึงการให้บริการของลูกค้าแต่ละรายอย่างเหมาะสม

2.4.2 การบริหารจัดการแคมเปญ (Campaign Management)

การบริหารจัดการแคมเปญที่เหมาะสม จะช่วยให้กลุ่มลูกค้าแต่ละกลุ่มทำกำไรให้องค์กรได้มาก สามารถใช้ข้อมูลด้านการขาย และการตลาดมาช่วยวิเคราะห์ไม่ว่าจะเป็น แคมเปญที่ทำให้ลูกค้ากลับมาใช้บริการให้มากที่สุด หรือแคมเปญที่เคยจัดขึ้นมาแล้วว่ามีผลสำเร็จมากน้อยเพียงใด ซึ่งขั้นตอนในการจัดการแคมเปญเพื่อให้มีประสิทธิภาพสูงสุด มีขั้นตอนดังต่อไปนี้

1) การวางแผน (Planning) ต้องมีการกำหนดวัตถุประสงค์ เพื่อให้ทราบกลุ่มเป้าหมาย รวมถึงจำนวนงบประมาณที่ใช้ และผลตอบแทนทางการลงทุน

2) การดำเนินการ (Action) เมื่อจัดการในส่วนของวางแผนเสร็จสิ้น จะทำการบันทึกติดตามผลของแคมเปญที่จัดขึ้น มีลูกค้าเข้ามาใช้บริการมากน้อยเพียงใด หรือกลุ่มลูกค้าเลือกใช้แคมเปญไหนสูงที่สุด เพื่อใช้เป็นแนวทางในการจัดแคมเปญในครั้งต่อไป

3) การติดตามผล (Tracking) เมื่อเสร็จสิ้นการจัดการแคมเปญ องค์กรหรือหน่วยงานจะมีการวิเคราะห์ค่าใช้จ่ายที่เกิดขึ้น มีรายจ่ายมากน้อยเพียงใดในการจัดแคมเปญนั้น ๆ สำหรับแคมเปญที่ไม่เป็นที่พึงพอใจของลูกค้า ก็ควรจะมีการระบุสาเหตุเอาไว้ เพื่อใช้ในการปรับปรุงแก้ไขแคมเปญในครั้งต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.3 ประโยชน์ของการบริหารลูกค้าสัมพันธ์ (CRM) (ชินจิตต์ แจ้งเจนกิต, 2554)

1) เพิ่มรายได้จากการขาย (Sale Revenue Increase) เพื่อให้ลูกค้าเกิดความพึงพอใจในการใช้สินค้าหรือบริการ ควรมุ่งเน้นไปที่การสร้างความสัมพันธ์ที่ดีระหว่างลูกค้ากับองค์กร การนำหลักการบริหารลูกค้าสัมพันธ์ สามารถสร้างรายได้ที่เพิ่มขึ้น รวมถึงปรับปรุงกระบวนการทำงานลดรายจ่ายและต้นทุนการหาลูกค้าใหม่ หรือดึงลูกค้ากลับมาใช้สินค้าและบริการอีกครั้ง

2) การบริหารของวงจรการทำธุรกิจของลูกค้า

- การค้นหาลูกค้าใหม่เข้าองค์กร (Customer Acquisition) สร้างความแตกต่างของสินค้าหรือการบริการที่ใหม่

- เพิ่มความพึงพอใจให้กับลูกค้า โดยผ่านขั้นตอนการทำงานที่กระชับถูกต้องและรวดเร็ว โดยผ่านหน่วยงานลูกค้าสัมพันธ์ (Customer Service)

- การรักษาฐานลูกค้า (Customer Retention) คือการใช้วิธีรับฟังความคิดเห็นจากลูกค้าและพนักงานภายในองค์กรว่ามีผลตอบรับเป็นเช่นใด เพื่อสามารถรักษาลูกค้าให้อยู่กับองค์กรได้นานที่สุดและดึงลูกค้ากลับมาใช้บริการเช่นเดิม รวมไปถึงการเสนอสินค้าหรือแคมเปญและการบริการรูปแบบใหม่

3) การเพิ่มประสิทธิภาพในกระบวนการตัดสินใจ (Improving Integration of Decision Making Process) การนำข้อมูลจากระบบต่าง ๆ มาประกอบการตัดสินใจของผู้บริหารหรือพนักงาน เช่น แนวโน้มของการใช้บริการของลูกค้า การตัดสินใจซื้อสินค้า หรือการควบคุมปริมาณของสินค้าในแต่ละช่วงจากฝ่ายสินค้าคงคลัง เป็นต้น

4) การเพิ่มประสิทธิภาพในการจัดการดำเนินงาน (Enhanced Operational Efficiency) การบริหารลูกค้าสัมพันธ์ สามารถเพิ่มประสิทธิภาพการทำงานของฝ่ายต่าง ๆ ในองค์กร โดยใช้การติดต่อสื่อสาร เช่น โทรศัพท์ อีเมล โทรสาร

- ฝ่ายขาย (Sale) ระบบเทคโนโลยีสารสนเทศที่เกี่ยวข้องกับฝ่ายขาย เช่นการขายแบบต่อเนื่องหรือต่อยอด รวมไปถึงการใช้ข้อมูลของลูกค้าเพื่อวิเคราะห์พฤติกรรมของลูกค้าและให้บริการที่เหมาะสมกับลูกค้าแต่ละกลุ่มหรือแต่ละราย

- ฝ่ายการตลาด (Marketing) สามารถวิเคราะห์ว่าวิธีใดที่ควรจัดจำหน่ายสินค้าผ่านช่องทางการขายใด รวมไปถึงการติดต่อสื่อสารกับลูกค้าวิธีใดที่เหมาะสมและสะดวกสำหรับลูกค้าแต่ละราย อีกทั้งการให้บริการของพนักงานที่เหมาะสมสำหรับลูกค้าแต่ละราย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ฝ่ายลูกค้าสัมพันธ์ (Customer Service) และฝ่ายสนับสนุน (Support) สิ่งสำคัญที่สุดคือ การดูแลลูกค้าให้มีความพึงพอใจ
- เพิ่มความรวดเร็วในการให้บริการ (Speed of Service) และสามารถปรับปรุงกระบวนการทำงานโดยมุ่งเน้นไปที่ การตอบสนองความต้องการของลูกค้าที่ถูกต้องและรวดเร็ว
- การรวบรวมรายละเอียดของลูกค้า (Gathering More Comprehensive Customer Profiles) ช่วยเพิ่มประสิทธิภาพการทำงานของฝ่ายต่าง ๆ ในองค์กรได้มาก
- การลดต้นทุนในด้านการขายและการจัดการ (Decrease General Sales and Marketing Administration Costs) เนื่องจากองค์กรมีความเข้าใจเกี่ยวกับความต้องการของลูกค้า และมีการจัดการเรื่องการสร้างความสัมพันธ์ที่ดีกับลูกค้า จึงทำให้องค์กรไม่สูญเสียต้นทุนในการดึงลูกค้ากลับมาใช้บริการ
- การสร้างมูลค่าเพิ่ม (Value Added) เนื่องจากการให้บริการหรือคุณภาพของสินค้า มีผลต่อความพึงพอใจของลูกค้า ดังนั้นผู้ประกอบการควรสร้างคุณค่าเพิ่มให้กับสินค้า และบริการทั้งในส่วนของผู้ค้า (Supply Chain) และส่วนของความต้องการของลูกค้า (Demand Chain) เพื่อให้ เกิดการบูรณาการที่ทำให้เกิดมูลค่าเพิ่มให้กับลูกค้าอย่างครบวงจร

2.5 โปรแกรม Weka

โปรแกรม Weka (Waikato Environment for Knowledge Analysis) เริ่มพัฒนามาตั้งแต่ปี 1997 โดยมหาวิทยาลัย Waikato ประเทศนิวซีแลนด์ เป็นซอฟต์แวร์สำเร็จรูป อยู่ภายใต้การควบคุมของ GPL License โปรแกรม Weka ได้ถูกพัฒนามาจากภาษาจาวาทั้งหมด ซึ่งเขียนมาโดยเน้นกับงานทางด้านการเรียนรู้ด้วยเครื่อง (Machine Learning) และการทำเหมืองข้อมูล (Data Mining) โปรแกรมจะประกอบไปด้วยโมดูลย่อย ๆ สำหรับใช้ในการจัดการข้อมูล และเป็นโปรแกรมที่สามารถใช้ Graphic User Interface (GUI) และใช้คำสั่งในการส่งค่าให้ซอฟต์แวร์ประมวลผล สามารถรัน (run) ได้หลายระบบปฏิบัติการ สามารถพัฒนาต่อยอดโปรแกรมได้ เป็นเครื่องมือที่ใช้ทำงานในด้านการทำดาตาไมนนิ่ง รวบรวมแนวคิดอัลกอริทึมมากมาย ซึ่งอัลกอริทึมสามารถเลือกใช้งานโดยตรงได้จาก 2 ทางคือ จากชุดเครื่องมือที่มีอัลกอริทึมมาให้ หรือเลือกใช้จากอัลกอริทึมที่ได้เขียนเป็นโปรแกรมลงไปเป็นชุดเครื่องมือเพิ่มเติม และชุดเครื่องมือมีฟังก์ชันสำหรับการทำงานร่วมกับข้อมูล ได้แก่ Pre-Processing, Classification, Regression,

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Clustering, Association rules, Selection และ Visualization ส่วนในเรื่องของการ Import ข้อมูล โดยข้อมูลนั้นจะต้องอยู่ในรูปแบบ ASCII อาจเป็น ARFF, CSV

ข้อดีโปรแกรม Weka (สายชล สินสมบูรณ์ทอง, 2558)

1. เป็นซอฟต์แวร์ที่เป็นฟรีแวร์
2. สามารถทำงานได้ทุกระบบปฏิบัติการ
3. เชื่อมต่อ SQL Database โดยใช้ Java Database Connectivity
4. มีลักษณะที่ง่ายต่อการใช้งานเนื่องจากใช้
5. สนับสนุนเกี่ยวกับการทำเหมืองข้อมูล (Data Mining)
6. มีการเตรียมข้อมูลและเทคนิคในการสร้างแบบจำลองที่ครอบคลุม

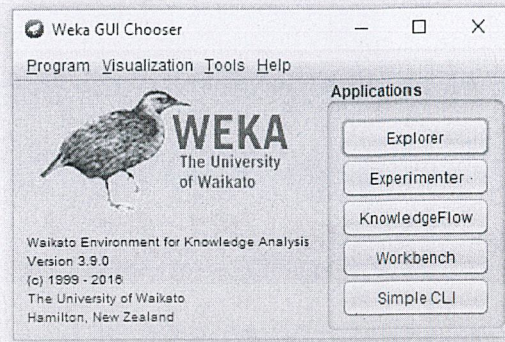
ข้อเสียโปรแกรม Weka

1. หาโหลดใช้งานได้ยาก
2. ฟังก์ชันอาจจะยังไม่ทันสมัย
3. ในฟังก์ชันบางตัวการประมวลผลของข้อมูลไม่ค่อยมีความแม่นยำ
4. การนำไฟล์ข้อมูลเข้าโปรแกรมมีความยากถ้าข้อมูลมีความผิดพลาด

การทำเหมืองข้อมูล คือ กระบวนการที่กระทำกับข้อมูล (โดยส่วนใหญ่จะมีจำนวนมาก) เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง และหลักคณิตศาสตร์

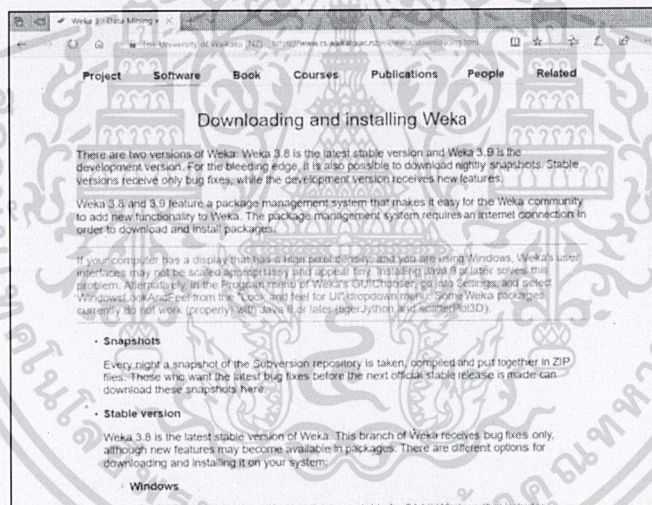
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.1 การติดตั้งโปรแกรม Weka



รูปที่ 2-3 การติดตั้งโปรแกรม Weka

1. เข้าเว็บไซต์ <https://www.cs.waikato.ac.nz/ml/weka/downloading.html> จากนั้นเลือกระบบปฏิบัติการและเวอร์ชันของโปรแกรม Weka ที่ต้องการติดตั้ง



รูปที่ 2-4 การติดตั้งโปรแกรม Weka

2. เมื่อดาวน์โหลดแล้วคลิกเลือกไฟล์ที่ดาวน์โหลด
3. ระบบจะแสดงหน้าจอการติดตั้งโปรแกรม Weka กด Next และอ่านเงื่อนไข เพื่อเลือกไฟล์ที่ต้องการการติดตั้งและกำหนดตำแหน่งที่ต้องการติดตั้งโปรแกรม
4. รอกการติดตั้ง และกดปุ่ม Finish เมื่อติดตั้งเสร็จสิ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 สถิติที่เกี่ยวข้อง

สถิติที่ใช้ในงานวิจัยมี 2 ประเภท คือ (กนกกรรณ ลีโรจนประภา, 2558)

สถิติเชิงพรรณนา (Descriptive statistics) เป็นสถิติที่ใช้ในการสรุปผลของข้อมูลที่ได้มาจากกลุ่มตัวอย่าง โดยไม่มีการอ้างอิงไปยังประชากร แต่เป็นการบรรยายลักษณะข้อมูลเท่านั้น เช่น การแจกแจงความถี่ (Frequency Distribution) เป็นการวัดแนวโน้มเข้าสู่ส่วนกลาง หรือการวัดการกระจายของข้อมูล ฯลฯ การนำเสนอผลการวิเคราะห์ข้อมูลพรรณนาจะอยู่ในรูปของตาราง (Table) และแผนภูมิ (Chart) ชนิดต่าง ๆ

สถิติอ้างอิง (Inferential statistics) หรือสถิติอนุมาน เป็นสถิติที่ใช้เพื่อนำผลสรุปที่คำนวณได้จากการสุ่มตัวอย่าง อธิบายหรือสรุปลักษณะของประชากรทั้งหมด วิธีที่ใช้ในการสรุปอ้างอิงไปยังกลุ่มประชากรนั้น คือ การประมาณค่า (Estimation) และการทดสอบสมมติฐาน (Hypothesis testing) การวิเคราะห์ค่าของความแปรปรวน (Analysis Of Variance) การวิเคราะห์ความถดถอย และสหสัมพันธ์ (Regression And Correlation Analysis) สถิติอ้างอิงจำแนกเป็น 2 ชนิดคือ แบบอ้างอิงพารามิเตอร์ (Parametric statistics) (มีการทดสอบสมมติฐานโดยใช้สถิติ T-test, Z-test, ANOVA, Regression Analysis ตัวแปรที่ต้องการวัดเป็น Interval Scale กลุ่มตัวอย่างจะต้องมีการแจกแจงเป็นโค้งปกติ กลุ่มประชากรจะต้องมีความแปรปรวนเท่ากัน) และแบบไม่อ้างอิงพารามิเตอร์ (Nonparametric statistics) (ใช้สถิติ Chi-Square, Median test, Sign test กลุ่มตัวอย่างเป็น Free Distribution ซึ่งเป็นกลุ่มตัวอย่างขนาดเล็ก ไม่ทราบลักษณะการแจกแจงของประชากรที่สนใจจะศึกษา)

2.6.1 ลักษณะข้อมูลที่ใช้ในงานวิจัย

2.6.1.1 ข้อมูลเชิงปริมาณ (Quantitative data)

ข้อมูลเชิงปริมาณ (กนกกรรณ ลีโรจนประภา, 2558) หมายถึงข้อมูลที่อยู่ในรูปตัวเลข (Numerical data) ที่แสดงถึงปริมาณ อาจเป็นค่าที่ไม่ต่อเนื่อง (Discrete) คือค่าที่เป็นจำนวนเต็มหรือจำนวนนับ เช่น จำนวน รถยนต์ในกรุงเทพมหานคร จำนวนบุตรในครอบครัว เป็นต้น หรือเป็นค่าที่ต่อเนื่อง (Continuous) คือค่าที่มีจุดทศนิยมได้ เช่น ความสูง น้ำหนัก อายุ อัตราเงินเฟ้อ สถิติน้ำฝนในปีต่าง ๆ เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.1.2 ข้อมูลเชิงคุณภาพ (Qualitative data)

ข้อมูลเชิงคุณภาพ หมายถึงข้อมูลที่แสดงถึงสถานภาพ คุณลักษณะ หรือคุณสมบัติ เช่น เพศ เชื้อชาติ สถานภาพสมรส ศาสนา กลุ่มเลือด เป็นต้น หรือจำแนกตัวแปรตามระดับการวัดได้แก่

1. นามบัญญัติ (Nominal scale) เป็นการจำแนกความแตกต่างของสิ่งที่ต้องการวัดซึ่งแบ่งออกเป็นกลุ่ม เช่น 1=ชาย 2=หญิง ตัวเลขไม่สามารถนำมาบวก ลบ คูณ หาร ได้
2. เรียงอันดับ (Ordinal scale) คือการใช้ในการจัดอันดับของสิ่งของหรือตำแหน่งของสิ่งที่ต้องการวัด เช่น ระดับการศึกษา ผลการเรียน ความเก่ง ตัวเลขอันดับที่แตกต่างกันไม่สามารถบ่งบอกถึงปริมาณความแตกต่างได้ เช่น ไม่สามารถบอกได้ว่าเก่งกว่ากันเท่าไร ตัวเลขสามารถนำมาบวกหรือลบกันได้
3. อันตรภาค หรือระดับช่วง (Interval scale) คือการกำหนดค่าตัวเลขโดยมีช่วงระยะห่างระหว่างตัวเลขเท่า ๆ กัน สามารถนำตัวเลขมาเปรียบเทียบกันได้ว่ามีปริมาณมากน้อยเท่าใด แต่ไม่สามารถบอกได้ว่าเป็นกี่เท่าของกันและกัน เพราะไม่มีศูนย์ที่แท้จริง เช่น คะแนนสอบ อุณหภูมิ (อุณหภูมิ 0 องศา มิได้หมายความว่าจะไม่มีความร้อน) ตัวเลขสามารถนำมาบวก ลบ คูณ หรือหารกันได้
4. อัตราส่วน (Ratio scale) สามารถกำหนดค่าตัวเลขให้กับสิ่งที่ต้องการวัด มีศูนย์แท้ เช่น น้ำหนัก ความสูง อายุ สามารถนำตัวเลขมาบวก ลบ คูณ หาร หรือหาอัตราส่วนกันได้ (เช่น ถนน 50 กิโลเมตร ยาวเป็น 2 เท่าของถนน 25 กิโลเมตร)

2.6.2 สถิติพื้นฐาน (กนกวรรณ สิริโรจนประภา, 2558)

ได้แก่ ค่าร้อยละ (Percentage) ดังสูตร

$$\text{สูตร } P = \frac{f \times 100}{n} \quad (2.5)$$

เมื่อ P แทน ค่าร้อยละ

f แทน จำนวนความถี่ที่เกิดขึ้น

n แทน จำนวนความถี่ทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าเฉลี่ย (mean)

$$\text{สูตร } \bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (2.6)$$

เมื่อ \bar{X} แทน ค่าเฉลี่ยของตัวอย่าง

X_i แทน ค่าของข้อมูลตัวที่ i

$\sum X_i$ แทน ผลรวมของ X_1, \dots, X_n

n แทน จำนวนข้อมูลทั้งหมด

ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard deviation หรือ S.D.)

$$\text{สูตร } S.D. = \sqrt{\frac{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}{n-1}} \quad (2.7)$$

เมื่อ S.D. แทน ส่วนเบี่ยงเบนมาตรฐาน

X_i แทน ค่าของข้อมูลตัวที่ i

\bar{X} แทน ค่าเฉลี่ยของตัวอย่าง

n แทน จำนวนข้อมูลทั้งหมด

2.7 งานวิจัยที่เกี่ยวข้อง

ฟูโดลระห์ ต้อมอง (2553) ได้กล่าวว่า ในปัจจุบันได้มีการพัฒนาขั้นตอนวิธีสำหรับการค้นหา กลุ่มข้อมูลที่ปรากฏร่วมกันบ่อยเพื่อใช้สำหรับการสร้างความสัมพันธ์ขึ้นมาอย่างมากมาย โดยแต่ ละขั้นตอนวิธีอาจคำนึงถึงการลดเนื้อที่หน่วยความจำที่ใช้ในการจัดเก็บข้อมูลในระหว่างการทำงาน การลดระยะเวลาในการทำงานให้มีความรวดเร็ว การสร้างขั้นตอนวิธีที่เหมาะสมกับรูปแบบของ ผลลัพธ์ที่ต้องการจากฐานข้อมูลหรือการสร้างขั้นตอนวิธีที่เหมาะสมกับลักษณะของข้อมูลใน ฐานข้อมูล ซึ่งขั้นตอนวิธี AMFIST เป็นขั้นตอนวิธีที่มีประสิทธิภาพในการค้นหากฎข้อมูลที่ปรากฏ ร่วมกันบ่อยสำหรับรายการข้อมูลที่มีความคล้ายคลึงกันในฐานข้อมูล และผลจากการวิเคราะห์และผล การทดลองจะเห็นได้ว่าขั้นตอนวิธี AMFIST มีประสิทธิภาพการทำงานที่ดีกว่าขั้นตอนวิธี Apriori ขั้นตอนวิธี Modified-Apriori และขั้นตอน Index-BitTable

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองจะเห็นได้ว่าขั้นตอนวิธี AMFIST มีประสิทธิภาพการทำงานที่ดีกว่าขั้นตอนวิธี Apriori ขั้นตอนวิธี Modified-Apriori และขั้นตอน Index-BitTable

บัณฑิตวิภา ทิพย์จักษ์ (2555) ได้กล่าวว่า ระหว่างอัลกอริทึม Apriori และอัลกอริทึม FP-Growth มีการกำหนดค่าสนับสนุน (Support Value) และค่าความเชื่อมั่น (Confidence Value) ซึ่งเป็นพารามิเตอร์ที่ใช้วัดประสิทธิภาพของกฎความสัมพันธ์ให้มีค่าเท่ากัน อัลกอริทึม Apriori เป็นการสร้างกฎความสัมพันธ์ด้วยการค้นหาความถี่ของไอเทมเซตที่เกิดขึ้น โดยอาศัยการทำซ้ำซึ่งใกล้เคียงกับการค้นหาแบบ Level wise โดยในการหาไอเทมเซตแต่ละครั้งนั้นต้องค้นหาข้อมูลทั้งหมดที่เกิดขึ้น ส่วนอัลกอริทึม FP-Growth เป็นการสร้างกฎความสัมพันธ์ด้วยการสร้างไอเทมเซตที่เกิดขึ้นพร้อมกันบ่อยครั้ง โดยไม่จำเป็นต้องสร้างจากไอเทมเซตที่เกิดขึ้นทั้งหมด นอกจากนี้ขั้นตอนวิธี FP-Growth ได้ใช้รูปแบบของ FP-Tree เพื่อหลีกเลี่ยงการอ่านข้อมูลจากฐานข้อมูลซ้ำกันหลายรอบ ข้อมูลที่ใช้ในการศึกษาคือ ข้อมูลเกี่ยวกับการซื้อสินค้าของลูกค้าในซูเปอร์มาเก็ต ผลการวิจัยพบว่า จำนวนกฎของทั้งสองอัลกอริทึมไม่แตกต่างกันและเวลาที่ใช้ในการหากฎความสัมพันธ์อัลกอริทึม FP-Growth ทำงานได้ดีกว่าอัลกอริทึม Apriori

ไพชยนต์ คงไชย (2554) ได้กล่าวว่า การหากฎความสัมพันธ์นั้นเป็นการความสัมพันธ์จากข้อมูลขนาดใหญ่ เพื่อให้ได้รูปแบบความสัมพันธ์ของข้อมูลเพื่อกระตุ้นยอดขายของสินค้ารวมไปถึงการวางแผนจัดรายการแนะนำสินค้า แต่ในการหาความสัมพันธ์ในปัจจุบันมีการใช้ระยะเวลาในการประมวลผลข้อมูลค่อนข้างมากเพราะกฎที่ได้จากการวิเคราะห์ความสัมพันธ์นั้นมีจำนวนมากและบางกฎที่ได้มานั้นไม่มีประโยชน์ จึงเสนอแนวทางแก้ไขปัญหาความล่าช้าของการประมวลผลข้อมูล โดยใช้วิธีการเพิ่มเงื่อนไขข้อจำกัด (Constraint) เข้าไปในการวิเคราะห์หาความสัมพันธ์ เพื่อให้ผู้ใช้ (User) สามารถระบุกฎที่ปรากฏเฉพาะไอเทม (Item) ที่ต้องการ กำหนดความยาวของกฎ และยังสามารถกำหนดไอเทมที่เป็นเป้าหมายได้ จากผลการวิจัย ถ้าไม่ใช้เงื่อนไขเพิ่มเติมนอกเหนือจากการระบุ Minimum Support และ Minimum Confidence ในการค้นหาความสัมพันธ์ ทั้งสามโปรแกรมจะให้ผลลัพธ์เป็นกฎความสัมพันธ์ที่มีจำนวนกฎที่เท่ากันและเวลาในการประมวลผลไม่แตกต่างกันมาก แต่ถ้าหากใช้เงื่อนไขกำหนดขนาดของสมาชิกโปรแกรม Apriori จะไม่สามารถกำหนดขนาดของกฎได้และโปรแกรม ACIF จะสามารถกำหนดขนาดได้ที่เงื่อนไขสมาชิกไม่เกิน 2 ชั้น เมื่อกำหนดขนาด

เอกสารนี้ถูกกักกั้นโปรแกรมไม่สามารถค้นหากฎได้ การศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สิทธิชัย วรโชติกำจร (2015) ทำการวิจัยเกี่ยวกับการทำเหมืองข้อมูลมาวิเคราะห์พฤติกรรมของผู้บริโภคโดยเน้นที่ผลิตภัณฑ์ 3 กลุ่มหลักคือ กลุ่มอาหารเสริม กลุ่มเครื่องสำอาง และกลุ่มอุปโภคบริโภค เพื่อหาความสัมพันธ์ของการเลือกซื้อผลิตภัณฑ์ของสมาชิก ซึ่งประกอบไปด้วยรายการการสั่งซื้อ และรายการสินค้า โดยใช้วิธีอัลกอริทึม Apriori ซึ่งพบกฎความสัมพันธ์ทั้งสิ้น 6 กฎ มีความเชื่อมั่น 61, 60, 59, 58, 54, 50 เปอร์เซนต์ ตามลำดับ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วิธีการดำเนินงานวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองกฎความสัมพันธ์และแก้ไขปัญหขององค์กร เพื่อเพิ่มยอดขาย เป็นการวิจัยเชิงทดลอง (Experimental Research) โดยนำข้อมูลจากฐานข้อมูลธุรกรรมการสั่งซื้อสินค้าผ่านแอปพลิเคชันขององค์กรแห่งหนึ่ง ซึ่งได้มีการจัดเก็บข้อมูลไว้แล้วเป็นข้อมูลประเภททุติยภูมิ (Secondary Data) มาทำการวิเคราะห์หาความสัมพันธ์ของสินค้าแต่ละชนิด เพื่อช่วยสนับสนุนการตัดสินใจในการจัดโปรโมชั่นให้แก่ลูกค้า และการวางแผนการตลาดของบริหารในอนาคตต่อไป โดยนำแบบจำลองที่ได้จากการวิเคราะห์ด้วยวิธีกฎความสัมพันธ์มาจัดโปรโมชั่นให้กับลูกค้าแต่ละหน่วยขายให้เหมาะสมทั้งหมด 6 หน่วย นำเสนอสินค้าได้ตามความต้องการของลูกค้าแต่ละคน เพื่อเป็นการตอบสนองความพึงพอใจของลูกค้าอีกทั้งยังสามารถคาดการณ์ความต้องการของสินค้าได้อย่างเหมาะสม โดยมีวิธีการทำเหมืองข้อมูลตามขั้นตอน CRISP-DM และได้ทำการค้นคว้าเพื่อหาหลักการแนวคิดและทฤษฎีที่เกี่ยวข้องมาใช้ในการอ้างอิง โดยทำการค้นคว้าจากบทความวิชาการและงานวิจัยที่เกี่ยวข้องนอกจากนี้ยังได้เก็บข้อมูลจากการสัมภาษณ์สอบถามจากผู้ดูแลหรือผู้ที่เกี่ยวข้องกับการดูแลแอปพลิเคชันซึ่งเป็นข้อมูลประเภทปฐมภูมิ (Primary Data) โดยมีขั้นตอนการดำเนินการวิจัยดังนี้

1. ทำความเข้าใจและศึกษาข้อมูลที่เกี่ยวข้อง
2. เตรียมข้อมูลและวิเคราะห์ข้อมูล
3. วิเคราะห์และสรุปผล

3.1 ทำความเข้าใจและศึกษาข้อมูลที่เกี่ยวข้อง

ในการทำวิจัยเรื่องการวิเคราะห์หาความสัมพันธ์ของการสั่งซื้อสินค้าผ่านช่องทางแอปพลิเคชันขององค์กรแห่งหนึ่งโดยวิธีกฎความสัมพันธ์ ผู้วิจัยได้ทำการศึกษาค้นคว้าข้อมูลและความรู้ที่จะนำมาใช้ในงานวิจัย ดังนี้

3.1.1 ทำการศึกษาและทบทวนวรรณกรรมหรือผลงานที่เกี่ยวข้อง (Review Literature) เพื่อใช้เป็นแนวทางในการดำเนินงานวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2 ทำการศึกษาขั้นตอนการวิเคราะห์ข้อมูลโดยวิธีกฎความสัมพันธ์ด้วยอัลกอริทึม Apriori ในการวิเคราะห์กฎความสัมพันธ์เริ่มจากขั้นตอนการหากฎความสัมพันธ์นั้นมี 2 ขั้นตอนคือ

1. หาความถี่ของไอเทมเซตที่เกิดขึ้นทั้งหมด (Frequent itemsets)

ในขั้นตอนแรกเป็นขั้นตอนของการนับทรานแซกชัน ซึ่งทำการสร้างและตรวจสอบเซตไอเทมที่เกิดขึ้นบ่อยทีละชั้นโดยเริ่มจากเซตไอเทมที่มีจำนวนสมาชิกเท่ากับหนึ่งถ้าเซตไอเทมใดมีค่าสนับสนุนน้อยกว่าค่าสนับสนุนขั้นต่ำจะทำการตัดเซตไอเทมนั้นออกไป ไม่นำไปสร้างเซตไอเทมในขั้นต่อไปและทำเช่นนี้วนไปเรื่อย ๆ

2. หากฎความสัมพันธ์ที่เกิดขึ้น

ในการหาความสัมพันธ์ที่เกิดขึ้นนั้นจะนำความถี่ของไอเทมเซตที่ได้จากขั้นตอนแรกที่ผ่านมาเกณฑ์ค่าสนับสนุนขั้นต่ำ และค่าความเชื่อมั่นขั้นต่ำที่กำหนดโดยการหาความถี่ของไอเทมเซตด้วยอัลกอริทึมเอโพรออริ ซึ่งค่าสนับสนุนและค่าความเชื่อมั่นดังกล่าวจะมีช่วงระหว่าง 0-100% หากกฎความสัมพันธ์นั้นที่เป็นไปตามค่าสนับสนุนขั้นต่ำและค่าความเชื่อมั่นขั้นต่ำดังกล่าว จัดว่ากฎนั้นเป็นกฎที่น่าสนใจหรือกฎเข้มแข็ง (Strong Rule) แต่ในทางกลับกันหากกฎความสัมพันธ์นั้นไม่เป็นไปตามค่าสนับสนุนขั้นต่ำและค่าความเชื่อมั่นขั้นต่ำดังกล่าว จัดว่ากฎนั้นเป็นกฎที่ไม่น่าสนใจหรือกฎอ่อนแอ (Weak Rule) โดยอัลกอริทึมเอโพรออริจะใช้ k-itemsets เพื่อหา (k+1) - itemsets เริ่มต้นการทำงานจะหา 1-itemsets โดยการอ่านค่าจากฐานข้อมูลเพื่อนับจำนวนหรือความถี่ของแต่ละ items (ข้อมูลหรือรายการสินค้า) เพื่อหา items ที่ผ่านเกณฑ์ของค่าสนับสนุนขั้นต่ำ ผลลัพธ์ที่ได้ สมมติให้เป็นเซตของ L1 จากนั้นนำเซตของ L1 (เซตของ 1-items ที่ผ่านเกณฑ์ค่าสนับสนุนขั้นต่ำ) ไปหาเซตของ L2 (เซตของ 2-items ที่ผ่านเกณฑ์ค่าสนับสนุนขั้นต่ำ) และนำเซตของ L2 ไปหาเซตของ L3 (เซตของ 3-items ที่ผ่านเกณฑ์ค่าสนับสนุนขั้นต่ำ) และทำซ้ำเช่นนี้จนกระทั่งไม่สามารถหา frequent k-itemsets ได้ (วิวัฒน์ ชินนาทศิริกุล, 2014)

กฎความสัมพันธ์ที่น่าสนใจ คือ กฎความสัมพันธ์ที่มีค่าสนับสนุนมากกว่าหรือเท่ากับค่าสนับสนุนที่กำหนด และค่าความเชื่อมั่นมีค่ามากกว่าหรือเท่ากับค่าความเชื่อมั่นที่กำหนด (Minimum Confidence Threshold) ซึ่งค่านี้คำนวณได้จากสมการดังนี้

ค่าสนับสนุน หาจากสูตร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น กรุณาอย่าให้นำไปใช้ประโยชน์ด้านการค้า
 Support(A \rightarrow B) = $\frac{\text{Frequency}(A, B)}{N}$ (2.8)
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าความเชื่อมั่น หาจกสูตร

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Frequency}(A,B)}{\text{Frequency}(A)} \quad (2.9)$$

ค่า Lift หาจกสูตร

$$\text{Lift}(A \rightarrow B) = \frac{\text{Support}(A \rightarrow B)}{\text{Support}(A) \times \text{Support}(B)} \quad (2.10)$$

3. เครื่องมือหรือซอฟต์แวร์ที่นำมาใช้ในการวิเคราะห์ข้อมูลเพื่อหากฎความสัมพันธ์ ผู้วิจัยเลือกใช้ซอฟต์แวร์ Weka Version 3.9.3 เป็นโปรแกรมที่สามารถนำมาใช้หากฎความสัมพันธ์ ซึ่งแสดงค่าสนับสนุนขั้นต่ำ ค่าความเชื่อมั่น และค่า Lift ในโปรแกรม Weka ได้อีกด้วย

3.2 การเก็บรวบรวมและจัดเตรียมข้อมูล

ในงานวิจัยนี้ใช้กฎความสัมพันธ์ด้วยขั้นตอนวิธี Apriori ซึ่งเป็นอีกเทคนิคหนึ่งในการทำเหมืองข้อมูลโดยใช้เทคนิคดังกล่าวในการสร้างตัวแบบความสัมพันธ์ของสินค้า ที่ลูกค้าซื้อบ่อย ๆ แล้วนำตัวแบบที่ได้มาช่วยสนับสนุนการตัดสินใจของผู้บริหารในการส่งเสริมการตลาด และเพื่อเพิ่มโอกาสทางธุรกิจในการขายสินค้า ฐานข้อมูลที่ใช้ในการทดสอบมีข้อมูลการสั่งซื้อสินค้า รวมทั้งสิ้นจำนวน 30,760 รายการ ในแต่ละรายการจะเป็นสินค้าที่ลูกค้าสั่งซื้อ

3.2.1 คัดเลือกข้อมูล (Data Preprocessing) นำข้อมูลจากแต่ละตารางที่เกี่ยวข้องมารวมกัน โดยข้อมูลจะประกอบไปด้วยระเบียบหน่วยขายและระเบียบการสั่งซื้อสินค้า

3.2.2 จัดรูปแบบข้อมูล (Data Transformation) ทำการจัดรูปแบบข้อมูลดิบ (Raw Data) ให้กลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นต่อไป โดยใช้วิธีการ Extract Transform Load (ETL) ผู้วิจัยได้เลือกใช้ซอฟต์แวร์คือโปรแกรม Excel และ Power BI ในการแปลงข้อมูลให้สามารถนำไปวิเคราะห์ได้

3.2.3 การเตรียมข้อมูล ชุดข้อมูลที่ใช้ในงานวิจัยนี้มาจากฐานข้อมูลการสั่งซื้อสินค้า เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า (Transaction database) เป็นข้อมูลที่ยังไม่สมบูรณ์ที่จะสามารถนำไปผ่านกระบวนการทำเหมืองไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามเผยแพร่และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลได้จึงต้องมีการจัดการข้อมูล โดยทำการจัดกลุ่มของสินค้า เนื่องจากข้อมูลชนิดของสินค้ามีความหลากหลายมาก จึงต้องทำการจัดกลุ่มสินค้าเพื่อที่จะลดการกระจายตัวของข้อมูล และกำหนดตัวแปรของการซื้อสินค้า ดังตารางที่ 3.1

ตารางที่ 3.1 ตัวแปรที่ใช้ในงานวิจัย

ชื่อข้อมูล	ชนิดข้อมูล	รายละเอียด
CD	Nominal	รหัสลูกค้า
SUD	Nominal	ชื่อหน่วยขาย
OD	Nominal	หมายเลขคำสั่งซื้อ
Soda I (S)	Nominal	โซดา I ขนาดเล็ก
Soda I (L)	Nominal	โซดา I ขนาดใหญ่
Soda II (S)	Nominal	โซดา II ขนาดเล็ก
WaterPet (S)	Nominal	น้ำดื่มพลาสติกขนาดเล็ก
WaterPet (M)	Nominal	น้ำดื่มพลาสติกขนาดกลาง
WaterPet (L)	Nominal	น้ำดื่มพลาสติกขนาดใหญ่
WaterGallon	Nominal	น้ำดื่มแกลลอน
WaterGlass	Nominal	น้ำดื่มขวดแก้ว
MineralWater (S)	Nominal	น้ำแร่ขนาดเล็ก
MineralWater (L)	Nominal	น้ำแร่ขนาดใหญ่

จากตารางที่ 3.1 ฐานข้อมูลประกอบด้วย 13 ฟิล ซึ่งสามารถอธิบายรายละเอียดได้ดังนี้

1. ฟิล CD จัดเก็บหมายเลขรหัสลูกค้า
2. ฟิล SUD จัดเก็บชื่อหน่วยขาย
3. ฟิล OD จัดเก็บหมายเลขคำสั่งซื้อ
4. ฟิล Soda I (S) จัดเก็บสินค้าโซดา I ขนาดเล็ก
5. ฟิล Soda I (L) จัดเก็บสินค้าโซดา I ขนาดใหญ่
6. ฟิล Soda II (S) จัดเก็บสินค้าโซดา II ขนาดเล็ก
7. ฟิล WaterPet (S) จัดเก็บสินค้าน้ำดื่มพลาสติกขนาดเล็ก
8. ฟิล WaterPet (M) จัดเก็บสินค้าน้ำดื่มพลาสติกขนาดกลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมีเหตุเปลี่ยนแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

9. ฟิล WaterPet (L) จัดเก็บสินค้าน้ำดื่มพลาสติกขนาดใหญ่
10. ฟิล WaterGallon จัดเก็บสินค้าน้ำแกลลอน
11. ฟิล WaterGlass จัดเก็บสินค้าน้ำดื่มขวดแก้ว
12. ฟิล MineralWater (S) จัดเก็บสินค้าน้ำแร่ขนาดเล็ก
13. ฟิล MineralWater (L) จัดเก็บสินค้าน้ำแร่ขนาดใหญ่

ตารางที่ 3.2 ตัวอย่างชุดข้อมูล

	A	B	C	D	E	F	G	H	I	J	K
1	TID	MineralWater (S)	MineralWater (L)	Soda I (S)	Soda I (L)	Soda II (S)	WaterPet (S)	WaterPet (M)	WaterPet (L)	WaterGallon	WaterGlass5
2	1	0	0	0	0	0	0	0	0	0	1
3	2	0	0	0	0	1	0	0	0	0	0
4	3	0	0	0	1	0	0	0	0	0	0
5	4	0	0	0	0	1	0	1	0	0	0
6	5	0	0	0	0	0	0	0	0	0	1
7	6	0	0	0	1	0	0	0	0	0	1
8	7	0	0	0	0	0	0	0	0	0	1
9	8	0	0	0	0	0	0	0	0	1	1
10	9	0	0	0	0	0	0	0	0	1	1
11	10	0	0	0	0	0	0	0	0	1	1
12	11	0	0	0	0	1	0	1	0	0	0
13	12	0	0	0	1	0	0	0	0	0	0
14	13	0	0	0	0	0	0	0	0	0	1
15	14	0	0	0	1	0	0	0	0	0	0
16	15	0	0	0	0	0	0	0	0	0	1
17	16	0	0	0	1	0	0	0	0	0	0
18	17	0	0	0	0	0	0	0	0	0	1
19	18	0	0	0	0	0	0	0	1	0	0
20	19	0	0	0	0	0	0	0	0	0	1
21	20	0	0	0	0	0	0	0	0	0	1
22	21	0	0	0	0	0	0	0	0	0	1
23	22	0	0	0	0	0	0	0	0	0	1
24	23	0	0	0	0	0	0	1	0	0	0
25	24	0	0	0	0	0	0	0	0	0	1
26	25	0	0	0	0	0	0	0	0	0	1
27	26	0	0	0	0	0	0	0	0	0	1
28	27	0	0	0	0	0	0	0	0	0	1
29	28	0	0	0	0	0	0	0	0	0	1
30	29	0	0	0	0	0	0	0	0	0	1
31	30	0	0	0	0	0	0	0	0	0	1
32	31	0	0	0	0	0	0	1	0	0	1
33	32	0	0	0	0	0	0	0	0	0	1

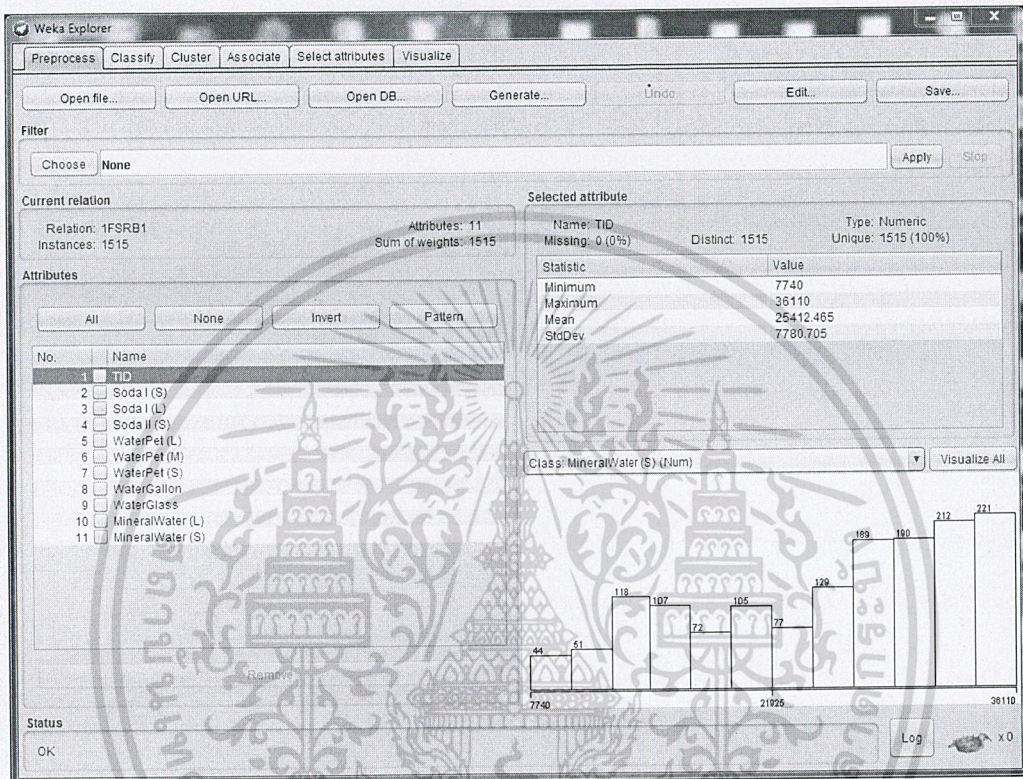
จากตารางที่ 3.2 สามารถอธิบายรายละเอียดได้ดังนี้

ในแต่ละ Transaction ที่มีการซื้อสินค้าจะกำหนดตัวแปรที่มีการซื้อจะแทนด้วย 1 และตัวแปรที่ไม่ได้ถูกซื้อจะแทนด้วย 0 มีข้อมูลการขายสินค้ารวมทั้งสิ้นจำนวน 30,760 รายการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

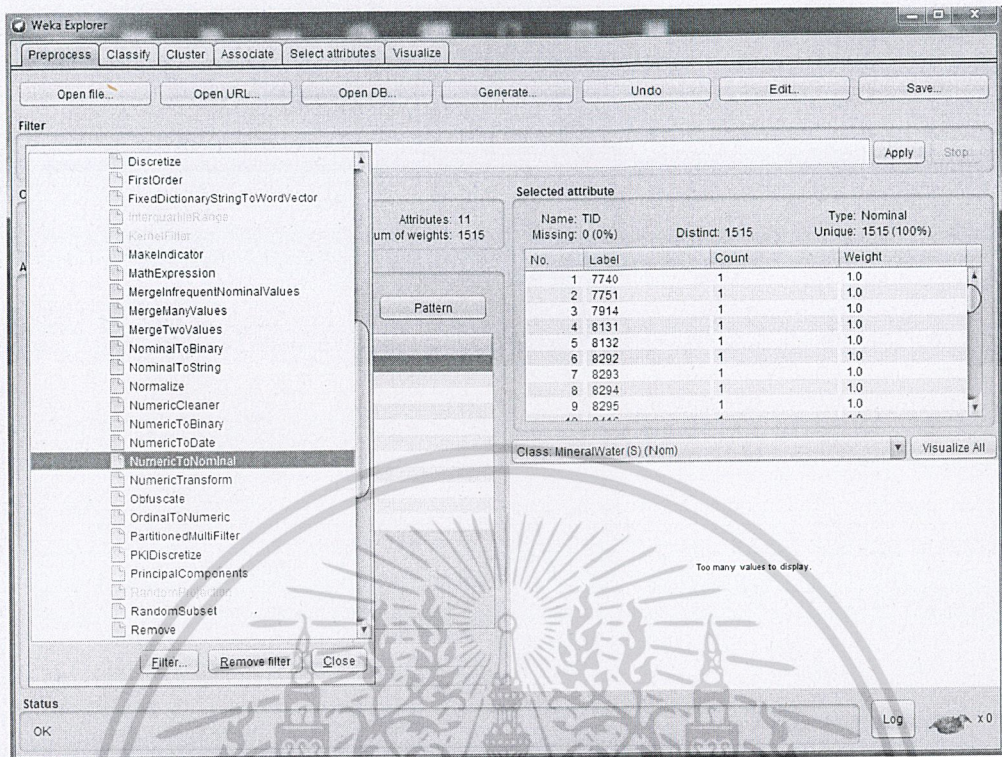
3.3 การสร้างแบบจำลอง

จากข้อมูลในตารางที่ 3.2 นำมาสร้างแบบจำลอง เมื่อนำเข้าโปรแกรม Weka แสดงข้อมูลได้ดังรูปภาพต่อไปนี้

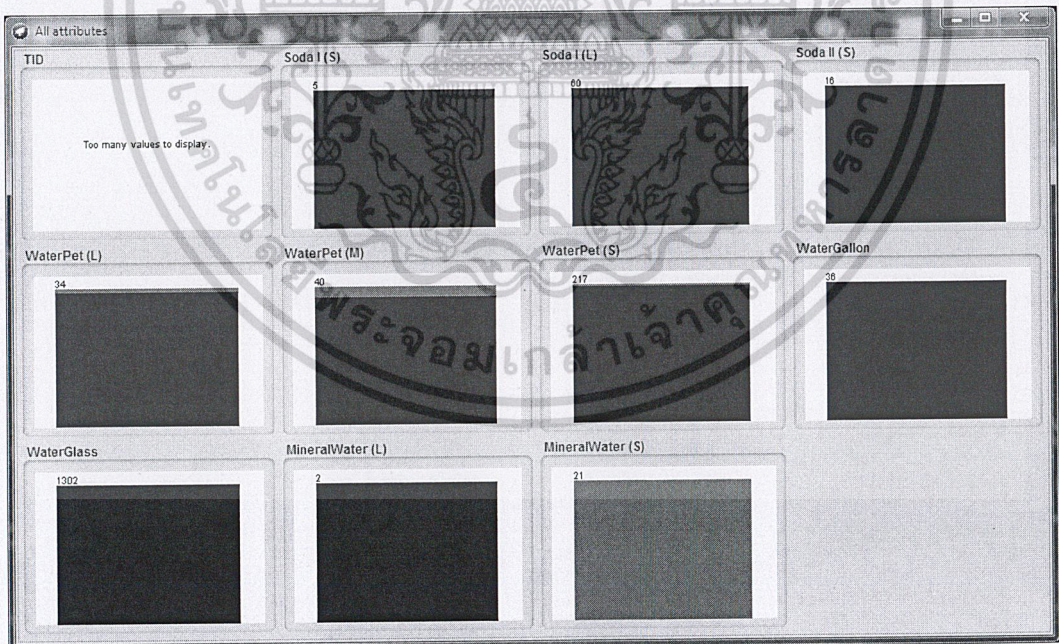


รูปที่ 3-1 แสดงชุดข้อมูลวิเคราะห์หากความสัมพันธ์จากตาราง 3.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



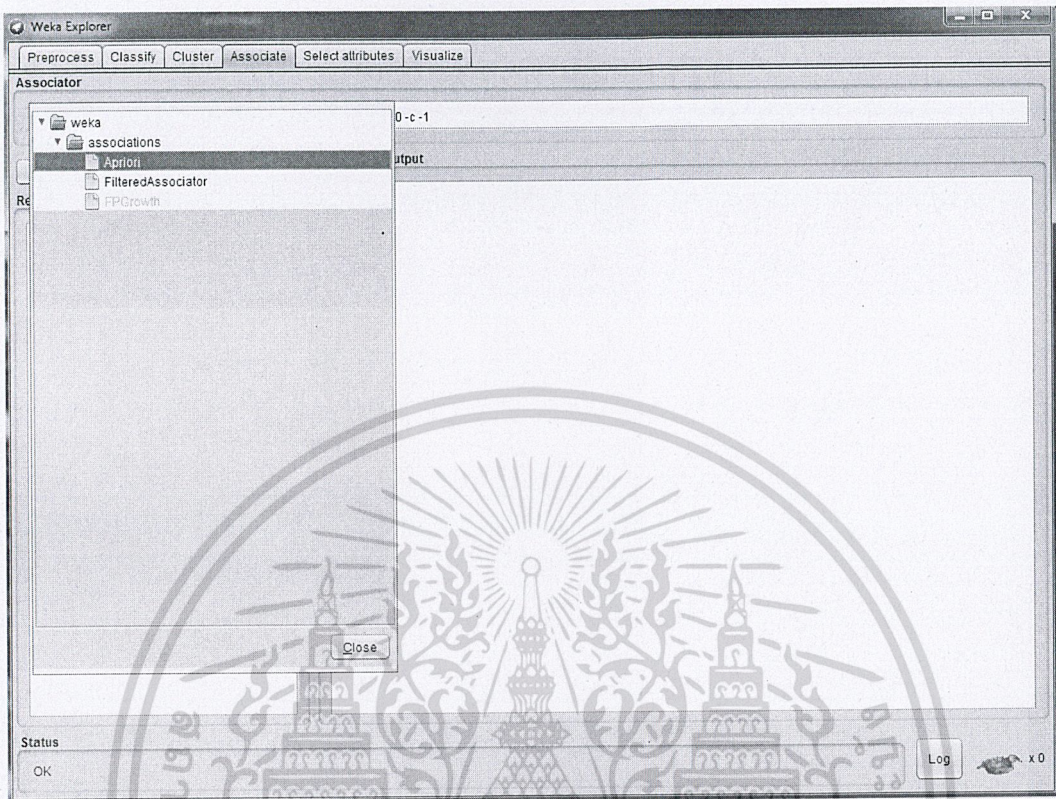
รูปที่ 3-2 แสดงลักษณะการปรับข้อมูลเชิงปริมาณเป็นข้อมูลเชิงคุณภาพ



รูปที่ 3-3 แสดงลักษณะข้อมูลทั้งหมดที่ใช้ในการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.4 การหาความสัมพันธ์โดยใช้อัลกอริทึม Apriori



รูปที่ 3-4 แสดงอัลกอริทึมที่ใช้ในการทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

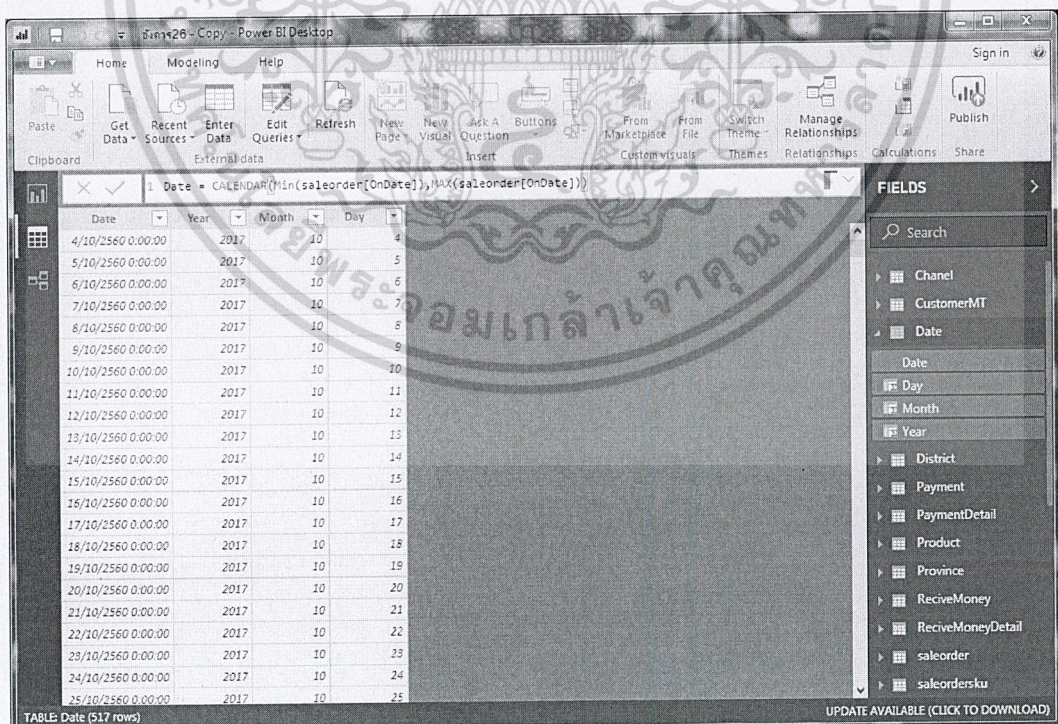
บทที่ 4

ผลการวิจัย

ผลการวิจัยการวิเคราะห์ความสัมพันธ์ของสินค้าจากรายการการสั่งซื้อของลูกค้า ซึ่งผู้วิจัยได้ทำการแบ่งผลการวิจัยออกเป็น 3 ส่วนแยกตามเทคนิคการทำเหมืองข้อมูล คือ ขั้นตอนการนำข้อมูลเข้าโปรแกรม ผลการดำเนินงาน และผลการทดสอบเพื่อวัดประสิทธิภาพ

4.1 ขั้นตอนการนำเข้าโปรแกรม

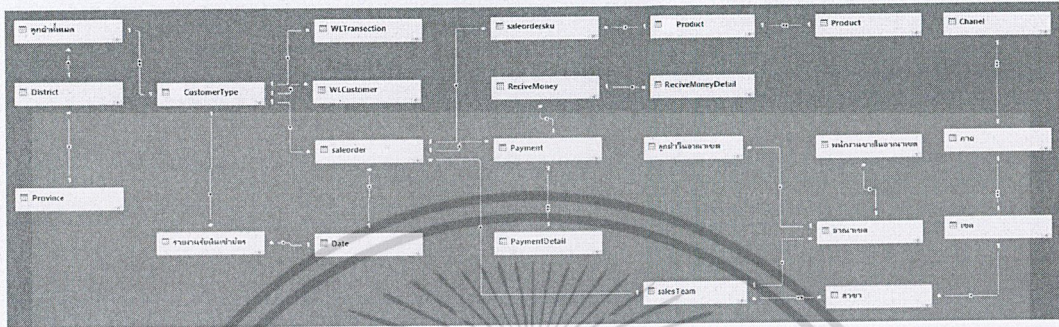
การแปลงข้อมูลจากบทก่อนหน้านี้ได้แสดงข้อมูลที่จะนำมาใช้ในการวิจัย ซึ่งข้อมูลที่ได้ประกอบด้วยข้อมูลที่มีหลากหลายรูปแบบ ซึ่งข้อมูลบางส่วนไม่สามารถนำมาใช้ในการวิเคราะห์ได้ จึงจำเป็นต้องมีการปรับปรุงแก้ไขข้อมูลเพื่อให้ข้อมูลตรงตามวัตถุประสงค์ของงานวิจัย และเลือกนำข้อมูลที่มีความเหมาะสมถูกต้องนำมาใช้ในงานวิจัยเพื่อหาผลลัพธ์คำตอบ และให้สอดคล้องกับรูปแบบอัลกอริทึม Apriori



Date	Year	Month	Day
4/10/2560 0:00:00	2017	10	4
5/10/2560 0:00:00	2017	10	5
6/10/2560 0:00:00	2017	10	6
7/10/2560 0:00:00	2017	10	7
8/10/2560 0:00:00	2017	10	8
9/10/2560 0:00:00	2017	10	9
10/10/2560 0:00:00	2017	10	10
11/10/2560 0:00:00	2017	10	11
12/10/2560 0:00:00	2017	10	12
13/10/2560 0:00:00	2017	10	13
14/10/2560 0:00:00	2017	10	14
15/10/2560 0:00:00	2017	10	15
16/10/2560 0:00:00	2017	10	16
17/10/2560 0:00:00	2017	10	17
18/10/2560 0:00:00	2017	10	18
19/10/2560 0:00:00	2017	10	19
20/10/2560 0:00:00	2017	10	20
21/10/2560 0:00:00	2017	10	21
22/10/2560 0:00:00	2017	10	22
23/10/2560 0:00:00	2017	10	23
24/10/2560 0:00:00	2017	10	24
25/10/2560 0:00:00	2017	10	25

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ 4-1 การทำ ETL ในโปรแกรม Power BI เท่านั้น ไม่ให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4-1 แสดงการทำ ETL หรือ Extract-Transform-Load เป็นการจัดหรือแปลงข้อมูลให้อยู่ในรูปแบบเดียวกัน รวมถึงการแก้ไขข้อมูลที่เกิดการผิดพลาดหรือไม่สมบูรณ์ ซึ่งเป็นการเพิ่มมูลค่าให้ข้อมูลเพื่อให้สามารถนำข้อมูลนั้นมาวิเคราะห์ได้



รูปที่ 4-2 Data Mapping

จากรูปที่ 4-2 แสดงตารางทั้งหมดที่ใช้ในการทำ ETL ซึ่ง data mapping เป็นส่วนหนึ่งของขั้นตอนการทำ ETL คือขั้นตอน Transform เป็นการนำข้อมูลที่ได้อาจจากการ Extract มาจัดรูปแบบให้ถูกต้องสอดคล้องกัน หรือนำข้อมูลที่มีความหมายเดียวกันแต่อยู่ในรูปแบบที่ต่างกั้มาจัดให้อยู่ในรูปแบบเดียวกัน

	A	B	C	D	E	F	G	H	I	J	K
1	TID	MineraWater (S)	MineraWater (L)	Soda I (S)	Soda I (L)	Soda II (S)	WaterPet (S)	WaterPet (M)	WaterPet (L)	WaterGalon	WaterGlass5
2	1	0	0	0	0	0	0	0	0	0	1
3	2	0	0	0	0	1	0	0	0	0	0
4	3	0	0	0	1	0	0	0	0	0	0
5	4	0	0	0	0	1	0	1	0	0	0
6	5	0	0	0	0	0	0	0	0	0	1
7	6	0	0	0	1	0	0	0	0	0	1
8	7	0	0	0	0	0	0	0	0	1	1
9	8	0	0	0	0	0	0	0	0	1	1
10	9	0	0	0	0	0	0	0	0	1	1
11	10	0	0	0	0	0	0	0	0	1	1
12	11	0	0	0	0	1	0	1	0	0	0
13	12	0	0	0	0	1	0	0	0	0	0
14	13	0	0	0	0	0	0	0	0	0	1
15	14	0	0	0	0	1	0	0	0	0	0
16	15	0	0	0	0	0	0	0	0	0	1
17	16	0	0	0	1	0	0	0	0	0	0
18	17	0	0	0	0	0	0	0	1	0	0
19	18	0	0	0	0	0	0	0	0	0	1
20	19	0	0	0	0	0	0	0	0	0	1
21	20	0	0	0	0	0	0	0	0	0	1
22	21	0	0	0	0	0	0	0	0	0	1
23	22	0	0	0	0	0	0	0	0	0	0
24	23	0	0	0	0	0	0	1	0	0	1
25	24	0	0	0	0	0	0	0	0	0	1
26	25	0	0	0	0	0	0	0	0	0	1
27	26	0	0	0	0	0	0	0	0	0	1
28	27	0	0	0	0	0	0	0	0	0	1
29	28	0	0	0	0	0	0	0	0	0	1
30	29	0	0	0	0	0	0	0	0	0	1
31	30	0	0	0	0	0	0	0	0	0	1
32	31	0	0	0	0	0	0	1	0	0	1
33	32	0	0	0	0	0	0	0	0	0	1

เอกสารนี้เป็นเอกสารรูปที่ 4-3 การแปลงข้อมูลในการเข้าโปรแกรม Weka ด้วย Association Rule ขนด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4-3 เป็นการแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมในการนำเข้าโปรแกรม Weka โดยเทคนิค Association Rule ด้วยอัลกอริทึม Apriori ประกอบด้วยข้อมูล 4 ส่วนได้แก่

1. ข้อมูลรายการการสั่งซื้อของลูกค้า
2. ข้อมูลผลิตภัณฑ์สินค้าที่วางจำหน่าย
3. ข้อมูลหน่วยขาย
4. ข้อมูลการขาย

โดยค่าข้อมูลนำมาใช้ในการวิเคราะห์ประกอบด้วย 2 รูปแบบคือ

ค่า 1 หมายถึง ลูกค้ามีการสั่งซื้อสินค้าชนิดนั้น

ค่า 0 หมายถึง ลูกค้าไม่มีการสั่งซื้อสินค้าชนิดนั้น

ตารางที่ 4.1 รายชื่อสินค้า

ชื่อตัวแปรสินค้า	คำอธิบาย
Soda I (S)	โซดา I (เล็ก)
Soda I (L)	โซดา I (ใหญ่)
Soda II (S)	โซดา II (เล็ก)
WaterPet (S)	น้ำดื่มพลาสติก (เล็ก)
WaterPet (M)	น้ำดื่มพลาสติก (กลาง)
WaterPet (L)	น้ำดื่มพลาสติก (ใหญ่)
WaterGallon	น้ำแกลลอน
WaterGlass	น้ำดื่มขวดแก้ว
MineralWater (S)	น้ำแร่ (เล็ก)
MineralWater (L)	น้ำแร่ (ใหญ่)

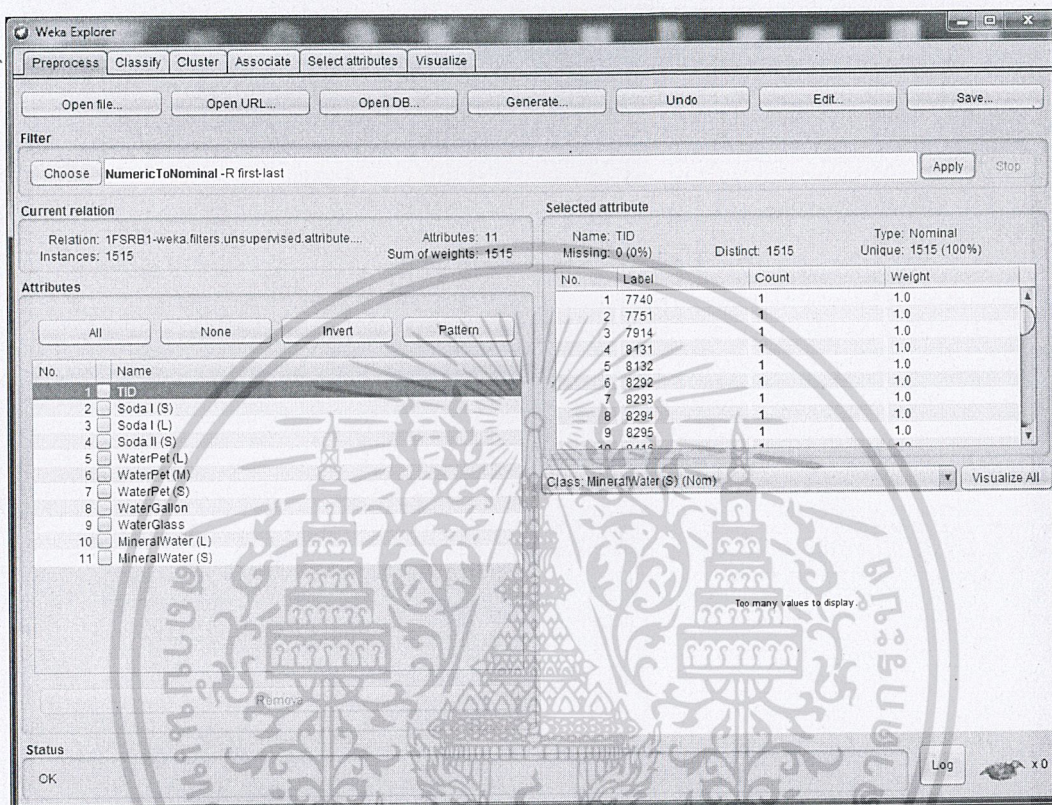
โดยค่าข้อมูลนำมาใช้ในการวิเคราะห์ประกอบด้วย 2 รูปแบบคือ

- a) ค่า 1 หมายถึง ลูกค้ามีการสั่งซื้อสินค้าชนิดนั้น
- b) ค่า 0 หมายถึง ลูกค้าไม่มีการสั่งซื้อสินค้าชนิดนั้น

การนำเข้าข้อมูลในโปรแกรม Weka

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 เนื่องจากรูปแบบข้อมูลที่ใช้ในการประมวลผลนั้นต้องอยู่ในรูปแบบไฟล์นามสกุล .arff หรือ .csv
 ไม่ว่าจะกรณีใดๆ ห้ามมิให้คัดลอกและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

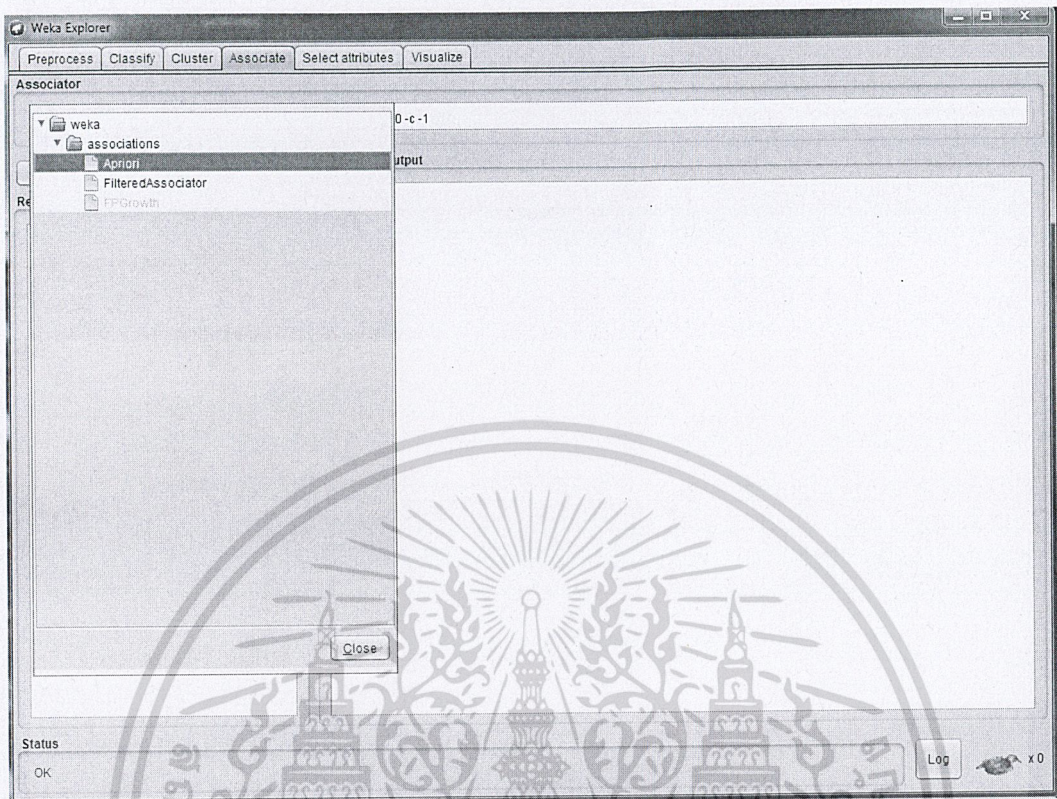
ดังนั้นผู้วิจัยได้ทำการแปลงข้อมูลให้อยู่ในรูปแบบ .csv เพื่อความเหมาะสมในการนำเข้าข้อมูล และสามารถแก้ไขข้อมูล ดังรูปที่ 4-3



รูปที่ 4-4 การนำข้อมูลเข้าโปรแกรม Weka

จากรูปที่ 4-4 เป็นการนำข้อมูลเข้าโปรแกรม Weka ด้วยเทคนิค Association ซึ่งข้อมูลจะถูกแบ่งออกเป็นกลุ่มของสินค้าแต่ละชนิดดังภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-5 การนำเข้าข้อมูลในโปรแกรม Weka ด้วยเทคนิค Association Rule โดยใช้อัลกอริทึม Apriori

จากรูปที่ 4-5 เป็นการนำเข้าข้อมูลในโปรแกรม Weka ด้วยเทคนิค Association Rule โดยใช้อัลกอริทึม Apriori ซึ่งผลลัพธ์ที่ได้จะมีการบ่งบอกถึงเปอร์เซ็นต์ความน่าเชื่อถือในการแสดงผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลการดำเนินงาน

ในการวิจัยข้อมูลเพื่อสร้างแบบจำลองกฎความสัมพันธ์ สำหรับฐานข้อมูลการสั่งซื้อสินค้าของลูกค้าด้วยการใช้อัลกอริทึม Apriori เพื่อหากฎความสัมพันธ์ของการสั่งซื้อสินค้าทั้ง 6 หน่วยขาย คือ Type 1 A, Type 1 B, Type 1 C, Type 2 A, Type 2 B และ Type 2 C ซึ่ง Type 1 คือหน่วยขายที่เป็นศูนย์กระจายสินค้าไปยังลูกค้าประเภทร้านอาหารหรือโรงแรม ส่วน Type 2 คือหน่วยขายที่กระจายสินค้าไปยังลูกค้าที่อาศัยอยู่บ้านเรือนหรือคอนโด โดยวิเคราะห์จากรายงานการสั่งซื้อของลูกค้า ได้แก่ ส่วนใหญ่ลูกค้าซื้อสินค้าชิ้นหนึ่งแล้วมักจะซื้อสินค้าอีกชิ้นหนึ่ง จึงจำเป็นต้องใช้กฎความสัมพันธ์ในการทำดาตาไมนนิ่ง ในบทนี้ผู้วิจัยได้นำเสนอผลการวิจัยของชุดข้อมูลที่ผู้วิจัยได้เตรียมข้อมูลชนิดสินค้า และความถี่ของการซื้อสินค้าในแต่ละชนิดเพื่อทำการทดลอง โดยแบ่งชุดข้อมูลการสั่งซื้อออกเป็น 12 หัวข้อ ดังต่อไปนี้คือ

1. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 A โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%
2. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 A โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%
3. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 B โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%
4. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 B โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%
5. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 C โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%
6. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 C โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

เอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 A โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

8. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 A โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

9. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 B โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

10. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 B โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

11. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 C โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

12. ข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 C โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองที่ 1

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 A โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่าการซื้อร่วมกันบ่อย แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.2 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด ดังภาพประกอบที่ 4-6

ตารางที่ 4.2 รูปแบบของสินค้าสองประเภทที่พบว่าการซื้อร่วมกัน MinSupport 0.01, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.01	0.5	น้ำดื่มแกลลอน	น้ำดื่มขวดแก้ว	0.0238	100%	1.16
		น้ำดื่มพลาสติก (ใหญ่)	น้ำดื่มพลาสติก (กลาง)	0.0158	71%	4.93

จากตารางที่ 4.2 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน 2 อันดับแรกที่มีความถี่ในการซื้อ 2.38% ซึ่งมากกว่าสินค้าซื้อร่วมกันกับประเภทอื่น ประกอบด้วย [น้ำดื่มแกลลอน : น้ำดื่มขวดแก้ว] ซึ่งค่าสนับสนุนรองลงมาที่มีความถี่ในการซื้อ 1.58% ประกอบด้วย [น้ำดื่มพลาสติก (ใหญ่) : น้ำดื่มพลาสติก (กลาง)]

Best rules found:

1. WaterGallon=1 36 ==> WaterGlass=1 36 <conf:(1)> lift:(1.16) lev:(0) [5] conv:(5.06)
2. WaterPet (L)=1 34 ==> WaterPet (M)=1 24 <conf:(0.71)> lift:(4.93) lev:(0.01) [19] conv:(2.65)

รูปที่ 4-6 แสดงกฎความสัมพันธ์ของประเภทสินค้ากำหนดค่า MinSupport 0.01 และค่า MinConfidence 0.5

จากรูปที่ 4-6 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่า

MinConfidence 0.5 ของสินค้าทั้งหมด 2 กฎ สามารถอธิบายความหมายได้ว่าไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule 1: สินค้า น้ำดื่มแกลลอน มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 100%

Rule 2: สินค้า น้ำดื่มพลาสติก (ใหญ่) มีความสัมพันธ์กับสินค้า น้ำดื่มพลาสติก (กลาง) ด้วยความเชื่อมั่น 71%

การทดลองที่ 2

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 A โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่ามี การซื้อร่วมกันบ่อยๆ แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.3 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด ดังรูปที่ 4-7

ตารางที่ 4.3 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน MinSupport 0.02, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.02	0.5	น้ำดื่มแกลลอน	น้ำดื่มขวดแก้ว	0.0238	100%	1.16

จากตารางที่ 4.3 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน 1 อันดับแรกที่มีความถี่ในการซื้อ 2.38% ซึ่งมากกว่าสินค้าซื้อ ร่วมกับประเภทอื่น ประกอบด้วย [น้ำดื่มแกลลอน : น้ำดื่มขวดแก้ว]

Best rules found:

```
1. WaterGallon=1 36 ==> WaterGlass=1 36 <conf: (1)> lift: (1.16) lev: (0) [5] conv: (5.06)
```

รูปที่ 4-7 แสดงกฎความสัมพันธ์ของประเภทสินค้ากำหนดค่า MinSupport 0.02 และค่า MinConfidence 0.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4-7 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.02 และค่า MinConfidence 0.5 ของสินค้าทั้งหมด 2 กฎ สามารถอธิบาย ความหมายได้ว่า

Rule 1: สินค้า น้ำดื่มแกลลอน มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 100%

การทดลองที่ 3

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 B โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่าการซื้อร่วมกันบ่อย ๆ แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.4 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด ดังรูปที่ 4-8

ตารางที่ 4.4 รูปแบบของสินค้าสองประเภทที่พบว่าการซื้อร่วมกัน MinSupport 0.01, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.01	0.5	น้ำดื่มแกลลอน	น้ำดื่มขวดแก้ว	0.0341	71%	0.91

จากตารางที่ 4.4 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน 1 อันดับแรกที่มีความถี่ในการซื้อ 3.41% ซึ่งมากกว่าสินค้าซื้อพร้อมกับสินค้าชนิดอื่น ประกอบด้วย [น้ำดื่มแกลลอน : น้ำดื่มขวดแก้ว]

Best rules found:

1. WaterGallon=1 128 ==> WaterGlass=1 91 <conf: (0.71)> lift: (0.91) lev: (-0) [-8] conv: (0.75)

รูปที่ 4-8 แสดงกฎความสัมพันธ์ของประเภทสินค้ากำหนดค่า MinSupport 0.01 และค่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งาน MinConfidence 0.5 ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4-8 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่า MinConfidence 0.5 ของสินค้าทั้งหมด 2 กฎ สามารถอธิบาย ความหมายได้ว่า

Rule 1: สินค้า น้ำดื่มเกลลอน มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 71%

การทดลองที่ 4

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 B โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่ามี การซื้อร่วมกันบ่อยๆ แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.5 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด ดังรูปที่ 4-9

ตารางที่ 4.5 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน MinSupport 0.02, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.02	0.5	น้ำดื่มเกลลอน	น้ำดื่มขวดแก้ว	0.0341	71%	0.91

จากตารางที่ 4.5 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน 1 อันดับแรกที่มีความถี่ในการซื้อ 3.41% ซึ่งมากกว่าสินค้าซื้อร่วมกันกับสินค้าชนิดอื่น ประกอบด้วย [น้ำดื่มเกลลอน : น้ำดื่มขวดแก้ว]

Best rules found:

```
1. WaterGallon=1 128 ==> WaterGlass=1 91 <conf: (0.71)> lift: (0.91) lev: (-0) [-8] conv: (0.75)
```

รูปที่ 4-9 แสดงกฎความสัมพันธ์ของประเภทสินค้ากำหนดค่า MinSupport 0.02 และค่า MinConfidence 0.5

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนำไปเผยแพร่บนสื่อออนไลน์โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย

จากรูปที่ 4-9 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.02 และค่า MinConfidence 0.5 ของสินค้าทั้งหมด 1 กฎ สามารถอธิบาย ความหมายได้ว่า

Rule 1: สินค้า น้ำดื่มแกลลอน มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 71%

การทดลองที่ 5

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 C โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่าการซื้อร่วมกันบ่อยๆ แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.6 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด ดังรูปที่ 4-10

ตารางที่ 4.6 รูปแบบของสินค้าสองประเภทที่พบว่าการซื้อร่วมกัน MinSupport 0.01, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.01	0.5	โซดา I (ใหญ่)	น้ำดื่มขวดแก้ว	0.0182	55%	0.65

จากตารางที่ 4.6 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน 1 อันดับแรกที่มีความถี่ในการซื้อ 1.82% ซึ่งมากกว่าสินค้าซื้อร่วมกันกับสินค้าชนิดอื่น ประกอบด้วย [โซดา I (ใหญ่) : น้ำดื่มขวดแก้ว]

Best rules found:

1. Soda I (L)=1 199 ==> WaterGlass=1 109 <conf: (0.55)> lift: (0.65) lev: (-0.01) [-58] conv: (0.34)

รูปที่ 4-10 แสดงกฎความสัมพันธ์ของประเภทสินค้ากำหนดค่า MinSupport 0.01 และค่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อวัตถุประสงค์เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
MinConfidence 0.5
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4-10 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่า MinConfidence 0.5 ของประเภทสินค้าทั้งหมด 1 กฎ สามารถอธิบาย ความหมายได้ว่า

Rule 1: สินค้า โซดา I (ใหญ่) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 55%

การทดลองที่ 6

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 1 C โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่ามี การซื้อร่วมกัน ดังตารางที่ 4.7

ตารางที่ 4.7 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน MinSupport 0.02, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1และ2	Support	Conf	Lift
0.02	0.5	ไม่พบกฎความสัมพันธ์			

จากตารางที่ 4.7 สามารถอธิบายความหมายได้ว่าไม่มีสินค้าที่ถูกซื้อร่วมกันที่ค่าซัพพอร์ตต่ำสุด 2%

การทดลองที่ 7

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 A โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่ามี การซื้อร่วมกันบ่อยๆ แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.8 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด ดัง

รูปที่ 4-11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.8 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน MinSupport 0.01,
MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.01	0.5	น้ำแร่(ใหญ่)	น้ำดื่มขวดแก้ว	0.0193	59%	1.39
		น้ำดื่มพลาสติก(ใหญ่)	น้ำดื่มแกลลอน	0.0166	55%	0.89
		น้ำดื่มพลาสติก(กลาง)	น้ำดื่มขวดแก้ว	0.0220	52%	1.24
		น้ำดื่มพลาสติก(เล็ก)	น้ำดื่มขวดแก้ว	0.0182	55%	0.65

จากตารางที่ 4.8 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน 2 อันดับแรกที่มีความถี่ในการซื้อ 2.20% ซึ่งมากกว่าสินค้าซื้อร่วมกันกับสินค้าอื่น ประกอบด้วย [น้ำดื่มพลาสติก(กลาง) : น้ำดื่มขวดแก้ว] ซึ่งค่าสนับสนุนรองลงมาที่มีความถี่ในการซื้อ 1.93% ประกอบด้วย [น้ำแร่(ใหญ่) : น้ำดื่มพลาสติก(กลาง)]

Best rules found:

1. MineralWater (S)=1 237 ==> WaterGlass=1 140 <conf:(0.59)> lift:(1.39) lev:(0.01) [39] conv:(1.39)
2. WaterPet (L)=1 218 ==> WaterGallon=1 120 <conf:(0.55)> lift:(0.89) lev:(-0) [-14] conv:(0.84)
3. WaterPet (M)=1 303 ==> WaterGlass=1 159 <conf:(0.52)> lift:(1.24) lev:(0) [30] conv:(1.2)
4. WaterPet (S)=1 176 ==> WaterGlass=1 88 <conf:(0.5)> lift:(1.18) lev:(0) [13] conv:(1.14)

รูปที่ 4-11 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่า
MinConfidence 0.5

จากรูปที่ 4-11 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่า MinConfidence 0.5 ของสินค้าทั้งหมด 1 กฎ สามารถอธิบาย ความหมายได้ว่า

Rule 1: สินค้า น้ำแร่ (ใหญ่) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 59%

Rule 2: สินค้า น้ำดื่มพลาสติก (ใหญ่) มีความสัมพันธ์กับสินค้า น้ำดื่มแกลลอน ด้วยความ
เชื่อมั่น 55%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule 3: สินค้า น้ำดื่มพลาสติก (กลาง) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 52%

Rule 4: สินค้า น้ำดื่มพลาสติก (เล็ก) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 50%

การทดลองที่ 8

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 A โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่าการซื้อร่วมกันบ่อยๆ แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.9 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด ดังรูปที่ 4-12

ตารางที่ 4.9 รูปแบบของสินค้าสองประเภทที่พบว่าการซื้อร่วมกัน MinSupport 0.02, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.02	0.5	น้ำดื่มพลาสติก(กลาง)	น้ำดื่มขวดแก้ว	0.0220	52%	1.24

จากตารางที่ 4.9 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน อันดับแรกที่มีความถี่ในการซื้อ 2.20% ซึ่งมากกว่าสินค้าซื้อร่วมกันกับสินค้าอื่น ประกอบด้วย [น้ำดื่มพลาสติก(กลาง) : น้ำดื่มขวดแก้ว]

Best rules found:

1. WaterPet (M)=1 303 ==> WaterGlass=1 159 <conf:(0.52)> lift:(1.24) lev:(0) [30] conv:(1.2)

รูปที่ 4-12 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.02 และค่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ MinConfidence 0.5 ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4-12 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.02 และค่า MinConfidence 0.5 ของประเภทสินค้าทั้งหมด 1 กฎ สามารถอธิบาย ความหมายได้ว่า

Rule 1: สินค้า น้ำดื่มพลาสติก (กลาง) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 52%

การทดลองที่ 9

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 B โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่าการซื้อร่วมกันบ่อยๆ แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.10 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด

ตารางที่ 4.10 รูปแบบของสินค้าสองประเภทที่พบว่าการซื้อร่วมกัน MinSupport 0.01, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.01	0.5	น้ำดื่มพลาสติก(กลาง)	น้ำดื่มขวดแก้ว	0.0141	60%	1.78
		น้ำดื่มพลาสติก(ใหญ่)	น้ำดื่มขวดแก้ว	0.0115	54%	1.63
		น้ำแร่(ใหญ่)	น้ำดื่มขวดแก้ว	0.0100	51%	1.52

จากตารางที่ 4.10 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน 3 อันดับที่มีความถี่ในการซื้อ 1.41% ซึ่งมากกว่าสินค้าซื้อร่วมกันกับประเภทอื่น ประกอบด้วย [น้ำดื่มพลาสติก (กลาง): น้ำดื่มขวดแก้ว] ซึ่งค่าสนับสนุนรองลงมาที่มีความถี่ในการซื้อ 1.15% ประกอบด้วย [น้ำดื่มพลาสติก (ใหญ่) : น้ำดื่มขวดแก้ว] และสุดท้ายค่าสนับสนุนในการซื้อสินค้านี้ร่วมกัน 0.01% คือ [น้ำแร่(ใหญ่) : น้ำดื่มขวดแก้ว]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Best rules found:

1. WaterPet (M)=1 128 ==> WaterGlass=1 80 <conf:(0.63)> lift:(1.85) lev:(0.01) [36] conv:(1.73)
2. WaterPet (L)=1 120 ==> WaterGlass=1 65 <conf:(0.54)> lift:(1.6) lev:(0) [24] conv:(1.42)
3. MineralWater (S)=1 111 ==> WaterGlass=1 57 <conf:(0.51)> lift:(1.52) lev:(0) [19] conv:(1.34)

รูปที่ 4-13 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่า MinConfidence 0.5

จากรูปที่ 4-13 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่า MinConfidence 0.5 ของประเภทสินค้าทั้งหมด 3 กฎ สามารถอธิบายความหมายได้ว่า

Rule 1: สินค้า น้ำดื่มพลาสติก (กลาง) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 63%

Rule 2: สินค้า น้ำดื่มพลาสติก (ใหญ่) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 54%

Rule 3: สินค้า น้ำแร่ (ใหญ่) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 51%

การทดลองที่ 10

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 B โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่ามี การซื้อร่วมกันบ่อย ดังตาราง ที่ 4.11

ตารางที่ 4.11 รูปแบบของสินค้าสองประเภทที่พบว่ามี การซื้อร่วมกัน MinSupport 0.02, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1 และ 2	Support	Conf	Lift
0.02	0.5	ไม่พบกฎความสัมพันธ์			

จากตารางที่ 4.11 สามารถอธิบายความหมายได้ว่าไม่มีสินค้าที่ถูกซื้อร่วมกันที่ค่า support เอกสารนี้ต่ำสุด 2% หน้าที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองที่ 11

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 C โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.01 หรือ 1% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่าการซื้อร่วมกันบ่อยๆ แบ่งออกเป็น 1 กลุ่ม ได้แก่ การซื้อสินค้าสองประเภท ดังตาราง ที่ 4.12 และกฎความสัมพันธ์จากข้อมูลรูปแบบที่ค้นพบทั้งหมด ดังรูปที่ 4-14

ตารางที่ 4.12 รูปแบบของสินค้าสองประเภทที่พบว่าการซื้อร่วมกัน MinSupport 0.01, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1	สินค้าที่ 2	Support	Conf	Lift
0.01	0.5	น้ำแร่ (ใหญ่)	น้ำดื่มขวดแก้ว	0.0154	67%	2.03
		น้ำดื่มพลาสติก (กลาง)	น้ำดื่มแกลลอน	0.0194	52%	0.74

จากตารางที่ 4.12 สามารถอธิบายความหมายได้ว่าสินค้าที่ถูกซื้อร่วมกันสองชนิด โดยพิจารณาจากค่าสนับสนุน 2 อันดับแรกที่มีความถี่ในการซื้อ 1.54% ซึ่งมากกว่าสินค้าซื้อพร้อมกันกับประเภทอื่น ประกอบด้วย [น้ำแร่(ใหญ่) : น้ำดื่มขวดแก้ว] และค่าสนับสนุนรองลงมาที่มีความถี่ในการซื้อ 1.49% ประกอบด้วย [น้ำดื่มพลาสติก(ใหญ่) : น้ำดื่มแกลลอน]

Best rules found:

1. MineralWater (S)=1 176 ==> WaterGlass=1 118 <conf:(0.67)> lift:(2.03) lev:(0.01) [59] conv:(2)
2. WaterPet (M)=1 286 ==> WaterGallon=1 149 <conf:(0.52)> lift:(0.74) lev:(-0.01) [-51] conv:(0.62)

รูปที่ 4-14 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่า MinConfidence 0.5

จากรูปที่ 4-14 แสดงกฎความสัมพันธ์จากค่าพารามิเตอร์ MinSupport 0.01 และค่าเอกสารนี้ MinConfidence 0.5 ของประเภทสินค้าทั้งหมด 2 กฎ สามารถอธิบาย ความหมายได้ว่าเงื่อนไขด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Rule 1: สินค้า น้ำแร่ I (ใหญ่) มีความสัมพันธ์กับสินค้า น้ำดื่มขวดแก้ว ด้วยความเชื่อมั่น 67%

Rule 2: สินค้า น้ำดื่มพลาสติก (กลาง) มีความสัมพันธ์กับสินค้า น้ำดื่มแกลลอน ด้วยความเชื่อมั่น 52%

การทดลองที่ 12

ทำการทดลองข้อมูลรายการสั่งซื้อของลูกค้าที่ผ่านหน่วยขาย Type 2 C โดยกำหนดค่าพารามิเตอร์ MinSupport ไว้ที่ 0.02 หรือ 2% และได้ทำการกำหนดค่าพารามิเตอร์ MinConfidence ไว้ที่ 0.5 หรือ 50%

จะสามารถพบรูปแบบของสินค้าที่พบว่าการซื้อร่วมกันบ่อย ดังตาราง ที่ 4.13

ตารางที่ 4.13 รูปแบบของสินค้าสองประเภทที่พบว่าการซื้อร่วมกัน Support 0.02, MinConfidence 0.5

MinSupport	Confidence	สินค้าที่ 1และ2	Support	Conf	lift
0.02	0.5	ไม่พบกฎความสัมพันธ์			

จากตารางที่ 4.13 สามารถอธิบายความหมายได้ว่าไม่มีสินค้าที่ถูกซื้อร่วมกันที่ค่าซัพพอร์ตต่ำสุด 2%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 ผลการทดสอบเพื่อวัดประสิทธิภาพ

จากผลการวิจัยโดยผ่านการประมวลผลด้วยอัลกอริทึม Apriori นั้นได้กฎที่ดีที่สุดสำหรับหน่วยขายทั้ง 6 หน่วยทั้งหมด 18 กฎที่มีค่า MinConfidence มากกว่า 0.5 ซึ่งผู้วิจัยได้สรุปว่าข้อมูลที่นำเข้าสู่โปรแกรม Weka ในการวิเคราะห์นั้น สามารถให้ผลในการวิเคราะห์และค่าความน่าเชื่อถือที่สอดคล้องกัน แสดงให้เห็นว่าข้อมูลในการนำเข้ามีความน่าเชื่อถือและเหมาะสมในการหาค่ากฎความสัมพันธ์

จากกฎความสัมพันธ์ที่ได้จากทั้ง 6 หน่วยขาย สามารถสรุปได้ว่า มี 4 ตัวแบบที่เหมาะสมในการนำไปใช้จัดโปรโมชั่นและวางแผนการตลาดดังต่อไปนี้

1. ตัวแบบที่ 1 ถ้าลูกค้าซื้อน้ำถัง แล้ว ซื้อน้ำดื่มขวดแก้ว
2. ตัวแบบที่ 2 ถ้าลูกค้าซื้อโซดา 1 (ใหญ่) แล้วจะซื้อน้ำดื่มขวดแก้ว
3. ตัวแบบที่ 3 ถ้าลูกค้าซื้อน้ำดื่มพลาสติก (กลาง) แล้ว ซื้อน้ำดื่มขวดแก้ว
4. ตัวแบบที่ 4 ถ้าลูกค้าซื้อน้ำแร่ (เล็ก) แล้ว ซื้อน้ำดื่มขวดแก้ว

หลังจากได้ตัวแบบการจำลองกฎความสัมพันธ์แล้ว ผู้วิจัยต้องการทดสอบตัวแบบโดยการนำตัวแบบไปทดสอบกับกลุ่มลูกค้าที่เคยได้รับโปรโมชั่นของทางบริษัท จำนวน 465 คน โดยทดสอบในโปรแกรม Excel ผ่านคำสั่ง if then และข้อมูลลูกค้าที่ใช้ในการทดสอบถูกจัดในรูปแบบเดียวกันกับข้อมูลที่ใช้ทดสอบ ดังต่อไปนี้

1. ข้อมูลของลูกค้า
2. ข้อมูลผลิตภัณฑ์สินค้าในการจำหน่าย
3. ข้อมูลการขาย

โดยค่าข้อมูลนำมาใช้ในการวิเคราะห์ประกอบด้วย 2 รูปแบบคือ

ค่า 1 หมายถึง ลูกค้ามีการสั่งซื้อสินค้าชนิดนั้น

ค่า 0 หมายถึง ลูกค้าไม่มีการสั่งซื้อสินค้าชนิดนั้น

	A	B	C	D	E	F	G	H	I	J	K
1	TID	MineralWater (S)	MineralWater (L)	Soda I (S)	Soda I (L)	Soda II (S)	WaterPet (S)	WaterPet (M)	WaterPet (L)	WaterGallon	WaterGlass5
2	1	0	0	0	0	0	0	0	0	0	1
3	2	0	0	0	0	1	0	0	0	0	0
4	3	0	0	0	1	0	0	0	0	0	0
5	4	0	0	0	0	1	0	1	0	0	0
6	5	0	0	0	0	0	0	0	0	0	1
7	6	0	0	0	1	0	0	0	0	0	1
8	7	0	0	0	0	0	0	0	0	0	1
9	8	0	0	0	0	0	0	0	0	1	1
10	9	0	0	0	0	0	0	0	0	1	1
11	10	0	0	0	0	0	0	0	0	1	1
12	11	0	0	0	0	1	0	1	0	0	0
13	12	0	0	0	1	0	0	0	0	0	0
14	13	0	0	0	0	0	0	0	0	0	1
15	14	0	0	0	1	0	0	0	0	0	0
16	15	0	0	0	0	0	0	0	0	0	1
17	16	0	0	0	1	0	0	0	0	0	0
18	17	0	0	0	0	0	0	0	0	0	1
19	18	0	0	0	0	0	0	0	1	0	0
20	19	0	0	0	0	0	0	0	0	0	1
21	20	0	0	0	0	0	0	0	0	0	1
22	21	0	0	0	0	0	0	0	0	0	1
23	22	0	0	0	0	0	0	0	0	0	1
24	23	0	0	0	0	0	0	1	0	0	0
25	24	0	0	0	0	0	0	0	0	0	1
26	25	0	0	0	0	0	0	0	0	0	1
27	26	0	0	0	0	0	0	0	0	0	1
28	27	0	0	0	0	0	0	0	0	0	1
29	28	0	0	0	0	0	0	0	0	0	1
30	29	0	0	0	0	0	0	0	0	0	1
31	30	0	0	0	0	0	0	0	0	0	1
32	31	0	0	0	0	0	0	1	0	0	1
33	32	0	0	0	0	0	0	0	0	0	1

รูปที่ 4-15 นำข้อมูลที่ใช้ในการทดสอบตัวแบบจำลองความสัมพันธ์เข้าสู่โปรแกรม Excel

จากนั้นสร้างคำสั่ง IF THEN โดยวางคำสั่งออกเป็น 4 แบบคือ

1. ถ้าลูกค้าน้ำถึงแล้วจะซื้อน้ำดื่มขวดแก้ว

คำสั่ง คือ =IF(AND(น้ำถึง=1,น้ำดื่มขวดแก้ว=1),1,0)

หมายความว่า ถ้าลูกค้าซื้อน้ำถึงและน้ำดื่มขวดแก้วให้แสดงค่าเป็น 1 อย่างอื่นให้แสดงว่าเป็น 0

2. ถ้าลูกค้าน้ำโซดา I (ใหญ่) แล้วจะซื้อน้ำดื่มขวดแก้ว

คำสั่ง คือ =IF(AND(โซดา I (ใหญ่) =1,น้ำดื่มขวดแก้ว=1),1,0)

หมายความว่า ถ้าลูกค้าซื้อโซดา I (ใหญ่) และน้ำดื่มขวดแก้วพร้อมกันให้แสดงค่าเป็น 1 ส่วนที่เหลือให้แสดงค่าเป็น 0

3. ถ้าลูกค้าน้ำดื่มพลาสติก (กลาง) แล้วซื้อน้ำดื่มขวดแก้ว

คำสั่ง คือ =IF(AND(น้ำดื่มพลาสติก (กลาง) =1,ขวดแก้ว=1),1,0)

หมายความว่า ถ้าลูกค้าซื้อน้ำดื่มพลาสติก (กลาง) และน้ำดื่มขวดแก้วพร้อมกันให้แสดงค่าเป็น 1 ส่วนที่เหลือให้แสดงค่าเป็น 0

4. ถ้าลูกค้าน้ำแร่(เล็ก)แล้วซื้อน้ำดื่มขวดแก้ว

คำสั่ง คือ =IF(AND(น้ำแร่ (เล็ก) =1,น้ำดื่มขวดแก้ว=1),1,0)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หมายความว่า ถ้าลูกค้าซื้อน้ำแร่ (เล็ก) และน้ำดื่มขวดแก้วพร้อมกันให้แสดงค่าเป็น 1 ส่วนที่เหลือให้แสดงค่าเป็น 0

จากนั้นทำการทดสอบกับลูกค้าทั้ง 465 คนแสดงได้ดังตารางที่ 4.14

ตารางที่ 4.14 แสดงการทดสอบตัวแบบจำลองของกฎความสัมพันธ์โดยใช้ข้อมูลลูกค้าที่เคยได้รับ

โปรโมชั่น

TID	ถ้าซื้อน้ำถังแล้วซื้อน้ำดื่มขวดแก้ว	ถ้าซื้อโซดา (ใหญ่) แล้วซื้อน้ำดื่มขวดแก้ว	ถ้าซื้อน้ำพลาสติก (กลาง) แล้วซื้อน้ำดื่มขวดแก้ว	ถ้าซื้อน้ำแร่ (เล็ก) แล้วซื้อน้ำดื่มขวดแก้ว
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0
9	0	0	0	0
10	0	0	1	0
11	0	0	0	0
12	0	0	0	0
13	0	0	0	0
14	0	0	1	0
15	0	0	0	0
16	0	0	1	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	1	0
...

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้แก้ไขประโยชน์ใด ๆ ในเอกสาร
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการทดสอบตัวแบบจำลองกับกลุ่มลูกค้าที่เคยรับโปรโมชั่นผลปรากฏว่า ทั้ง 4 ตัวแบบ ให้ผลที่ไม่แตกต่างกันมากดังต่อไปนี้

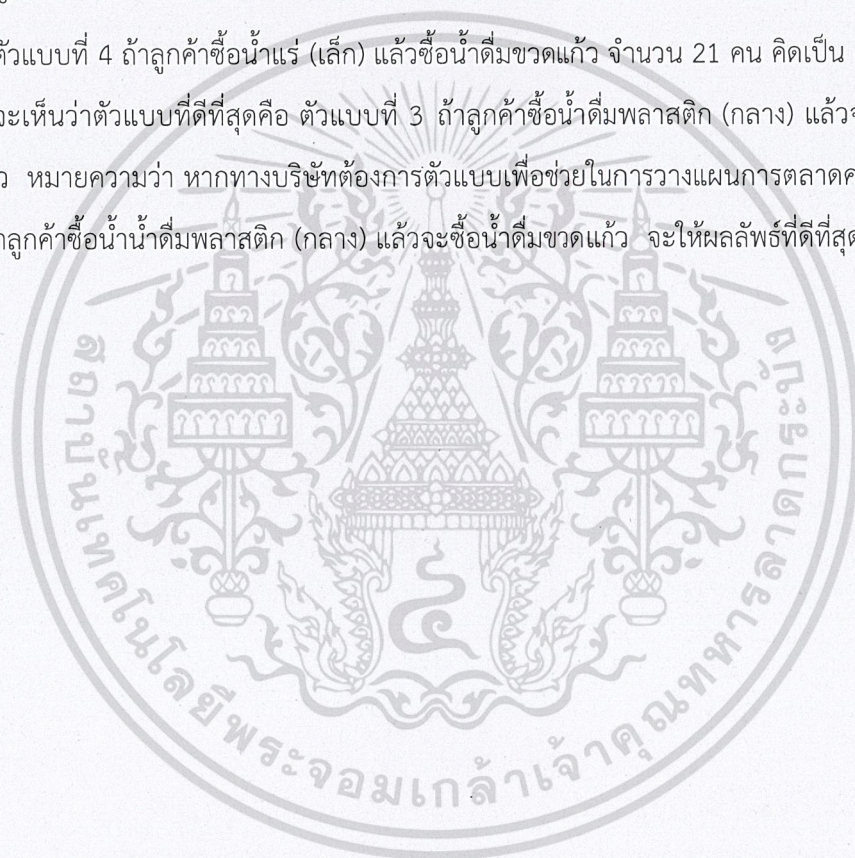
ตัวแบบที่ 1 ถ้าลูกค้าซื้อน้ำถัง แล้วซื้อน้ำดื่มขวดแก้ว จำนวน 15 คน คิดเป็น 3.23%

ตัวแบบที่ 2 ถ้าลูกค้าซื้อน้ำโซดา 1 (ใหญ่) แล้วซื้อน้ำดื่มขวดแก้ว จำนวน 17 คน คิดเป็น 3.66%

ตัวแบบที่ 3 ถ้าลูกค้าซื้อน้ำดื่มพลาสติก (กลาง) แล้วซื้อน้ำดื่มขวดแก้ว จำนวน 76 คน คิดเป็น 16.34%

ตัวแบบที่ 4 ถ้าลูกค้าซื้อน้ำแร่ (เล็ก) แล้วซื้อน้ำดื่มขวดแก้ว จำนวน 21 คน คิดเป็น 4.52%

จะเห็นว่าตัวแบบที่ดีที่สุดคือ ตัวแบบที่ 3 ถ้าลูกค้าซื้อน้ำดื่มพลาสติก (กลาง) แล้วจะซื้อน้ำดื่มขวดแก้ว หมายความว่า หากทางบริษัทต้องการตัวแบบเพื่อช่วยในการวางแผนการตลาดควรเลือกตัวแบบ ถ้าลูกค้าซื้อน้ำดื่มพลาสติก (กลาง) แล้วจะซื้อน้ำดื่มขวดแก้ว จะให้ผลลัพธ์ที่ดีที่สุด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ในบทนี้เป็นกรอภิปรายเพื่อสรุปผลที่ได้จากการทดลองงานวิจัย รวมทั้งข้อเสนอแนะสำหรับแนวทางในการพัฒนางานวิจัยนี้ต่อไป เพื่อแก้ข้อบกพร่องของแบบจำลองให้มีประสิทธิภาพมากขึ้น จากการวิจัยเกี่ยวกับการนำเอาระบบเหมืองข้อมูล เข้ามาช่วยในการสร้างแบบจำลองกฎความสัมพันธ์ สำหรับฐานข้อมูลการสั่งซื้อสินค้าของลูกค้าโดยใช้อัลกอริทึม Apriori โดยใช้ชุดข้อมูลมาตรฐานจากฐานข้อมูลของแอปพลิเคชัน เพื่อใช้ประกอบการทดลองในการหาความสัมพันธ์ โปรแกรมที่นำมาช่วยในการสร้างแบบจำลอง ผู้วิจัยได้นำโปรแกรม Weka เข้ามาช่วยในการสร้างแบบจำลอง โดยใช้แบบจำลองข้อมูลเพื่อหาความสัมพันธ์ของการสั่งซื้อสินค้า โดยใช้วิธีกฎความสัมพันธ์ ซึ่งจะสามารถหาความสัมพันธ์ของการสั่งซื้อสินค้า โดยวิเคราะห์จากการสั่งซื้อของลูกค้า แล้วนำกฎความสัมพันธ์ที่ได้มาจัดโปรแกรมขึ้นให้กับลูกค้า และนำเสนอสินค้าได้ตามความต้องการของลูกค้าในแต่ละราย เพื่อเป็นการตอบสนองความพึงพอใจของลูกค้าต่อไป อีกทั้งยังสามารถพยากรณ์ความต้องการของสินค้าได้อย่างเหมาะสม เพื่อวางแผนการตลาด ผู้วิจัยได้สรุปผลการทดลอง ดังนี้

5.1.1 สรุปผลตามวัตถุประสงค์ของงานวิจัย

1. สามารถสร้างแบบจำลองเพื่อหาความสัมพันธ์ สำหรับเป็นข้อมูลสนับสนุนการตัดสินใจของบริษัทในการวางแผนการตลาด
2. สามารถวิเคราะห์กฎความสัมพันธ์ของสินค้าแต่ละชนิดสำหรับทุกหน่วยขายในการวางแผนการจัดการโปรแกรมขึ้นทางการตลาด โดยใช้อัลกอริทึม Apriori

5.1.2 สรุปผลการทดลองงานวิจัย

งานวิจัยนี้มีวัตถุประสงค์ในการสร้างแบบจำลองเพื่อหาความสัมพันธ์การสั่งซื้อสินค้าผ่าน Application ของบริษัทแห่งหนึ่ง โดยการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์ของการสั่งซื้อสินค้าโดยใช้กฎความสัมพันธ์ (Association Rules) ด้วยอัลกอริทึม Apriori โดยวิเคราะห์จากการสั่งซื้อสินค้าของลูกค้าในแต่ละรายบุคคล ซึ่งผลลัพธ์ที่ได้จากงานวิจัยนี้ สามารถพยากรณ์ความต้องการของลูกค้าออกมาเป็นกฎความสัมพันธ์ ได้ 2 รูปแบบประกอบด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ค่าพารามิเตอร์ Support ไว้ที่ 0.01 และค่าพารามิเตอร์ Confidence ไว้ที่ 0.5 กฎความสัมพันธ์ของสินค้าทั้งหมด 13 กฎ

2. ค่าพารามิเตอร์ Support ไว้ที่ 0.02 และค่าพารามิเตอร์ Confidence ไว้ที่ 0.5 กฎความสัมพันธ์ของสินค้าทั้งหมด 5 กฎ

ซึ่งสามารถนำไปใช้ในการสนับสนุนการตัดสินใจ การนำเสนอสินค้าได้ตามความต้องการของลูกค้าในแต่ละราย เป็นการสร้างความพึงพอใจให้กับลูกค้า ซึ่งเป็นกลยุทธ์หนึ่งของการวางแผนการตลาดได้อย่างเหมาะสม และมีประสิทธิภาพ ดังแสดงในตารางที่ 5.1

ตารางที่ 5.1 สรุปผลการทดลองงานวิจัย

การทดลอง	ค่าพารามิเตอร์	ค่าพารามิเตอร์	กฎ
1	1%	50%	2
2	2%	50%	1
3	1%	50%	1
4	2%	50%	1
5	1%	50%	1
6	2%	50%	-
7	1%	50%	4
8	2%	50%	1
9	1%	50%	3
10	2%	50%	-
11	1%	50%	2
12	2%	50%	-

จากตารางที่ 5.1 พบว่าการสร้างแบบจำลองเพื่อหาความสัมพันธ์การสั่งซื้อสินค้าผ่านแอปพลิเคชัน ของบริษัทแห่งหนึ่ง โดยการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์ของการสั่งซื้อสินค้าโดยใช้ กฎความสัมพันธ์ด้วยอัลกอริทึม Apriori จาก 12 การทดลองพบกฎความสัมพันธ์ 18 กฎซึ่งสามารถสรุปได้เป็น 4 ตัวแบบคือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวแบบที่ 1 ถ้าลูกค้าซื้อน้ำถัง แล้วซื้อน้ำดื่มขวดแก้ว

ตัวแบบที่ 2 ถ้าลูกค้าซื้อน้ำโซดา I (ใหญ่) แล้วซื้อน้ำดื่มขวดแก้ว

ตัวแบบที่ 3 ถ้าลูกค้าซื้อน้ำดื่มพลาสติก (กลาง) แล้วซื้อน้ำดื่มขวดแก้ว

ตัวแบบที่ 4 ถ้าลูกค้าซื้อน้ำแร่ (เล็ก) แล้วซื้อน้ำดื่มขวดแก้ว

เมื่อเปรียบเทียบทั้ง 4 ตัวแบบพบว่า กฎความสัมพันธ์ถ้าลูกค้าซื้อค่าน้ำดื่มพลาสติก (กลาง) แล้วซื้อน้ำดื่มขวดแก้ว เป็นกฎที่ดีที่สุดที่เหมาะสมในการนำไปใช้ในการวางแผนการตลาดของบริษัท โดยหน่วยขาย Type 1 และ Type 2 ได้กฎที่มีการสั่งซื้อน้ำดื่มขวดแก้วเหมือนกัน ที่แตกต่างกันคือ Type 1 สั่งน้ำถังและโซดาคู่กับน้ำดื่มขวดแก้ว เพราะ Type 1 เป็นลูกค้าประเภทร้านอาหารหรือโรงแรม การสั่งสินค้าถังจะทำให้ประหยัดมากกว่า ส่วนสินค้าน้ำดื่มขวดแก้วอาจใช้วางไว้บนโต๊ะอาหารช่วยตกแต่งโต๊ะอาหารให้ดูมีระดับมากกว่าน้ำดื่มขวดพลาสติก ในส่วนของหน่วยขาย Type 2 มีการสั่งซื้อน้ำดื่มพลาสติก (กลาง) และน้ำแร่ (เล็ก) คู่กับน้ำดื่มขวดแก้ว เนื่องจากสินค้าน้ำดื่มพลาสติก (กลาง) น้ำแร่และขวดแก้ว มีขนาดที่พอเหมาะสามารถแช่ตู้เย็นได้สะดวกกว่าการรอกน้ำจากเครื่องกรองน้ำ หรือจากน้ำถังใหญ่โดยตรง ทำให้สินค้าน้ำดื่มขวดแก้วและน้ำถังเป็นสินค้าที่ขายดี ดังนั้นตัวแบบที่ 1 ถ้าลูกค้าซื้อน้ำถังแล้วซื้อน้ำดื่มขวดแก้ว เหมาะสำหรับหน่วยขาย Type 1 A และ Type 1 B ตัวแบบที่ 2 ถ้าลูกค้าซื้อน้ำโซดา I (ใหญ่) แล้วซื้อน้ำดื่มขวดแก้ว เหมาะสำหรับหน่วยขาย Type 1 C ตัวแบบที่ 3 ถ้าลูกค้าซื้อน้ำดื่มพลาสติก (กลาง) แล้วซื้อน้ำดื่มขวดแก้ว เหมาะสำหรับหน่วยขาย Type 2 A และ Type 2 B และตัวแบบที่ 4 ถ้าลูกค้าซื้อน้ำแร่ (เล็ก) แล้วซื้อน้ำดื่มขวดแก้ว เหมาะสำหรับหน่วยขาย Type 1 C หากต้องการตอบสนองความต้องการลูกค้าให้ได้ตรงตามความต้องการ ควรเลือกโปรโมชั่นให้เหมาะกับประเภทลูกค้าและพื้นที่นั้น ๆ ด้วย ส่วนตัวแบบที่ดีที่สุดหากเปรียบเทียบจากทั้ง 4 ตัวแบบ คือ ตัวแบบที่ 3 หมายความว่า หากทางบริษัทต้องการตัวแบบเพื่อช่วยในการวางแผนการตลาดควรเลือกตัวแบบ ถ้าลูกค้าซื้อน้ำดื่มพลาสติกแล้วจะซื้อน้ำดื่มขวดแก้ว จะให้ผลลัพธ์ที่ดีที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ข้อเสนอแนะ

5.2.1 หากมีการจัดทำแบบสอบถามให้ลูกค้าประเมินหรือลงพื้นที่สำรวจมาช่วยสนับสนุนแบบจำลอง จะทำให้กฎความสัมพันธ์ที่ได้มีความแม่นยำมากขึ้น

5.2.2 การหากฎความสัมพันธ์ของการสั่งซื้อสินค้าของลูกค้าควรค้นหาความสัมพันธ์จากฐานข้อมูลขนาดใหญ่เพื่อที่จะได้กฎความสัมพันธ์ที่มีความน่าเชื่อถือที่สูง

5.2.3 เพื่อให้แบบจำลองกฎความสัมพันธ์สำหรับฐานข้อมูลการสั่งซื้อสินค้าของลูกค้า โดยใช้ อัลกอริทึม Apriori มีประสิทธิภาพมากยิ่งขึ้น ทางผู้วิจัยจึงเห็นควรที่จะศึกษา การใช้อัลกอริทึมอื่น มาช่วยในการพยากรณ์ให้มีความถูกต้องและแม่นยำมากขึ้น



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- กนกวรรณ ลีโรจนประภา, (2558). สถิติเบื้องต้น. กรุงเทพฯ : ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- ชื่นจิตต์ แจ่มเจนกิจ. (2554). การบริหารลูกค้าสัมพันธ์. กรุงเทพฯ : เจริญบุญการพิมพ์.
- ณัฐญาพร ชื่นมัจฉา และนิเวศ จิรวชิตชัย. (2559). การหากฎความสัมพันธ์จากฐานข้อมูลการซื้อสินค้าของลูกค้าโดยใช้ FP-Growth. Science and Tecnology RMUTT Journal, 6(1), 122-131.
- นิเวศ จิรวชิตชัย. (2559). การประยุกต์ใช้ซอฟต์แวร์เวก้าในการจำแนกประเภทข้อมูล. บัณฑิตวิทยาลัย มหาวิทยาลัยศรีปทุม วิทยาเขตชลบุรี.
- บัณฑิตทิศา ทิพย์จักขุ (2553). การเปรียบเทียบประสิทธิภาพของอัลกอริทึม Association ระหว่างอัลกอริทึม Apriori และอัลกอริทึม FP-Growth. Mahidol University.
- พนิดา ทรงรัมย์. (2555). การสร้างกฎความสัมพันธ์จากเซตรายการความถี่และเซตรายการความถี่แบบปิด. Information Technology Journal Mahasarakham University.
- ไพชยนต์ คงไชย. (2557). การพัฒนาขั้นตอนวิธีเพื่อจำแนกประเภทข้อมูลด้วยกฎความสัมพันธ์แบบคลุมเครือที่กะทัดรัด (THE DEVELOPMENT OF DATA CLASSIFICATION ALGORITHM WITH COMPACT FUZZY ASSOCIATION RULES). มหาวิทยาลัยเทคโนโลยีสุรนารี.
- ฟูโตะห์ ต้อมอง. (2553). AMFIST : ขั้นตอนวิธีสำหรับการค้นหากลุ่มข้อมูลที่ปรากฏร่วมกันบ่อย โดยรองรับรายการข้อมูลที่คล้ายคลึงกัน. มหาวิทยาลัยสงขลานครินทร์.
- ภควัต คุปต์ธนโรจน์. (2554). การทำเหมืองข้อมูลเพื่อการขายต่อเนื่องของบริการการเงินสด กรณีศึกษา ธนาคารพาณิชย์แห่งหนึ่ง. Science and Technology Thammasat University Journal, 7-15.
- สายชล สีนสมบุญทอง. (2558). การทำเหมืองข้อมูล. กรุงเทพฯ : จามจุรีโปรดักท์.
- สุรพงศ์ เอื้อวัฒนามงคล. (2557). ดาต้าไมนิ่งการค้นหารูขุมทรัพย์. กรุงเทพฯ : สถาบันบัณฑิตบริหารศาสตร คณะสถิติประยุกต์.
- อนุพงศ์ สุขประเสริฐ. (2017). Introduction to Data mining. [Online]. Available : <https://boykap.files.wordpress.com/2007/10/chapter-1-20173.pdf>
- อัชฌา อระวีพร. (2557). เอกสารประกอบการสอนความน่าจะเป็น. กรุงเทพฯ : ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- เอกสิทธิ์ พัทธวงศ์ศักดิ์. 2558. Practical Data Mining with RapidMiner Studio 6. พิมพ์ครั้งที่ 34. กรุงเทพมหานคร : สำนักพิมพ์เอเชีย ดิจิตอลการพิมพ์ จำกัด.

Agrawal et al. (1993). Mining Association Rules Between Sets of item in Lard Databases. In P.Buneman and S.Jaordai, eds. New York: ACM.

- Berberidis et al. (2005). **Mining for Contiguous Frequent itemsets in Transaction Database**. [Online]. Available : https://www.researchgate.net/publication/252596891_Mining_for_Contiguous_Frequent_Itemsets_in_Transaction_Databases.html
- Gintamaz. (2555). **เทคนิคของ Data mining**, [Online]. Available : http://lem0nz.blogspot.com/2012/03/data-mining_8699.html
- Han, J., Pei, J., and Yin, Y. 2000. **Mining frequent patterns without candidate**
- Han, J. Kamber, M. Pei, J. (2011). **Data Mining: Concepts and Techniques**. 2 ed. Morgan Kaufmann. United States of America.
- HOSSEINI, S. M. S., MALEKI, A. & GHOLAMIAN, M. R. 2010. **Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty**. *Expert Systems with Applications*, 37, 5259-5264.
- Inmon W.H. (2005), **Building the Data Warehouse**, Fourth Edition, Wiley Publishing Inc.
- Kotler, P. (1997). **Marketing management: analysis, planning implementation, and control** Englewood Cliffs, New Jersey: Prentice - Hall.
- Piyanut pintasiri, (2558), **เหมืองข้อมูล (Data Mining)**, [Online]. Available : <http://piyanutpintasiri.blogspot.com/2015/10/datamining-mining-discovery-in-database.html>
- Ridings, C.M., Gefen, D. and Arinze, B. (2002). **Some Antecedents and Effects of Trust in Virtual Communities**. *The Journal of Strategic Information Systems*, 11, 271-295.
- Rygielski et al. (2002). **Data mining techniques for customer relationship management**. [Online]. Available : https://www.researchgate.net/publication/222522103_Data_mining_techniques_for_customer_relationship_management
- Sayad, S. (2019). **Association Rules**. [Online]. Available : https://www.saesayad.com/association_rules.html (เข้าถึงเมื่อวันที่ 12 มิถุนายน 2562)
- Smart Vision Europe. (2015). **CRISP-DM CrossIndustry Standard Process for Data Mining**. [Online]. Available : <http://crisp-dm.eu/home/aboutcrisp-dm/>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ข้อมูลที่ใช้ในการวิเคราะห์หากฎความสัมพันธ์ของสินค้า

ตารางที่ ก-1 ข้อมูลการสั่งซื้อสินค้า เดือนมกราคม 2561 – เดือนธันวาคม 2561

	A	B	C	D	E	F	G	H	I	J	K
1	TID	MineralWater (S)	MineralWater (L)	Soda I (S)	Soda I (L)	Soda II (S)	WaterPet (S)	WaterPet (M)	WaterPet (L)	WaterGallon	WaterGlass5
2	1	0	0	0	0	0	0	0	0	0	1
3	2	0	0	0	0	1	0	0	0	0	0
4	3	0	0	0	1	0	0	0	0	0	0
5	4	0	0	0	0	1	0	1	0	0	0
6	5	0	0	0	0	0	0	0	0	0	1
7	6	0	0	0	1	0	0	0	0	0	1
8	7	0	0	0	0	0	0	0	0	0	1
9	8	0	0	0	0	0	0	0	0	1	1
10	9	0	0	0	0	0	0	0	0	1	1
11	10	0	0	0	0	0	0	0	0	1	1
12	11	0	0	0	0	1	0	1	0	0	0
13	12	0	0	0	1	0	0	0	0	0	0
14	13	0	0	0	0	0	0	0	0	0	1
15	14	0	0	0	1	0	0	0	0	0	0
16	15	0	0	0	0	0	0	0	0	0	1
17	16	0	0	0	1	0	0	0	0	0	0
18	17	0	0	0	0	0	0	0	0	0	1
19	18	0	0	0	0	0	0	0	1	0	0
20	19	0	0	0	0	0	0	0	0	0	1
21	20	0	0	0	0	0	0	0	0	0	1
22	21	0	0	0	0	0	0	0	0	0	1
23	22	0	0	0	0	0	0	0	0	0	1
24	23	0	0	0	0	0	0	1	0	0	0
25	24	0	0	0	0	0	0	0	0	0	1
26	25	0	0	0	0	0	0	0	0	0	1
27	26	0	0	0	0	0	0	0	0	0	1
28	27	0	0	0	0	0	0	0	0	0	1
29	28	0	0	0	0	0	0	0	0	0	1
30	29	0	0	0	0	0	0	0	0	0	1
31	30	0	0	0	0	0	0	0	0	0	1
32	31	0	0	0	0	0	0	1	0	0	1
33	32	0	0	0	0	0	0	0	0	0	1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

รูปที่ ข-1 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori -I-N 10-T0-C 0.5-D 0.05-U 1.0-M 0.02-S-1.0-c-1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 8

Large Itemsets L(1):
Soda I (L)=1 60
Soda II (S)=1 16
WaterPet (L)=1 34
WaterPet (S)=1 40
WaterPet (M)=1 217
WaterGallon=1 36
WaterGlass=1 1302
MineralWater (S)=1 21

Size of set of large itemsets L(2): 4

Large Itemsets L(2):
Soda I (L)=1 WaterGlass=1 21
WaterPet (L)=1 WaterPet (M)=1 24
WaterPet (M)=1 WaterGlass=1 100
WaterGallon=1 WaterGlass=1 36

Best rules found:

1. WaterGallon=1 36 ==> WaterGlass=1 36 <conf:(1)> lift:(1.16) lev:(0) [5] conv:(5.06)
2. WaterPet (L)=1 34 ==> WaterPet (M)=1 24 <conf:(0.71)> lift:(4.93) lev:(0.01) [19] conv:(2.65)

Status

OK Log x0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-2 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori -I-N 10 -T 0 -C 0.5 -D 0.05 -U 1.0 -M 0.02 -S -1.0 -c -1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

Apriori

Minimum support: 0.02 (30 instances)
Minimum metric <confidence>: 0.5
Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Large Itemsets L(1):
Soda I (L)=1 60
WaterPet (L)=1 34
WaterPet (S)=1 40
WaterPet (M)=1 217
WaterGallon=1 36
WaterGlass=1 1302

Size of set of large itemsets L(2): 2

Large Itemsets L(2):
WaterPet (M)=1 WaterGlass=1 100
WaterGallon=1 WaterGlass=1 36

Best rules found:

1. WaterGallon=1 36 ==> WaterGlass=1 36 <conf:(1)> lift:(1.16) lev:(0) [5] conv:(5.06)

Status

OK Log x0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-3 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori -I-N 10 -T 0 -C 0.5 -D 0.05 -U 1.0 -M 0.02 -S 1.0 -c 1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

```
Soda I (L)=1 208
Soda II (S)=1 59
WaterPet (L)=1 93
WaterPet (S)=1 73
WaterPet (M)=1 403
WaterGallon=1 128
WaterGlass=1 2100
MineralWater (S)=1 130

Size of set of large itemsets L(2): 7

Large Itemsets L(2):
Soda I (L)=1 WaterPet (M)=1 82
Soda I (L)=1 WaterGlass=1 86
Soda II (S)=1 WaterPet (M)=1 28
WaterPet (L)=1 WaterPet (M)=1 34
WaterPet (M)=1 WaterGlass=1 118
WaterGallon=1 WaterGlass=1 91
WaterGlass=1 MineralWater (S)=1 38

Size of set of large itemsets L(3): 1

Large Itemsets L(3):
Soda I (L)=1 WaterPet (M)=1 WaterGlass=1 37

Best rules found:
1. WaterGallon=1 128 ==> WaterGlass=1 91 <conf: (0.71)> lift: (0.91) lev: (-0) [-E] conv: (0.75)
```

Status

OK Log x 0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-4 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Associator

Choose Apriori -I-N 10-T 0-C 0.5-D 0.05-U 1.0-M 0.02-S-1.0-c-1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

Minimum metric <confidence>: 0.5
Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 8

Large Itemsets L(1):
Soda I (L)=1 208
Soda II (S)=1 59
WaterPet (L)=1 93
WaterPet (S)=1 73
WaterPet (M)=1 403
WaterGallon=1 128
WaterGlass=1 2100
MineralWater (S)=1 130

Size of set of large itemsets L(2): 4

Large Itemsets L(2):
Soda I (L)=1 WaterPet (M)=1 82
Soda I (L)=1 WaterGlass=1 86
WaterPet (M)=1 WaterGlass=1 113
WaterGallon=1 WaterGlass=1 91

Best rules found:

1. WaterGallon=1 128 ==> WaterGlass=1 91 <conf:(0.71)> lift:(0.91) lev:(-0) [-8] conv:(0.75)

Status

OK Log x0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-5 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori -I-N 10 -T 0 -C 0.5 -D 0.05 -U 1.0 -M 0.02 -S -1.0 -c -1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

Minimum support: 0.01 (60 instances)
Minimum metric <confidence>: 0.5
Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 7

Large Itemsets L(1):
Soda I (L)=1 199
WaterPet (L)=1 231
WaterPet (S)=1 182
WaterPet (M)=1 658
WaterGallon=1 80
WaterGlass=1 5043
MineralWater (S)=1 86

Size of set of large itemsets L(2): 3

Large Itemsets L(2):
Soda I (L)=1 WaterGlass=1 109
WaterPet (L)=1 WaterPet (M)=1 91
WaterPet (M)=1 WaterGlass=1 162

Best rules found:

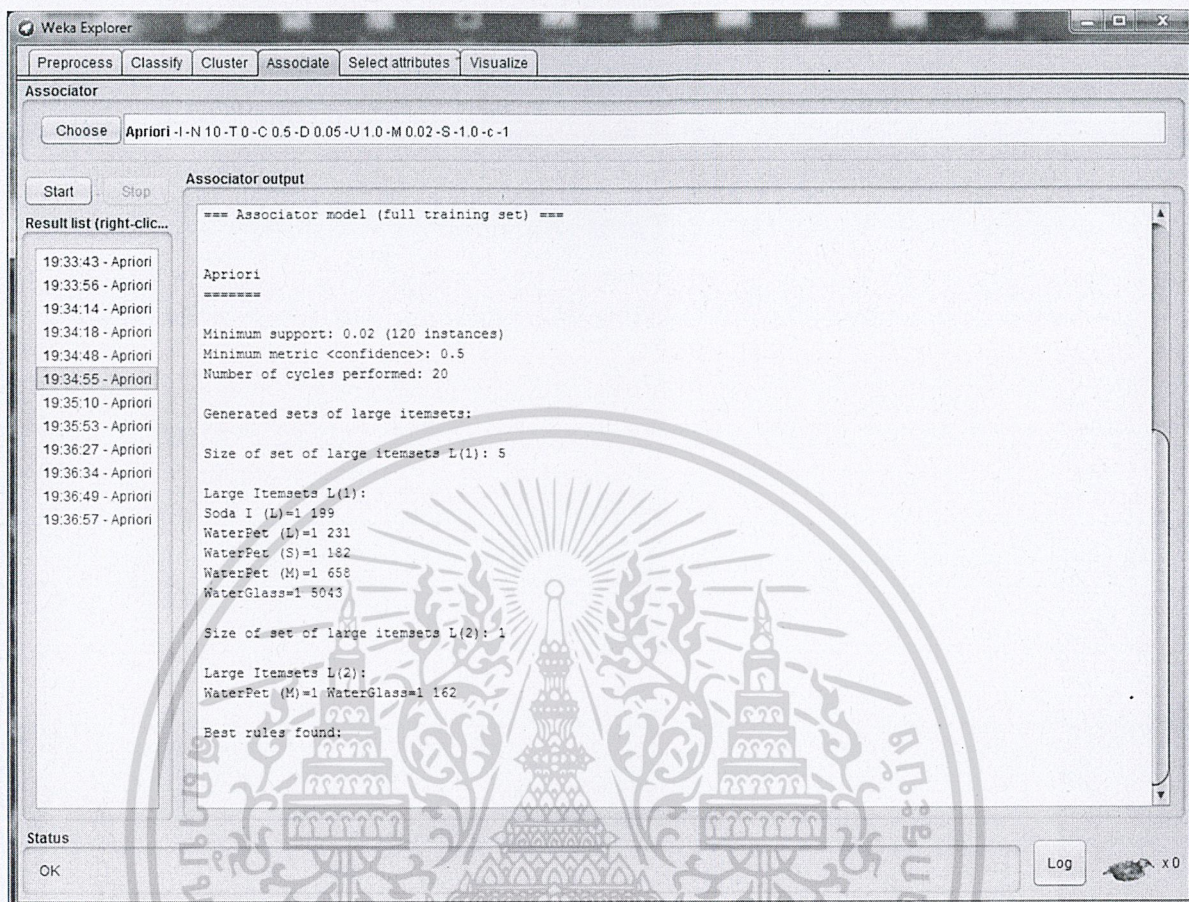
1. Soda I (L)=1 199 ==> WaterGlass=1 109 <conf: (0.55)> lift: (0.65) lev: (-0.01) [-58] conv: (0.34)

Status

OK Log x 0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-6 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-7 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Associator

Choose Apriori -I-N 10-T 0-C 0.5-D 0.05-U 1.0-M 0.02-S-1.0-c-1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

```

Large Itemsets L(1):
WaterPet (L)=1 218
WaterPet (S)=1 176
WaterPet (M)=1 303
WaterGallon=1 4457
WaterGlass=1 3066
MineralWater (S)=1 237

Size of set of large itemsets L(2): 9

Large Itemsets L(2):
WaterPet (L)=1 WaterPet (M)=1 72
WaterPet (L)=1 WaterGallon=1 120
WaterPet (L)=1 WaterGlass=1 73
WaterPet (S)=1 WaterGlass=1 88
WaterPet (M)=1 WaterGallon=1 115
WaterPet (M)=1 WaterGlass=1 159
WaterGallon=1 WaterGlass=1 369
WaterGallon=1 MineralWater (S)=1 74
WaterGlass=1 MineralWater (S)=1 140

Best rules found:
1. MineralWater (S)=1 237 ==> WaterGlass=1 140 <conf:(0.59)> lift:(1.39) lev:(0.01) [39] conv:(1.39)
2. WaterPet (L)=1 218 ==> WaterGallon=1 120 <conf:(0.55)> lift:(0.89) lev:(-0) [-14] conv:(0.84)
3. WaterPet (M)=1 303 ==> WaterGlass=1 159 <conf:(0.52)> lift:(1.24) lev:(0) [30] conv:(1.2)
4. WaterPet (S)=1 176 ==> WaterGlass=1 88 <conf:(0.5)> lift:(1.18) lev:(0) [13] conv:(1.14)

```

Status

OK Log x0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-8 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori -I -N 10 -T 0 -C 0.5 -D 0.05 -U 1.0 -M 0.02 -S -1.0 -c -1

Start Stop

Associator output

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Apriori
=====

Minimum support: 0.02 (144 instances)
Minimum metric <confidence>: 0.5
Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Large Itemsets L(1):
WaterPet (L)=1 218
WaterPet (S)=1 176
WaterPet (M)=1 303
WaterGallon=1 4457
WaterGlass=1 3066
MineralWater (S)=1 237

Size of set of large itemsets L(2): 2

Large Itemsets L(2):
WaterPet (M)=1 WaterGlass=1 159
WaterGallon=1 WaterGlass=1 369

Best rules found:

1. WaterPet (M)=1 303 ==> WaterGlass=1 159 <conf: (0.52)> lift: (1.24) lev: (0) [30] conv: (1.2)

Status

OK Log x0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-9 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori -I-N 10-T0-C 0.5-D 0.05-U 1.0-M 0.02-S-1.0-c-1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

Minimum metric <confidence>: 0.5
Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Large Itemsets L(1):
WaterPet (L)=1 120
WaterPet (S)=1 81
WaterPet (M)=1 128
WaterGallon=1 3814
WaterGlass=1 1918
MineralWater (S)=1 111

Size of set of large itemsets L(2): 4

Large Itemsets L(2):
WaterPet (L)=1 WaterGlass=1 65
WaterPet (M)=1 WaterGlass=1 80
WaterGallon=1 WaterGlass=1 123
WaterGlass=1 MineralWater (S)=1 57

Best rules found:

1. WaterPet (M)=1 128 ==> WaterGlass=1 80 <conf: (0.63)> lift: (1.85) lev: (0.01) [36] conv: (1.73)
2. WaterPet (L)=1 120 ==> WaterGlass=1 65 <conf: (0.54)> lift: (1.6) lev: (0) [24] conv: (1.42)
3. MineralWater (S)=1 111 ==> WaterGlass=1 57 <conf: (0.51)> lift: (1.52) lev: (0) [19] conv: (1.34)

Status

OK Log x0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-10 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori -I-N 10-T 0-C 0.5-D 0.05-U 1.0-M 0.02-S-1.0-c-1

Start Stop

Associator output

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

MineralWater (S)
=== Associator model (full training set) ===

Apriori
=====

Minimum support: 0.02 (113 instances)
Minimum metric <confidence>: 0.5
Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 4

Large Itemsets L(1):
WaterPet (L)=1 120
WaterPet (M)=1 128
WaterGallon=1 3814
WaterGlass=1 1918

Size of set of large itemsets L(2): 1

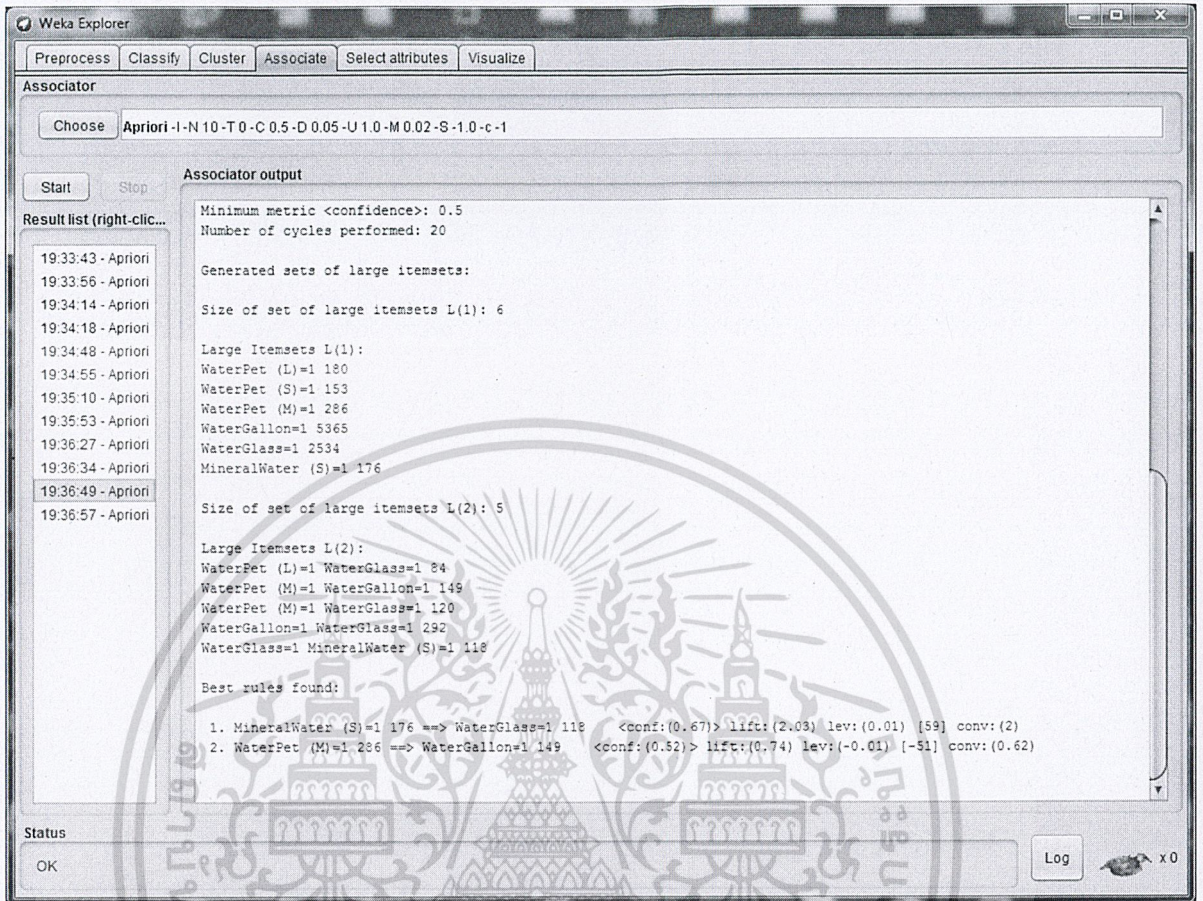
Large Itemsets L(2):
WaterGallon=1 WaterGlass=1 125

Best rules found:

Status
OK Log x 0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-11 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)



Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori -I-N 10 -T 0 -C 0.5 -D 0.05 -U 1.0 -M 0.02 -S -1.0 -c -1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

Minimum metric <confidence>: 0.5
Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Large Itemsets L(1):
WaterPet (L)=1 180
WaterPet (S)=1 153
WaterPet (M)=1 286
WaterGallon=1 5365
WaterGlass=1 2534
MineralWater (S)=1 176

Size of set of large itemsets L(2): 5

Large Itemsets L(2):
WaterPet (L)=1 WaterGlass=1 84
WaterPet (M)=1 WaterGallon=1 149
WaterPet (M)=1 WaterGlass=1 120
WaterGallon=1 WaterGlass=1 292
WaterGlass=1 MineralWater (S)=1 118

Best rules found:

1. MineralWater (S)=1 176 ==> WaterGlass=1 118 <conf:(0.67)> lift:(2.03) lev:(0.01) [59] conv:(2)
2. WaterPet (M)=1 286 ==> WaterGallon=1 149 <conf:(0.52)> lift:(0.74) lev:(-0.01) [-51] conv:(0.62)

Status

OK Log x0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ ข-12 ผลการวิเคราะห์กฎความสัมพันธ์ (Association Rules)

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Associator

Choose Apriori-I-N10-T0-C0.5-D0.05-U1.0-M0.02-S-1.0-c-1

Start Stop

Result list (right-click...)

19:33:43 - Apriori
19:33:56 - Apriori
19:34:14 - Apriori
19:34:18 - Apriori
19:34:48 - Apriori
19:34:55 - Apriori
19:35:10 - Apriori
19:35:53 - Apriori
19:36:27 - Apriori
19:36:34 - Apriori
19:36:49 - Apriori
19:36:57 - Apriori

Associator output

Minimum metric <confidence>: 0.5
Number of cycles performed: 20

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6

Large Itemsets L(1):
WaterPet (L)=1 180
WaterPet (S)=1 153
WaterPet (M)=1 286
WaterGallon=1 5365
WaterGlass=1 2534
MineralWater (S)=1 176

Size of set of large itemsets L(2): 5

Large Itemsets L(2):
WaterPet (L)=1 WaterGlass=1 84
WaterPet (M)=1 WaterGallon=1 149
WaterPet (M)=1 WaterGlass=1 120
WaterGallon=1 WaterGlass=1 292
WaterGlass=1 MineralWater (S)=1 118

Best rules found:

1. MineralWater (S)=1 176 ==> WaterGlass=1 118 <conf: (0.67)> lift: (2.03) lev: (0.01) [59] conv: (2)
2. WaterPet (M)=1 286 ==> WaterGallon=1 149 <conf: (0.52)> lift: (0.74) lev: (-0.01) [-51] conv: (0.62)

Status

OK Log x0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้