

การหาความสัมพันธ์จากฐานข้อมูลการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ มายเฮลท์ ของลูกค้าโดยใช้อัลกอริทึม เอฟพี โกรท และการแบ่งกลุ่ม ลูกค้าตามพฤติกรรมการซื้อสินค้ามายเฮลท์ โดยเทคนิคอาร์เอฟเอ็ม ด้วยโปรแกรม แรพพิดไมเนอร์ : กรณีศึกษาของร้านยาเซนแห่งหนึ่ง

สุภาพรณ คงมณีพรรณ*

บริษัท เอ็มไอเอส ฟาร์มา เซน จำกัด

*Correspondence: supapun.kon@gmail.com

doi: XXXXX

วันที่รับบทความ: 9 ธ.ค. 2562

วันแก้ไขบทความ: 17 ธ.ค. 2562

วันที่ตอบรับบทความ: 19 ธ.ค. 2562

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อค้นหาความสัมพันธ์ด้วยอัลกอริทึม FP-Growth และจัดกลุ่มลูกค้าตามพฤติกรรมการซื้อสินค้า ด้วยเทคนิค RFM รวมทั้งการค้นหาความสัมพันธ์ของการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH ของแต่ละกลุ่มลูกค้า โดยใช้วิธีการทำเหมืองข้อมูล ด้วยข้อมูลของลูกค้าที่ซื้อสินค้า MYHEALTH ภายใต้กรอบ CRISP-DM โดยข้อมูลจำนวน 93,278 แถว 2 คอลัมน์ นำมาใช้เพื่อค้นหาความสัมพันธ์ และข้อมูลจำนวน 63,860 แถว 5 คอลัมน์ นำมาใช้เพื่อจัดกลุ่มลูกค้าตามลำดับ ผลการวิจัยแสดงให้เห็นถึงความสัมพันธ์ของการซื้อผลิตภัณฑ์เสริมอาหาร ยี่ห้อ MYHEALTH จำนวน 7 กฎ ตัวอย่างหนึ่งของกฎความสัมพันธ์ คือ จากรายการซื้อผลิตภัณฑ์หนึ่งพันรายการ พบว่ามี 1 รายการที่ หากลูกค้าซื้อ L-Carnitine และ Royal Jelly แล้ว มีโอกาสที่จะซื้อ Carbo-X ร่วมด้วย ถึง 71.4% และสามารถแบ่งกลุ่มลูกค้าเป็น 6 กลุ่ม ได้แก่ Blue Diamond, Diamond, Platinum, Gold, Silver, Welcome กล่าวคือ (1) ลูกค้ากลุ่ม Blue Diamond เมื่อซื้อผลิตภัณฑ์ CALCIUM และ CARBO-X มีโอกาสจะซื้อผลิตภัณฑ์ L-CARNITINE ร่วมด้วย (2) ลูกค้ากลุ่ม Diamond เมื่อซื้อผลิตภัณฑ์ CARBO-X และ Royal Jelly มีโอกาสจะซื้อผลิตภัณฑ์ L-CARNITINE ร่วมด้วย (3) ลูกค้ากลุ่ม Platinum เมื่อซื้อผลิตภัณฑ์ Fish Oil และ Bilberry มีโอกาสจะซื้อผลิตภัณฑ์ Golden oil ร่วมด้วย (4) ลูกค้ากลุ่ม Gold เมื่อซื้อผลิตภัณฑ์ L-CARNITINE และ Acerola Cherry มีโอกาสจะซื้อผลิตภัณฑ์ CARBO-X ร่วมด้วย (5) ลูกค้ากลุ่ม Silver เมื่อซื้อผลิตภัณฑ์ Calcium และ CARBO-X มีโอกาสจะซื้อผลิตภัณฑ์ L-CARNITINE ร่วมด้วย (6) ลูกค้ากลุ่ม Welcome เมื่อซื้อผลิตภัณฑ์ Rice Bran Oil และ CARBO-X มีโอกาสจะซื้อผลิตภัณฑ์ L-CARNITINE ร่วมด้วย ซึ่งความสัมพันธ์เหล่านี้สามารถช่วยในการจัดทำแผนการตลาด เพื่อตอบสนองความต้องการของลูกค้า และส่งเสริมให้เกิดการซื้ออย่างต่อเนื่องของลูกค้าได้ นอกจากนี้ บริษัทยังสามารถสร้างแผนจัดตำแหน่งสินค้า เพื่อเพิ่มยอดขายให้กับธุรกิจรวมถึงการแบ่งกลุ่มลูกค้าตามลักษณะเฉพาะของพวกเขาเพื่อส่งเสริมพฤติกรรมการซื้อของลูกค้าได้

คำสำคัญ: การทำเหมืองข้อมูล, การวิเคราะห์แบบ RFM, การแบ่งกลุ่มลูกค้า, กฎความสัมพันธ์

**Finding Association Rule from the Database of Purchasing MYHEALTH
Supplement to Customer using the FP-GROWTH Algorithm and Customer
Segmentation according to Purchasing Behavior of MYHEALTH Products with
RFM Techniques of Rapidminer: A Case Study of a Pharmacy Chain**

Supapun Kongmaneepon*

MIS Pharma Chain Co., Ltd.

*Correspondence: supapun.kon@gmail.com

doi: XXXXX

Received: 9 Dec 2019

Revised: 17 Dec 2019

Accepted: 19 Dec 2019

Abstract

The objective of this research is finding the association rules and cluster customers according to their purchasing behavior using FP-Growth and RFM techniques respectively. In addition, this research also finds the association rules of purchasing MYHEALTH products in each customer group of clusters. This research follows the CRISP-DM framework. This research creates association rules from customer purchasing data which comprise 93,278 rows, 2 columns. The clusters customer by RFM model is analyzed from 63,860 rows, 5 columns of customer purchasing data. The research results show 7 association rules of purchasing MYHEALTH products. One of the association rules is that one purchasing transaction out of one thousand purchasing transaction shows the association among Carnitine, Royal Jelly, and Carbo-X with 71.4%. In addition, this pharmacy chain can segment customers into 6 groups which are Blue Diamond, Diamond, Platinum, Gold, Silver and Welcome. The association rules of each cluster are (1) Blue Diamond customers who bought CALCIUM and CARBO-X, will also buy L-CARNITINE ; (2) Diamond customers who bought CARBO-X and Royal Jelly, will also buy L-CARNITINE ; (3) Platinum customers who bought Fish Oil and Bilberry, will also buy Golden oil ; (4) Gold customers who bought L-CARNITINE and Acerola Cherry, will also buy CARBO-X ; (5) Silver customers who bought Calcium, CARBO-X, will also buy L-CARNITINE ; and (6) Welcome customers who bought Rice Bran Oil and CARBO-X, will also buy L-CARNITINE. This association rules can make marketing plans to response customer needs by recommending products that suit to customers' need and to promote the continuous purchase of customers. Furthermore, the company can create a good planogram to increase sales for the business together with the segmentation of customers according to their specific characteristics to promote customers' purchasing behavior.

Keywords: Data mining, RFM Analysis, Customer Segmentation, Association Rules

1. บทนำ

1.1 ความสำคัญและที่มาของปัญหาการวิจัย

ร้านขายยาเป็นหนึ่งในธุรกิจในระบบบริการสุขภาพที่อยู่ใกล้ชิดประชาชน ไม่เพียงแต่ทำหน้าที่ด้านการกระจายยาเท่านั้น แต่ยังมีบทบาทสำคัญ เปรียบเสมือนเป็นที่พึ่งด้านสุขภาพของชุมชน และเป็นอีกหนึ่งทางเลือก ของประชาชน ในการเลือกใช้บริการเมื่อมีอาการเจ็บป่วยเบื้องต้น ไม่รุนแรงถึงขั้นต้องไปโรงพยาบาล เช่น ไข้หวัด ปวดศีรษะ ท้องเสีย ท้องผูก เป็นต้น อย่างไรก็ตามท่ามกลางภาวะเศรษฐกิจในยุคปัจจุบันที่ผู้ประกอบการธุรกิจ ร้านขายยาจะต้องเผชิญการแข่งขันที่สูงขึ้น ทั้งจากการลงทุนของต่างชาติ ซึ่งคิดเป็น ร้อยละ 46.18 ของการลงทุนในตลาดร้านขายยา (ศูนย์วิจัยกสิกรไทย, 2558) และจากการลงทุนของคนไทยเองซึ่งปี พ.ศ. 2561 มีร้านยาที่ขึ้นทะเบียนถูกกฎหมายทั้งหมด 21,840 ร้าน ซึ่งเพิ่มจากปี พ.ศ. 2560 เป็นจำนวน 1,819 ร้าน ส่วนใหญ่เป็นการลงทุนในพื้นที่กรุงเทพมหานคร เป็นสัดส่วนที่สูงถึง ร้อยละ 52.91 (สำนักยา, 2562) สัญญาดังกล่าวนำมาซึ่งโอกาสและความท้าทายของธุรกิจในการวางธุรกิจร้านขายยา ดังนั้น ผู้ประกอบการร้านขายยาจะต้องมีการปรับตัวเพื่อให้สามารถแข่งขันกับร้านยาที่ปริมาณเพิ่มขึ้นเป็นลำดับ เช่น หาช่องทางและสร้างกลยุทธ์การตลาดใหม่ๆ ใหม่ ๆ (ศูนย์วิจัยกสิกรไทย, 2558) หรือการยกระดับร้านขายยาให้มีมาตรฐานตามหลัก GPP ซึ่งคงจะปฏิเสธไม่ได้ว่าการที่ธุรกิจจะสร้างกลยุทธ์ใหม่ หรือการหาช่องทางทางการตลาดใหม่จำเป็นที่จะต้องใช้ข้อมูลเดิมเป็นสิ่งนำทางเพื่อให้ธุรกิจสามารถต่อสู้ และแข่งขันได้อย่างมีประสิทธิภาพ แต่เนื่องจากยาเป็นสินค้าที่ไม่สามารถทำการโฆษณา การเชิญชวน หรือทำการตลาดให้ผู้บริโภคต้องการซื้อหรืออยากซื้อในปริมาณที่มากขึ้นได้ตาม พระราชบัญญัติยา พ.ศ. 2510

จากการศึกษาข้อมูลเกี่ยวกับ ผลิตภัณฑ์เสริมอาหาร ภายใต้ชื่อการค้า “MYHEALTH” ซึ่งเป็นผลิตภัณฑ์เสริมอาหารในกลุ่มเสริมสุขภาพและรักษาโรค ตามการแบ่งกลุ่มของ EuroMonitor พบว่า ธุรกิจอาหารเสริมของไทยมีมูลค่ามากกว่า 6.67 แสนล้านบาท โดยธุรกิจอาหารเสริมนี้แบ่งออกเป็น 3 ประเภท ได้แก่ อาหารเสริมความงาม อาหารเสริมสุขภาพและรักษาโรค และอาหารเสริมเพิ่มสมรรถภาพทางร่างกาย อาหารเสริมสุขภาพ และรักษาโรคเป็นตลาดที่มีมูลค่ามากที่สุด ด้วยมูลค่ากว่า 518,000 ล้านบาท รองลงมาเป็นอาหารเสริมความงาม มีมูลค่าตลาดประมาณ 142,000 ล้านบาท และอันดับสุดท้ายตลาดอาหารเสริมเพิ่มสมรรถภาพทางร่างกายมีมูลค่าตลาด 66,700 ล้านบาท (นริศรัศร ตูลาผลม, 2559) นอกจากนี้ศูนย์วิจัยธนาคารไทยพาณิชย์ ให้ข้อมูลว่า ธุรกิจอาหารเสริมความงามจะเผชิญกับความท้าทายมากขึ้น จากการที่สินค้าเริ่มอึมตัวในกลุ่มผู้บริโภคหลักซึ่งส่วนใหญ่เป็นกลุ่มวัยรุ่น และวัยเริ่มต้นทำงานที่มีอายุต่ำกว่า 30 ปี แต่คนกลุ่มนี้มีแนวโน้มลดลงตามการเปลี่ยนแปลงเข้าสู่ยุค Aging Society ของสังคมไทย ทำให้ตลาดอาหารเสริมความงามที่เติบโตมาระยะหนึ่งนั้นมีการเติบโตที่ลดลง ขณะที่อาหารเสริมสุขภาพ และรักษาโรคที่เป็นกระแสหลักของกลุ่มอาหารเสริมยังคงมีการเติบโตอย่างต่อเนื่อง (นริศรัศร ตูลาผลม, 2559) ดังนั้นผู้ประกอบการจะต้องเข้าใจถึงพฤติกรรมและความต้องการ ของกลุ่มลูกค้าเป้าหมายที่จะเป็นตัวขับเคลื่อนตลาด และสามารถเสนอสิ่งที่ลูกค้าแต่ละรายต้องการได้อย่างตรงจุด (ศูนย์วิจัยกสิกรไทย, 2560)

จากความสำคัญข้างต้น ผู้วิจัยจึงต้องการศึกษาและวิเคราะห์ข้อมูลรายการทางการค้าของบริษัท MIS PHARMA CHAIN จำกัด ที่เกี่ยวข้องกับผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH และนำผลการวิเคราะห์ทั้งด้านการหาโมเดลความสัมพันธ์ในการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อดังกล่าว โดยใช้เทคนิค FP-Growth และแบ่งกลุ่มลูกค้าตามพฤติกรรมการซื้อสินค้ายี่ห้อ MYHEALTH ด้วยเทคนิค RMF เพื่อให้ได้โมเดลที่จะนำมาใช้ให้เกิดความได้เปรียบเชิงกลยุทธ์ และนำองค์ความรู้ไปสู่ตำแหน่งผู้นำในระบบบริการสุขภาพหรือร้านยาในอนาคตต่อไป

โดยการศึกษาครั้งนี้ ผู้จัดทำศึกษาข้อมูลจากมุมมองสองด้าน คือ ด้านธุรกิจ เป็นการศึกษาถึงโครงสร้าง การดำเนินกิจกรรมต่างๆ ของบริษัท และกลยุทธ์ทางธุรกิจขององค์กร ส่วนอีกด้าน คือด้านการนำเทคโนโลยีมาช่วยประกอบการตัดสินใจ เพื่อให้องค์กรขับเคลื่อนด้วยข้อมูล (data-driving)

1.2 วัตถุประสงค์ของการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาถึงเพื่อสร้างแบบจำลองกฎความสัมพันธ์สำหรับฐานข้อมูลการสั่งซื้อสินค้า MYHEALTH ของลูกค้าโดยใช้อัลกอริทึม FP-Growth เพื่อค้นหารูปแบบกฎความสัมพันธ์ของการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH โดยใช้เทคนิคอัลกอริทึม FP-Growth และเพื่อแบ่งกลุ่มลูกค้าจากพฤติกรรมการสั่งซื้อสินค้า MYHEALTH โดยใช้เทคนิค Recency (R) Frequency (F) และ Monetary (M)

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาวิจัย การหาความสัมพันธ์จากฐานข้อมูลการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH ของลูกค้าร้านยาในเครือบริษัท MIS PHARMA CHAIN ผู้วิจัยได้ศึกษา ทฤษฎี และงานวิจัยในอดีตที่เกี่ยวข้องเพื่อเป็นแนวทางในงานวิจัยดังนี้

2.1 ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับพฤติกรรมผู้บริโภคอาหารเสริม

ทฤษฎีและงานวิจัยที่เกี่ยวข้องที่สามารถนำมาอธิบายปัจจัยและความสัมพันธ์ของปัจจัยเพื่อกำหนดรูปแบบการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH ของลูกค้าในเบื้องต้น มีดังนี้

พฤติกรรมผู้บริโภคยาของคนไทย จากการศึกษาของ พิณทิรา ตันเกียรติ (2552) พบว่าคนไทยบริโภคยาโดยผ่านการตัดสินใจหรือคำแนะนำจากผู้ประกอบวิชาชีพแพทย์ เกษัชกร และเจ้าหน้าที่สาธารณสุขอื่นๆ ประมาณ 2 ใน 3 และบริโภคโดยการตัดสินใจของตนเอง โดยอาศัยคำแนะนำจากญาติ เพื่อน หรือการโฆษณาประมาณ 1 ใน 3 โดยมีแนวโน้มที่จะบริโภคผ่านการตัดสินใจและแนะนำจากผู้ประกอบวิชาชีพมากขึ้นเรื่อยๆ

โมเดลพฤติกรรมผู้บริโภค (consumer behavior model) โมเดลพฤติกรรมผู้บริโภคหรือ S-R Theory กล่าวว่า พฤติกรรมของผู้บริโภคจะเกิดจากสิ่งเร้า หรือสิ่งกระตุ้น (stimulus) เพื่อเกิดความต้องการ (need) หรือความรู้สึกนึกคิดของผู้ซื้อ (buyer's black box) ซึ่งจะมีอิทธิพลทำให้เกิดการตอบสนอง (buyer's response) หรือ การตัดสินใจของผู้ซื้อ (buyer's purchase decision) กล่าวโดยสรุปโมเดลพฤติกรรมผู้บริโภคจะประกอบด้วย 3 ส่วนสำคัญ (Kotler, 1997) ได้แก่ (1) สิ่งกระตุ้นทั้งภายในและภายนอก นักการตลาดจะสนใจ เน้นการสร้างสิ่งกระตุ้นทางการตลาดซึ่งควบคุมได้ และสิ่งกระตุ้นอื่นที่ควบคุมไม่ได้ (2) ความรู้สึกนึกคิดของผู้ซื้อ เปรียบเสมือนกล่องดำซึ่งผู้ผลิตหรือผู้ขายไม่สามารถทราบได้ ต้องพยายามค้นหาความรู้สึกนึกคิดของผู้ซื้อ ที่ได้รับอิทธิพลจากลักษณะของผู้ซื้อ และกระบวนการตัดสินใจของผู้ซื้อ (3) การตอบสนอง การตอบสนองหรือการตัดสินใจซื้อของผู้ซื้อหรือ ผู้บริโภคจะมีการตัดสินใจในประเด็นต่างๆ

ทฤษฎีการเรียนรู้ทางปัญญาสังคม (social cognitive theory) เป็นโมเดลที่ทำให้เข้าใจและพยากรณ์พฤติกรรมของมนุษย์ โดยทฤษฎีนี้กล่าวว่า พฤติกรรมของคนเกิดจากการปฏิสัมพันธ์ของสามปัจจัยคือ บุคคล (personal) พฤติกรรม (behavioral) และสภาพแวดล้อม (environmental) (Bandura, 1999) โดยปัจจัยบุคคล ประกอบด้วย ความรู้ ความเข้าใจ (cognition) อารมณ์ (affective) และลักษณะทางชีวภาพ (biological events) ส่วนปัจจัยพฤติกรรมหมายถึง การตอบสนองต่อสภาพแวดล้อม ซึ่งการตอบสนองนี้เป็นการกระทำที่ขึ้นอยู่กับแรงจูงใจและทัศนคติ ทำได้ดีที่สุด ปัจจัยสภาพแวดล้อม ประกอบด้วย สภาพของที่ทำงานและเงื่อนไขต่างๆ เช่น งบประมาณและระยะเวลาการฝึกอบรม เป็นต้น นอกจากนี้ Bandura (1999) ยังกล่าวว่าปัจจัยที่สำคัญที่อยู่ภายในปัจเจกบุคคล ได้แก่ ความเชื่อในความสามารถของตนเอง (self-efficacy beliefs) ความคาดหวังต่อผลลัพธ์ (outcome expectations) และการควบคุมการเรียนรู้ด้วยตนเอง (self-regulated learning)

2.2 แนวคิดเบื้องต้นของเหมืองข้อมูล

เทคนิคการทำเหมืองข้อมูล (data mining) คือ ขั้นตอนการวิเคราะห์หรือสกัดข้อมูลจากฐานข้อมูลขนาดใหญ่หรือหลายคนอาจเข้าใจว่าเป็นขั้นตอนหนึ่งของกระบวนการค้นหาความรู้ที่เป็นประโยชน์จากฐานข้อมูลหรือที่เรียกว่า Knowledge Discovery in Databases (KDD) (Han, 2012) โดยเทคนิคการวิเคราะห์เหมืองข้อมูลแบ่งออกเป็น 2 ประเภทหลัก คือ (1) การวิเคราะห์เหมืองข้อมูลเพื่อการอธิบาย (descriptive) ซึ่งเป็นการเรียนรู้แบบไม่มีผู้สอน

(unsupervised learning) เป็นการค้นหารูปแบบความสัมพันธ์หรือลักษณะที่เชื่อมโยงกันของข้อมูล ให้โปรแกรมวิเคราะห์หารูปแบบที่น่าสนใจจากข้อมูล เทคนิคการวิเคราะห์ได้แก่ การจัดกลุ่ม (cluster) และการหากฎความสัมพันธ์ (association rule) (2) การวิเคราะห์เหมืองข้อมูลเพื่อการทำนาย (predictive) เป็นการเรียนรู้แบบมีการสอน (supervised learning) เป็นการเรียนรู้จากข้อมูลเดิมที่มีอยู่ สร้างโมเดลเพื่อทำนายผลของข้อมูลใหม่หรือการทำนายข้อมูลใหม่ที่จะเกิดในอนาคต เทคนิคการวิเคราะห์ได้แก่ การจำแนกประเภท (classification) โดยวิธีที่นิยมใช้ได้แก่ การวิเคราะห์ต้นไม้ตัดสินใจ (decision tree) และการวิเคราะห์เครือข่ายประสาท (neuron networks) (Dangeti, 2017)

ในส่วนของงานวิจัยที่เกี่ยวข้องนี้ งานวิจัยนี้จะจัดกลุ่มงานวิจัยที่เกี่ยวข้องโดยจัดกลุ่มตามเทคนิคการวิเคราะห์เหมืองข้อมูล ซึ่งงานวิจัยในอดีตที่เกี่ยวกับพฤติกรรมผู้บริโภคของลูกค้าเทคนิคการวิเคราะห์เหมืองข้อมูลที่ใช้ในการวิเคราะห์ข้อมูลมีดังนี้

การจัดกลุ่ม (cluster analysis) หรือ Segmentation หรือ Partitioning (Tampakis, 2019) เป็นเทคนิคการวิเคราะห์เหมืองข้อมูลในลักษณะการเรียนรู้แบบไม่มีผู้สอน (unsupervised learning) โดยการจัดกลุ่มจะจัดข้อมูลที่มีลักษณะคล้ายกันจัดเข้ากลุ่มเดียวกัน ต่างกันอยู่คนละกลุ่ม การจัดกลุ่มนี้สามารถทำได้หลายวิธี เช่น Partitioning methods (K-means, K-medoids), Hierarchical Agglomerative methods (single-link, complete-link, average-link), Density based methods, Grid-based methods, Model-based method (Bharati, 2010) เช่น การแบ่งผู้บริโภคหรือลูกค้าตามพฤติกรรมผู้บริโภคสินค้าต่าง ๆ โดยให้ลูกค้าที่มีพฤติกรรมผู้บริโภคหรือการซื้อสินค้าที่คล้ายกันอยู่ในกลุ่มเดียวกัน ส่วนลูกค้าที่มีพฤติกรรมผู้บริโภคต่างกันจะอยู่ต่างกลุ่มกัน เมื่อจัดกลุ่มแล้วจะทำให้สามารถวางแผนกลยุทธ์ทางการตลาดสำหรับลูกค้าแต่ละกลุ่มได้อย่างมีประสิทธิภาพ ตัวแปรที่นำมาใช้ในการจัดกลุ่มอาจใช้ตัวแปรด้านพฤติกรรมต่าง ๆ ของลูกค้า (อำนาจ มณีศรีวงศ์กุล, 2541)

การค้นหากฎความสัมพันธ์ของข้อมูล (association analysis) หรือบางครั้งอาจเรียกเทคนิคนี้ว่า การวิเคราะห์ตะกร้าตลาด (market basket analysis) เป็นเทคนิคการวิเคราะห์เหมืองข้อมูลในลักษณะการเรียนรู้แบบไม่มีผู้สอน (unsupervised learning) เช่นเดียวกับการจัดกลุ่ม โดยการค้นหากฎความสัมพันธ์ของข้อมูลมีการวัดความแม่นยำของกฎความสัมพันธ์ด้วยค่าสนับสนุน (support) และค่าความมั่นใจ (confidence) เป็นเกณฑ์ในการคัดเลือกรูปแบบทั่วไปของการค้นหากฎความสัมพันธ์ซึ่งจะอยู่ในรูปแบบ $A \Rightarrow B$ โดยที่ A, B เป็นเซตของ ไอเทม (Itemset) ที่ประกอบอยู่ในฐานข้อมูล โดยค่าสนับสนุนจะวัดความน่าจะเป็นของจำนวนรายการของข้อมูลที่เกิดร่วมกันเทียบกับจำนวนรายการทั้งหมด ส่วนค่าความมั่นใจจะวัดความน่าจะเป็นเมื่อเกิดเหตุการณ์หนึ่ง (A) แล้วจะเกิด อีกเหตุการณ์หนึ่งตามมา (B) ในการหากฎความสัมพันธ์นั้นจะมีขั้นตอนวิธีการหาหลายวิธีด้วยกัน ขั้นตอนวิธีที่เป็นที่รู้จักและถูกใช้อย่างแพร่หลายคือ อัลกอริทึมเอพริออริ (apriori algorithm) และ FP-Growth (FP-Growth algorithm) (Chen, Han, and YU, 1996)

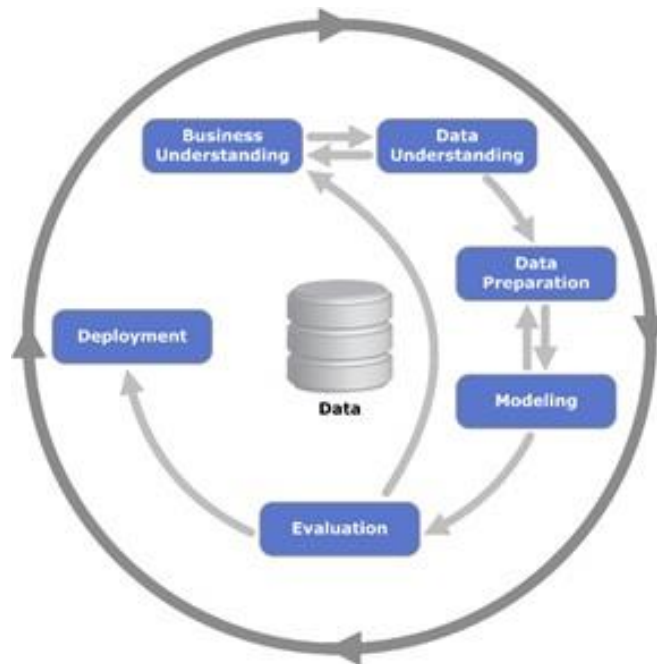
การจำแนกประเภท (classification) การจำแนกประเภท เป็นเทคนิคการวิเคราะห์เหมืองข้อมูล ที่นิยมใช้กันมากที่สุด (Bharati, 2010) ซึ่งเป็นการเรียนรู้แบบมีผู้สอน (supervised learning) คือการนำข้อมูลในอดีตมาสร้างตัวแบบเพื่อทำนายอนาคตโดยมีการใช้ข้อมูลฝึกหัด (training data) ซึ่งข้อมูลทุกตัวจะมีคุณสมบัติที่ใช้ในการทำนาย อัลกอริทึมประเภทนี้จะมุ่งเน้นการแบ่งแยกข้อมูลออกเป็นกลุ่มตามค่าคุณสมบัติของข้อมูล ซึ่งถ้าค่าคุณสมบัติของข้อมูลมีค่าไม่ต่อเนื่อง จะเรียกกระบวนการที่ใช้แบ่งแยกว่า การจำแนกประเภท (classification) แต่ถ้าคุณสมบัติของข้อมูลมีค่าต่อเนื่องจะเรียกกระบวนการที่ใช้แบ่งแยกว่า การถดถอย (regression) หรือ การพยากรณ์ (forecasting) โดยเทคนิคที่ใช้ได้แก่ KNN, Decision tree, Neuron network, Support vector machine, Regression และ Logistic regression (Tan, 2019)

Recency Frequency Monetary Analysis (RFM analysis) การวิเคราะห์แบบ RFM เป็นเทคนิคหนึ่งที่ยิมนำมาประยุกต์ใช้กับงานด้านการตลาดแบบทางตรง เช่น การแบ่งกลุ่มลูกค้าตามพฤติกรรมการซื้อสินค้าของลูกค้า โดยพิจารณาจากข้อมูล 3 ค่า ดังนี้ Recency หมายถึง ความเป็นปัจจุบัน โดยพิจารณาจากระยะเวลา (จำนวนวัน) ที่ซื้อสินค้าล่าสุด Frequency หมายถึง ความถี่ โดยพิจารณาจากความถี่ของการซื้อสินค้าว่ามีมากน้อยเพียงใด และ Monetary หมายถึง มูลค่าในการซื้อ หรือปริมาณการใช้จ่ายของลูกค้า โดยจะทำการแบ่งค่า Recency (R) Frequency

(F) และ Monetary (M) ออกเป็นส่วนๆ เท่าๆ กัน เช่น ถ้าแบ่งออกเป็น 5 ส่วน โดยเลขที่มีค่ามากที่สุด คือ เลข 5 จะมีความสำคัญมากที่สุด และใช้ค่าตัวเลข 3 หลักมาแทนค่าข้อมูลในแต่ละกลุ่ม เช่น กลุ่ม 555 หมายถึง กลุ่มที่มีค่า $R = 5$, $F = 5$ และ $M = 5$ ซึ่งหมายถึงกลุ่มของลูกค้าที่มีพฤติกรรมการซื้อสินค้าไม่นานนี้ มีการซื้อบ่อย และมีค่าใช้จ่ายที่สูง โดยการแบ่งกลุ่มลูกค้าด้วยวิธี RFM มีการจัดเรียงลำดับข้อมูล เช่น Recency ทำการเรียงจากน้อยไปหามาก ค่าน้อยหมายถึง ลูกค้ามาซื้อสินค้าเมื่อไม่นานมานี้ ทำให้ทราบว่าเป็นลูกค้าที่มาซื้อสินค้าล่าสุดในช่วง เวลานานแล้วหรือไม่ Frequency ทำการเรียงค่าจากมากไปหาน้อย ลูกค้าที่มีค่า Frequency สูง แสดงว่ามีความถี่ในการซื้อสินค้าบ่อยหรือซื้อสินค้าเป็นประจำ และเรียงค่า Monetary จากมากไปหาน้อย ลูกค้าที่มีค่า Monetary สูง แสดงว่ามียอดการใช้จ่ายในการซื้อสินค้าในปริมาณมาก (เอกสิทธิ์ พืชรวงศ์ศักดิ์ดา, 2557) RFM เป็นวิธีที่มีประโยชน์ในการแบ่งกลุ่มลูกค้าตามพฤติกรรมการซื้อ เพื่อช่วยในการให้บริการหรือสามารถส่งเสริมการขายได้เฉพาะเจาะจงกับลูกค้ามากขึ้น (Birant, 2011)

3. วิธีการวิจัย

งานวิจัยนี้เป็นงานวิจัยเชิงปริมาณ เพื่อค้นหารูปแบบกฎความสัมพันธ์ของการซื้อผลิตภัณฑ์โดยใช้เทคนิคอัลกอริทึม FP-Growth แบ่งกลุ่มลูกค้าจากพฤติกรรมการซื้อสินค้าโดยใช้เทคนิค RFM โดยมีขั้นตอนในการดำเนินการตามแบบ Cross-Industry Standard Process for Data Mining (CRISP-DM) ซึ่งเป็นกระบวนการมาตรฐานสำหรับการทำเหมืองข้อมูล ดังภาพที่ 1 ดังนี้



ภาพที่ 1 ขั้นตอนในการดำเนินการตามแบบ CRISP-DM

ที่มา: Jensen (2012)

3.1 การทำความเข้าใจธุรกิจ (business understanding)

ในการทำความเข้าใจธุรกิจนั้น งานวิจัยนี้จะกล่าวถึงภาพรวมของบริษัท กลยุทธ์และเป้าหมายขององค์กรที่มีต่อผลิตภัณฑ์เสริมอาหาร MYHEALTH และความต้องการของผู้บริหาร

3.1.1 ภาพรวมของบริษัท

บริษัท MIS PHARMA CHAIN จำกัด ก่อตั้งขึ้นในปี พ.ศ.2540 ประกอบธุรกิจให้บริการขายยาแผนปัจจุบัน ผลิตภัณฑ์เสริมอาหาร เวชสำอางค์และอุปกรณ์การแพทย์ครบวงจร โดยให้บริการภายใต้แบรนด์ “MIS PHARMA CHAIN” , “MIS CLUB” , “MIS DRUG” โดยกลุ่มของผลิตภัณฑ์ สามารถแบ่งเป็น 5 กลุ่มหลัก ได้แก่

- (1) ยา
- (2) อาหารเสริม
- (3) อุปกรณ์การแพทย์
- (4) ผลิตภัณฑ์ดูแลและบำรุงผิว
- (5) อื่นๆ

จากแนวโน้มการเติบโตของตลาดผลิตภัณฑ์อาหารเสริม ที่ผู้บริหารได้มีการจัดตั้ง บริษัท 123 จำกัด ในปี พ.ศ. 2558 เพื่อจัดจำหน่ายผลิตภัณฑ์เสริมอาหารภายใต้ชื่อการค้า “MYHEALTH” ภายใต้คอนเซ็ปต์ สุขภาพดี เริ่มต้นที่ MYHEALTH ซึ่งปัจจุบันมีสินค้าทั้งหมด 13 ชนิด 15 ไอเทม

ปัจจุบัน MYHEALTH มีวางจำหน่ายเฉพาะร้านยาในเครือบริษัท MIS PHARMA CHAIN จำกัด เช่น MIS PHARMA CHAIN, MIS CLUB, MIS DRUG และ MIS MAX นอกจากนี้ยังมีการจำหน่ายผ่านช่องทางเฟซบุ๊ก (Facebook) และช้อปปี้ (Shopee) เพื่อตอบสนองไลฟ์สไตล์ของลูกค้าในยุคปัจจุบัน และเป็นช่องทางโปรโมทสินค้าให้ลูกค้าได้รู้จักอีกด้วย

3.1.2 กลยุทธ์และเป้าหมายขององค์กรที่มีต่อผลิตภัณฑ์เสริมอาหาร MYHEALTH

เป้าหมายระยะสั้น เน้นการเพิ่มจำนวนผลิตภัณฑ์ให้มีความหลากหลาย ครอบคลุมความต้องการของผู้บริโภค สร้างพนักงานขายที่มีความรู้เกี่ยวกับสารสำคัญทุกตัวที่อยู่ในผลิตภัณฑ์แต่ละชนิดเพื่อเพิ่มความน่าเชื่อถือให้แบรนด์และสอดคล้องกับเป้าหมายใหญ่ขององค์กรคือ ต้องการให้ร้านยาในเครือบริษัทเป็นที่พึ่งและเป็นส่วนหนึ่งของชุมชนในบริเวณนั้น

เป้าหมายระยะยาว (5 ปีขึ้นไป) ขยายตลาดหรือกลุ่มลูกค้าให้เพิ่มมากขึ้น สร้างแบรนด์ของผลิตภัณฑ์ให้เป็นที่รู้จัก เข้าใจพฤติกรรมการณ์ซื้อผลิตภัณฑ์เสริมอาหาร สามารถจัดโปรโมชั่นส่งเสริมการขาย แยกตามสาขาได้ สามารถจัดการส่งเสริมการขายทั้งออฟไลน์ (Offline) และออนไลน์ (Online) ได้อย่างมีประสิทธิภาพ

3.1.3 ความต้องการของผู้บริหาร

จากการสัมภาษณ์ กรรมการผู้จัดการ บริษัท MIS PHARMA CHAIN จำกัด เมื่อวันที่ 6 กรกฎาคม พ.ศ. 2562 ทำให้ทราบว่า ผู้บริหารมีความต้องการ

- (1) ทำ Promotion แยกให้เหมาะสมกับลูกค้าแต่ละสาขาได้
- (2) อยากทราบรูปแบบการซื้อผลิตภัณฑ์เสริมอาหารของลูกค้า เพื่อไปผลิตสินค้าตัวใหม่ที่ตรงกับความต้องการของผู้บริโภค
- (3) อยากทราบและเข้าใจพฤติกรรมกลุ่มลูกค้าเป้าหมายแต่ละกลุ่ม เพื่อนำเสนอสิ่งที่ลูกค้าแต่ละกลุ่มต้องการ
- (4) สามารถตรวจสอบได้ว่าลูกค้าท่านใดมีแนวโน้มที่จะเลิกใช้ผลิตภัณฑ์ และสามารถตอบสนองสิ่งที่ลูกค้าเหล่านั้นต้องการได้ทัน
- (5) มียอดจำหน่ายภายในร้านขายยาในเครือบริษัท สูงกว่า GoodHealth ภายใน 2 ปี

3.2 การทำความเข้าใจเกี่ยวกับข้อมูล (data Understanding)

3.2.1 การเก็บรวบรวมข้อมูลขั้นต้น (initial data collection)

ข้อมูลที่น่ามาใช้ในการสร้างแบบจำลองจะใช้ข้อมูลการซื้อขายสินค้าของลูกค้า (transaction data) ของบริษัท MIS PHARMA CHAIN ที่จัดเก็บใน Microsoft SQL Server และได้รับอนุญาตจากผู้บริหารของบริษัทให้นำมาใช้ในการวิเคราะห์ได้ โดยผู้วิจัยให้เจ้าหน้าที่ที่ดูแลระบบเป็นผู้ดึงข้อมูลมาจากฐานข้อมูลและบันทึกเป็น excel ทั้งหมด 3 ไฟล์ ประกอบด้วย ข้อมูลการขายสินค้า ข้อมูลลูกค้า และข้อมูลสินค้า

3.2.2 การอธิบายข้อมูล (describe data)

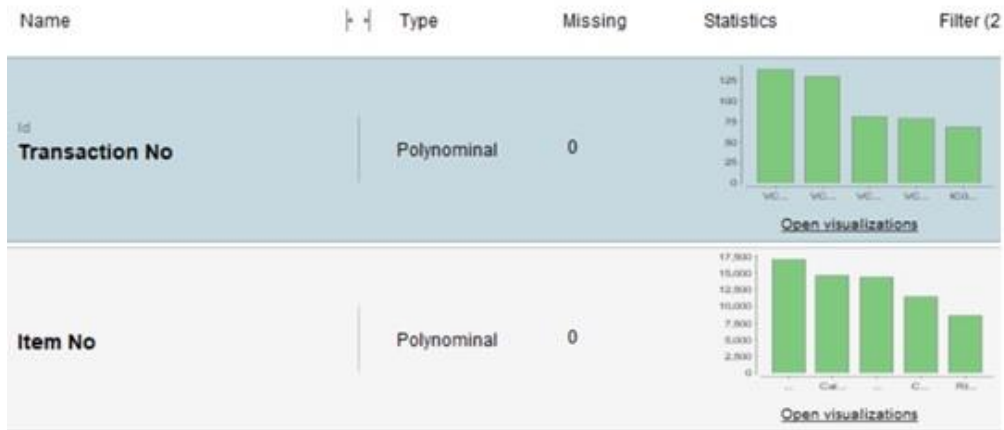
ข้อมูลการขายสินค้า ข้อมูลลูกค้า และข้อมูลสินค้า ข้างต้น ประกอบด้วยรายละเอียดดังแสดงในตารางที่ 1 ซึ่งแสดงถึงชื่อฟิลด์ คำอธิบาย ชนิดข้อมูล และชื่อแฟ้มข้อมูลใน Microsoft SQL Server

ตารางที่ 1 รายละเอียดข้อมูลที่น่ามาจัดทำเหมืองข้อมูล

ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	ชื่อแฟ้มข้อมูลใน Microsoft SQL Server
Transaction No	เลขที่ทำการซื้อขาย	Integer	Trans Sales Entry
Store No	รหัสสาขา	nvarchar(10)	Trans Sales Entry
POS Terminal No	รหัสเครื่องชำระเงิน	nvarchar(10)	Trans Sales Entry
Date	วันที่ทำการซื้อขาย	datetime	Trans Sales Entry
Time	เวลาที่ทำการซื้อขาย	datetime	Trans Sales Entry
Item No	รหัสสินค้า	nvarchar(20)	Trans Sales Entry
Price per unit	ราคาสินค้าต่อชิ้น	decimal(38,20)	Trans Sales Entry
Quantity	จำนวนสินค้าที่ลูกค้าซื้อ	decimal(38,20)	Trans Sales Entry
Member Card No	รหัสลูกค้า	nvarchar(30)	Transaction Header
Description	ชื่อสินค้า	nvarchar(50)	Item

3.2.3 การตรวจสอบคุณภาพของข้อมูล (data quality verification)

จากการตรวจสอบในเบื้องต้น ข้อมูลที่สำคัญและจำเป็นต่อการทำเหมืองข้อมูล มีความสมบูรณ์ประมาณ 95% มีขาดหายเพียง รหัสลูกค้า (member card no) ในบางรายการขาย ซึ่งหมายถึงรายการขายนั้นถูกซื้อไปโดยลูกค้าที่ไม่มีสมาชิกกับทางบริษัท และจากภาพที่ 2 แสดงให้เห็นว่าข้อมูลทั้งหมด 93,278 แถว 2 คอลัมน์ แต่ละคอลัมน์ไม่มีค่าขาดหาย และอยู่ในรูปแบบที่เหมาะสมต่อการนำทำเหมืองข้อมูล แบบการหากฎความสัมพันธ์



ภาพที่ 2 ความสมบูรณ์ของข้อมูลสำหรับการหาความสัมพันธ์

ภาพที่ 3 แสดงให้เห็นว่าข้อมูลในคอลัมน์ รหัสสมาชิก (member card no) มีค่าขาดหาย ทั้งหมด 29,418 แถว และข้อมูลทุกคอลัมน์อยู่ในรูปแบบที่เหมาะสมต่อการนำทำเหมืองข้อมูล แบบการจัดกลุ่ม โดยใช้เทคนิค RFM เนื่องจากการจัดกลุ่มลูกค้าเป็นการจัดกลุ่มลูกค้าที่เป็นสมาชิกกับทางร้านอยู่แล้ว ส่วนค่าขาดหายของ Member card no เป็นลูกค้าที่ไม่ได้เป็นสมาชิก ดังนั้นสามารถตัดค่าขาดหายนี้ออกได้โดยไม่ส่งผลต่อผลการวิเคราะห์ข้อมูล



ภาพที่ 3 ความสมบูรณ์ของข้อมูลสำหรับการจัดกลุ่มโดยใช้เทคนิค RFM

3.3 การเตรียมข้อมูล (data preparation)

การเตรียมข้อมูลเป็นการจัดข้อมูลเพื่อให้อยู่ในรูปแบบที่สามารถนำไปหากฎความสัมพันธ์และจัดกลุ่มลูกค้าด้วยโปรแกรมเหมืองข้อมูลได้ ดังนี้

3.3.1 การคัดเลือกข้อมูลสำหรับการหากฎความสัมพันธ์ (association rules)

การหากฎความสัมพันธ์จะใช้ข้อมูลจากการซื้อสินค้าแต่ละครั้ง นำมาแปลงข้อมูลมาจัดกลุ่มข้อมูลตามคอลัมน์ Transaction no (group by transaction no) จากนั้นนำข้อมูลที่ได้ไปหารูปแบบของการซื้อสินค้าที่เกิดขึ้นร่วมกันบ่อยๆ และนำไปสร้างกฎความสัมพันธ์ (association rule) โดยข้อมูลหลักที่นำมาใช้สำหรับหากฎความสัมพันธ์สำหรับงานวิจัยนี้ คือ Transaction no หรือเลขที่ทำการซื้อขาย และ Item No หรือรหัสสินค้าที่ลูกค้าซื้อในใบเสร็จนั้นๆ ดังตารางที่ 1

ตารางที่ 1 ข้อมูลหลักสำหรับการหากฎความสัมพันธ์ของการขายผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH

ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	ชื่อเพิ่มข้อมูลใน Microsoft SQL Server
Transaction No	เลขที่ทำการซื้อขาย	Integer	Trans Sales Entry
Item No	รหัสสินค้า	nvarchar(20)	Trans Sales Entry

3.3.2 การคัดเลือกข้อมูลสำหรับการจัดกลุ่มลูกค้าตาม RFM model (clustering by RFM model)

การจัดกลุ่มลูกค้าตามพฤติกรรมกรรมการซื้อสินค้าโดยใช้ โมเดล RFM คือ Recency ระยะเวลาล่าสุดที่ลูกค้ามาซื้อสินค้า Frequency ความถี่ในการมาซื้อสินค้า และ Monetary จำนวนเงินที่ลูกค้าใช้จ่ายในการซื้อสินค้า ดังนั้นข้อมูลที่นำมาใช้ในการวิเคราะห์สำหรับงานวิจัยนี้ คือ เลขที่ทำการซื้อขาย วันที่ที่ลูกค้ามาซื้อสินค้าเพื่อใช้ในการคำนวณหา Recency รหัสสินค้าที่ลูกค้าซื้อ จำนวนสินค้า และราคาสินค้าที่ซื้อเพื่อใช้ในการคำนวณหา Monetary และรหัสลูกค้าเพื่อใช้เป็นคีย์หลักและเพื่อหา Frequency ดังตาราง 2

ตารางที่ 2 ข้อมูลหลักสำหรับการจัดกลุ่มลูกค้าที่ซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH

ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	ชื่อเพิ่มข้อมูลใน Microsoft SQL Server
Transaction No	เลขที่ทำการซื้อขาย	Integer	Trans Sales Entry
Item No	รหัสสินค้า	nvarchar(20)	Trans Sales Entry
Date	วันที่ทำการซื้อขาย	datetime	Trans Sales Entry
Time	เวลาที่ทำการซื้อขาย	datetime	Trans Sales Entry
Price per unit	ราคาสินค้าต่อชิ้น	decimal(38,20)	Trans Sales Entry
Quantity	จำนวนสินค้าที่ลูกค้าซื้อ	decimal(38,20)	Trans Sales Entry
Member Card No	รหัสลูกค้า	nvarchar(30)	Transaction Header

3.4 การสร้างโมเดลและเลือกเทคนิคที่เหมาะสม (modeling)

ขั้นตอนต่อไปนี้จะเป็นการเตรียมพร้อมข้อมูลให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ข้อมูลได้ จนถึงขั้นการสร้างแบบจำลอง โดยผู้วิจัยแบ่งออกเป็น 2 ส่วน คือ

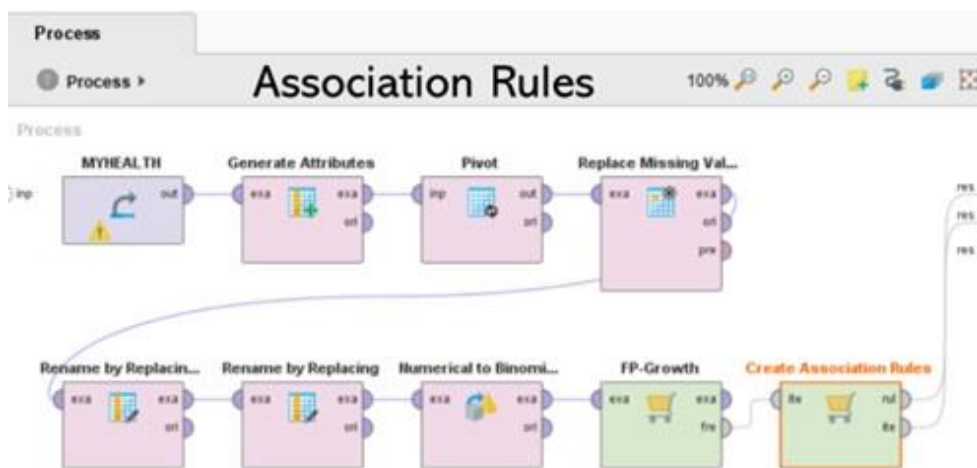
3.4.1 ขั้นตอนของการหาความสัมพันธ์

ในการศึกษาวิจัยเลือกการทำเหมืองข้อมูลแบบการค้นหาความสัมพันธ์ระหว่างข้อมูลด้วยเทคนิค FP-Growth เพื่อใช้วางแผนสำหรับการจัดทำโปรโมชั่นหรือพัฒนาสินค้าให้ตรงตามความต้องการของลูกค้า และการแบ่งกลุ่มลูกค้าตามพฤติกรรมการซื้อด้วยเทคนิค K-Means เนื่องจากมีข้อดีหลายประการ ได้แก่

- (1) เป็นวิธีพื้นฐานที่ใช้กันอย่างแพร่หลาย
- (2) เวลาที่ใช้ในการเรียนรู้สั้นกว่าวิธีอื่น
- (3) ผลลัพธ์ที่ได้เป็นกฎและความสัมพันธ์ที่เข้าใจได้ง่าย
- (4) ให้ผลตรงกับวัตถุประสงค์การทำเหมืองข้อมูล เมื่อเทียบกับวิธีอื่น

ขั้นตอนในการค้นหาความสัมพันธ์ระหว่างข้อมูลด้วยเทคนิค FP-Growth สำหรับงานวิจัยนี้ (ภาพที่ 4) มีดังนี้

1. สร้าง Column เพิ่ม 1 column และกำหนดค่าของ column ให้มีค่าเป็น 1 เพื่อแสดงว่ามีการขายสินค้านั้นเกิดขึ้นใน Transaction นั้น
2. ทำการ Group transaction ID คือเปลี่ยนจาก POS database ให้เป็น Transaction database เพื่อแสดงให้เห็นว่าในแต่ละ Transaction มีการซื้อสินค้า MYHEALTH ชนิดใดบ้างและไม่ซื้อชนิดใดบ้าง
3. ทำการหารูปแบบของข้อมูลที่เกิดการซื้อพร้อมกันบ่อยๆ (frequent itemset) และให้โปรแกรมทำการหา minimum support ให้
4. นำ Frequent itemset มาหา กฎความสัมพันธ์ และกำหนดค่า Minimum confidence ที่ 70% คือความสัมพันธ์นี้มีโอกาสเกิดขึ้นตั้งแต่ 70% ขึ้นไปจะนำมาพิจารณาแต่อย่างไรก็ตาม ความสัมพันธ์ที่จะได้รับคัดเลือกหรือพิจารณาก่อนจะเป็นความสัมพันธ์ที่มีค่า Confidence สูงที่สุด



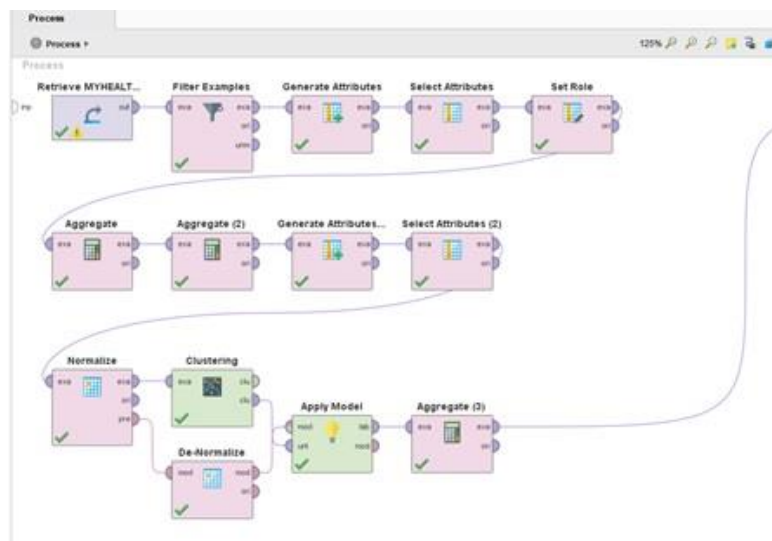
ภาพที่ 4 ภาพรวมขั้นตอนการสร้างกฎความสัมพันธ์

3.4.2 ขั้นตอนการแบ่งกลุ่มลูกค้าและการหาความสัมพันธ์ของแต่ละกลุ่ม

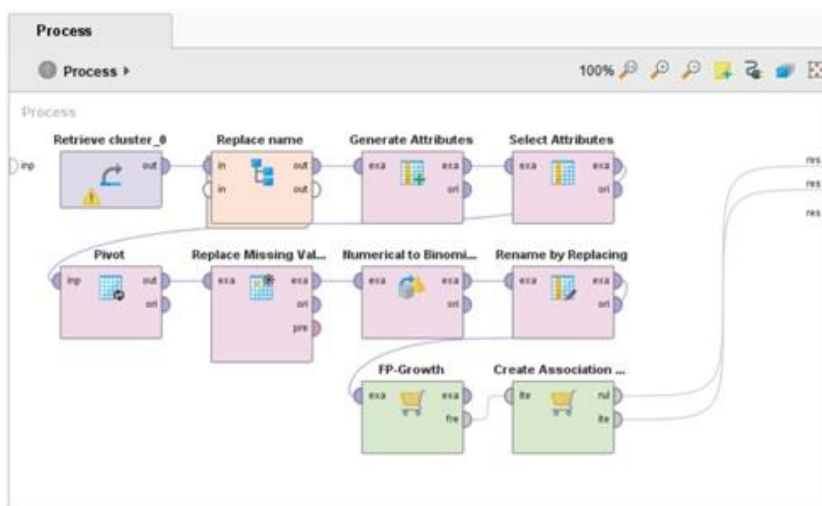
มีดังนี้

ขั้นตอนการแบ่งกลุ่มลูกค้าและการหาความสัมพันธ์ของแต่ละกลุ่ม สำหรับงานวิจัยนี้ (ภาพที่ 5 และ 6)

1. เนื่องจากข้อมูลที่นำมาจาก Microsoft SQL Server จะมีข้อมูลของลูกค้าที่ไม่เป็นสมาชิกของทางร้าน ซึ่งไม่มีรหัสสมาชิก งานวิจัยจึงนำข้อมูลลูกค้าที่ไม่มีรหัสสมาชิกออกจากการวิเคราะห์ โดยงานวิจัยนี้จะจัดกลุ่มลูกค้าเฉพาะลูกค้าที่มีสมาชิกกับทางร้านเท่านั้นซึ่งเป็นไปตามวัตถุประสงค์ของการวิจัย
2. สร้างคอลัมน์ TotalPrice เพื่อคำนวณยอดใช้จ่ายแต่ละ Transaction
3. แบ่งกลุ่มลูกค้าเป็น 6 กลุ่มด้วย เทคนิค K-Mean และหาค่าเฉลี่ยของ RFM แต่ละกลุ่มลูกค้า
4. กรองลูกค้าแต่ละกลุ่มและหาความสัมพันธ์ของลูกค้าแต่ละกลุ่ม



ภาพที่ 5 ภาพรวมขั้นตอนการแบ่งกลุ่มลูกค้าด้วยเทคนิค RFM



ภาพที่ 6 ภาพรวมขั้นตอนการสร้างความสัมพันธ์ในแต่ละกลุ่มลูกค้า

3.5 การตรวจสอบและประเมินผล (evaluation)

เนื่องจากงานวิจัยนี้ทำการค้นหาความสัมพันธ์ของการซื้อผลิตภัณฑ์เสริมอาหาร ยี่ห้อ MYHEALTH และการแบ่งกลุ่มลูกค้าตามพฤติกรรมการซื้อ ดังนั้นตัวชี้วัดประสิทธิภาพของกฎความสัมพันธ์จะตรวจสอบได้จากค่าความเชื่อมั่น (confidence) งานวิจัยนี้ได้กำหนดค่าความเชื่อมั่นที่ยอมรับได้ต้องมีค่ามากกว่า 70% และค่า lift ที่มีค่ามากกว่า 1 ส่วนการแบ่งกลุ่มตามพฤติกรรมการซื้อจะใช้ Brute force อัลกอริทึม คือให้โปรแกรมวนลูปหาค่า K ต่างๆ และเลือกค่า K ที่เหมาะสมที่สุดจาก Elbow method และพบว่าค่า K ที่เหมาะสมที่สุดคือ 6 ผลจากการวิเคราะห์ข้อมูลพบว่าค่าความเชื่อมั่น ค่า lift และค่า K เป็นไปตามเกณฑ์ที่กำหนด ดังนั้นสามารถกล่าวได้ว่าโมเดลของงานวิจัยนี้มีความน่าเชื่อถือ

3.6 การนำแบบจำลองไปใช้ (deployment)

การนำแบบจำลองไปใช้ ประกอบด้วย การวางแผนการปรับใช้แบบจำลอง และแผนการติดตามและการบำรุงรักษา ดังนี้

3.6.1 การวางแผนการปรับใช้แบบจำลอง (plan deployment)

การวิจัยครั้งนี้ได้นำเทคนิคการทำเหมืองข้อมูลมาประยุกต์ใช้ในการวิเคราะห์หากฎความสัมพันธ์ของการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ “MYHEALTH” และวิเคราะห์พฤติกรรมการซื้อโดยพิจารณาจาก ระยะเวลาล่าสุด ความถี่ และยอดใช้จ่ายของลูกค้า ซึ่งถือว่ามีผลสำคัญต่อองค์กร เนื่องจากการทราบความต้องการของลูกค้าจะทำให้สามารถนำเสนอโปรโมชั่นหรือสินค้าที่ตรงกับความต้องการของลูกค้าแต่ละกลุ่มได้ โดยรายละเอียดข้อเสนอแนะจะกล่าวต่อไป

3.6.2 แผนการติดตามและการบำรุงรักษา (plan monitoring and maintenance)

จากผลการวิเคราะห์ข้อมูลทำให้ทราบถึงสิ่งที่ควรติดตามและการบำรุงรักษาระบบสำหรับบริษัท MIS PHARMA CHAIN จำกัด ดังนี้

- (1) ควรเชื่อมโยงข้อมูลจากฐานข้อมูลเข้าสู่โปรแกรมเพื่อวิเคราะห์ข้อมูลในลักษณะ Online นอกจากนี้ควรจัดรูปแบบของผลลัพธ์ (Visualization) ให้ผู้บริหาร สามารถดูผลลัพธ์เพื่อใช้ในการตัดสินใจได้รวดเร็วขึ้น
- (2) เสนอแนะให้มีการจัดเก็บข้อมูลบางส่วนเพิ่มเติม เพื่อให้สามารถวิเคราะห์ข้อมูลในลักษณะ Classification ได้มากขึ้น โดยรายละเอียดข้อมูลที่ควรจัดเก็บเพิ่มเติมจะกล่าวต่อไป
- (3) ควรเพิ่มตำแหน่งเจ้าหน้าที่วิเคราะห์ข้อมูลเพื่อติดตามข้อมูลได้อย่างเหมาะสมต่อไป

4. สรุปผลการศึกษา

4.1 อภิปรายผลการศึกษา

ในการจัดทำวิจัยนี้ นักวิจัยได้ประยุกต์ใช้กระบวนการทำเหมืองข้อมูลแบบ CRISP-DM มาเป็นแนวทางในการวิจัย เนื่องจากเป็นมาตรฐานสากลที่ทั่วโลกยอมรับ โดยงานวิจัยนี้ใช้วิธีการหาความสัมพันธ์ของผลิตภัณฑ์กับอีกผลิตภัณฑ์หนึ่ง เพื่อนำไปสู่การเพิ่มยอดขายและการขายข้ามผลิตภัณฑ์ และใช้ขั้นตอนวิธีการแบ่งกลุ่มเพื่อแบ่งกลุ่มลูกค้า เพื่อที่จะตอบสนองและเสนอโปรโมชั่นได้ตรงจุด และนำมาซึ่งการเพิ่มยอดขาย โดยเริ่มตั้งแต่การทำความเข้าใจกับวัตถุประสงค์ทางธุรกิจ ในการทำเหมืองข้อมูลเพื่อประเมินสถานการณ์ปัจจุบันขององค์กร ศึกษาถึงเงื่อนไขข้อจำกัดต่างๆ และสำรวจแหล่งที่มาของข้อมูลที่จะนำมาศึกษารวบรวมข้อมูลต่างๆ เข้าด้วยกัน

ผลการศึกษาความสัมพันธ์สำหรับผลิตภัณฑ์เสริมอาหาร ยี่ห้อ MYHEALTH ที่ค่า confidence เท่ากับ 70% พบ 7 กฎดังนี้

1. จากรายการซื้อผลิตภัณฑ์หนึ่งพันรายการ พบว่ามี 1 รายการ ที่ หากลูกค้าซื้อ L-Carnitine และ Royal Jelly แล้ว มีโอกาสจะซื้อ Carbo-X ด้วย ถึง 71.4%
2. จากรายการซื้อผลิตภัณฑ์หนึ่งพันรายการ พบว่ามี 1 รายการ ที่ หากลูกค้าซื้อ Rice Bran Oil และ Carbo-X แล้ว มีโอกาสจะซื้อ L-Carnitine ด้วย ถึง 73.9%
3. จากรายการซื้อผลิตภัณฑ์หนึ่งหมื่นรายการ พบว่ามี 5 รายการ ที่ หากลูกค้าซื้อ Multi Plus Q10 , Carbo-X และ Fish Oil แล้ว มีโอกาสจะซื้อ L-Carnitine ด้วย ถึง 75.0%
4. จากรายการซื้อผลิตภัณฑ์หนึ่งพันรายการ พบว่ามี 1 รายการ ที่ หากลูกค้าซื้อ Carbo-X และ Royal Jelly แล้ว มีโอกาสจะซื้อ L-Carnitine ด้วยถึง 90.9%
5. จากรายการซื้อผลิตภัณฑ์หนึ่งหมื่นรายการ พบว่ามี 2 รายการ ที่ หากลูกค้าซื้อ Carbo-X และ Golden Oil แล้ว มีโอกาสจะซื้อ L-Carnitine ด้วยถึง 100.0%
6. จากรายการซื้อผลิตภัณฑ์หนึ่งหมื่นรายการ พบว่ามี 4 รายการ ที่ หากลูกค้าซื้อ L-Carnitine ,Royal Jelly และ Fish Oil แล้ว มีโอกาสจะซื้อ Carbo-X ด้วยถึง 100.0%
7. จากรายการซื้อผลิตภัณฑ์หนึ่งหมื่นรายการ พบว่ามี 4 รายการ ที่ หากลูกค้าซื้อ Carbo-X ,Royal Jelly และ Fish Oil แล้ว มีโอกาสจะซื้อ L-Carnitine ด้วยถึง 100.0%

นอกจากนี้พบว่าทุกความสัมพันธ์มีค่า lift สูงกว่า 1 ซึ่งสามารถตีความได้ว่าสินค้าในแต่ละชุดความสัมพันธ์นั้น ลูกค้ามักซื้อพร้อมกันมากกว่าซื้อแยกชนิดกัน จากกฎความสัมพันธ์นี้ สามารถนำไปวางแผนการจัดโปรโมชั่นทางการตลาดสำหรับทุกสาขาได้ และสามารถใช้อ้างอิงเหล่านี้เพื่อสนับสนุนการตัดสินใจในการวางแผนการตลาด

สำหรับการแบ่งกลุ่มลูกค้าผลการวิเคราะห์ข้อมูลพบว่า สามารถแบ่งกลุ่มลูกค้าสำหรับผลิตภัณฑ์เสริมอาหาร ยี่ห้อ MYHEALTH ได้ทั้งหมด 6 กลุ่ม ดังนี้

ชื่อกลุ่ม	Recency(วัน)	Frequency(ครั้ง)	Monetary(บาท)	จำนวนสมาชิก ในกลุ่ม(คน)
Blue Diamond	90.733	5.616	4436.868	2976
Diamond	302.399	2.662	1974.779	8054
Platinum	536.743	1.684	1150.520	4747
Gold	872.018	1.411	1020.514	2621
Silver	1326.154	1.248	1077.067	1832
Welcome	1676.667	1.090	778.943	1618

1. Blue Diamond คือ กลุ่มลูกค้าที่จะกลับมารับบริการซ้ำภายใน 91 วัน โดยมีค่าเฉลี่ยของจำนวนครั้งที่มาใช้ 5.7 ครั้ง และยอดใช้จ่ายเฉลี่ยประมาณ 4,437 บาท ซึ่งลูกค้าในกลุ่มนี้มีจำนวน 2,976 คน
2. Diamond คือ กลุ่มลูกค้าที่จะกลับมารับบริการซ้ำภายใน 302 วัน โดยมีค่าเฉลี่ยของจำนวนครั้งที่มาใช้ 2.7 ครั้ง และยอดใช้จ่ายเฉลี่ยประมาณ 1,975 บาท ซึ่งลูกค้าในกลุ่มนี้มีจำนวน 8,054 คน
3. Platinum คือ กลุ่มลูกค้าที่จะกลับมารับบริการซ้ำภายใน 505 วัน โดยมีค่าเฉลี่ยของจำนวนครั้งที่มาใช้ 1.7 ครั้ง และยอดใช้จ่ายเฉลี่ยประมาณ 1,151 บาท ซึ่งลูกค้าในกลุ่มนี้มีจำนวน 4,747 คน
4. Gold คือ กลุ่มลูกค้าที่จะกลับมารับบริการซ้ำภายใน 872 วัน โดยมีค่าเฉลี่ยของจำนวนครั้งที่มาใช้ 1.4 ครั้ง และยอดใช้จ่ายเฉลี่ยประมาณ 1,020 บาท ซึ่งลูกค้าในกลุ่มนี้มีจำนวน 2,621 คน
5. Silver คือ กลุ่มลูกค้าที่จะกลับมารับบริการซ้ำภายใน 1,326 วัน โดยมีค่าเฉลี่ยของจำนวนครั้งที่มาใช้ 1.2 ครั้ง และยอดใช้จ่ายเฉลี่ยประมาณ 1,077 บาท ซึ่งลูกค้าในกลุ่มนี้มีจำนวน 1,832 คน
6. Welcome คือ กลุ่มลูกค้าที่จะกลับมารับบริการซ้ำภายใน 1,677 วัน โดยมีค่าเฉลี่ยของจำนวนครั้งที่มาใช้ 1 ครั้ง และยอดใช้จ่ายเฉลี่ยประมาณ 779 บาท ซึ่งลูกค้าในกลุ่มนี้มีจำนวน 1,618 คน

จากการแบ่งกลุ่มลูกค้าตามมูลค่าของลูกค้า โดยพิจารณาจาก RFM model ข้างต้นทำให้บริษัทเข้าใจพฤติกรรม การมาซื้อสินค้าของลูกค้าแต่ละกลุ่มมากขึ้น และสามารถพิจารณาได้ว่าการออกโปรโมชั่นครั้งต่อไปจะเลือก กลุ่มเป้าหมายใดเป็นกลุ่มเป้าหมายแรก หรือเป็นกลุ่มเป้าหมายที่เป็นกลุ่มหัวหัด (beachhead segment) ที่บริษัท จะต้องรักษาและเพิ่มมูลค่าลูกค้ากลุ่มนี้ให้ได้ หรือการสร้างโปรโมชั่นหรือกิจกรรมเพื่อดึงดูดลูกค้าแต่ละกลุ่ม เช่นการเพิ่มความถี่ในการมาซื้อสินค้าของลูกค้ากลุ่ม Platinum เป็นต้น

นอกจากนี้งานวิจัยนี้ยังทำการหาความสัมพันธ์ของแต่ละกลุ่มลูกค้าเพื่อเพิ่มประสิทธิภาพในการออกโปรโมชั่น หรือกิจกรรมส่งเสริมการขายและดึงดูดลูกค้าแต่ละกลุ่ม ดังนี้

ชื่อกลุ่ม (จำนวนสมาชิก)	กฎความสัมพันธ์ที่พบ		Support (%)	Confidence (%)
	Premises	Conclusion		
Blue Diamond (2,976)	CALCIUM, CARBO-X	L-CARNITINE	0.1	76.9
Diamond (8,054)	RICE BRAN OIL, CARBO-X	L-CARNITINE	0.1	71.4
	L-CARNITINE, Royal Jelly	CARBO-X	0.1	78.6
	CALCIUM, CARBO-X	L-CARNITINE	0.1	80.0
	MULTI plus Q10, CARBO-X, Fish Oil	L-CARNITINE	0.1	83.3
	CARBO-X, Royal Jelly	L-CARNITINE	0.1	100.0
	CARBO-X, Golden Oil	L-CARNITINE	0.03	100.0
	Royal Jelly, Fish Oil, Bilberry	MULTI plus Q10	0.03	100.0
	L-CARNITINE, Royal Jelly, Fish Oil	CARBO-X	0.1	100.0
	CARBO-X, Royal Jelly, Fish Oil	L-CARNITINE	0.1	100.0
Platinum (4,747)	CALCIUM, CARBO-X	L-CARNITINE	0.1	71.4
	L-CARNITINE, Royal Jelly	CARBO-X	0.1	75.0
	CARBO-X, Royal Jelly	L-CARNITINE	0.1	75.0
	CALCIUM, L-CARNITINE	CARBO-X	0.1	83.3
	Rice Bran Oil, Golden Oil	Royal Jelly	0.05	100.0
	Royal Jelly, Golden oil	Rice Bran Oil	0.05	100.0
	Acerola Cherry, Fish Oil	Royal Jelly	0.02	100.0
	Fish Oil, Bilberry	Golden oil	0.02	100.0
	Bilberry, Golden oil	Fish Oil	0.02	100.0
	Calcium, L-CARNITINE, Royal Jelly	CARBO-X	0.02	100.0
	Calcium, CARBO-X, Royal Jelly	L-CARNITINE	0.02	100.0

ชื่อกลุ่ม (จำนวนสมาชิก)	กฎความสัมพันธ์ที่พบ		Support (%)	Confidence (%)
	Premises	Conclusion		
Gold (2,621)	L-CARNITINE, Acerola Cherry	CARBO-X	0.1	100.0
Silver (1,832)	L-CARNITINE, Acerola Cherry	CARBO-X	0.2	75.0
	Calcium, CARBO-X	L-CARNITINE	0.3	80.0
Welcome (1,618)	CARBO-X	L-CARNITINE	9	65.0
	Rice Bran Oil, CARBO-X	L-CARNITINE	0.5	100.0
	L-CARNITINE, Acerola Cherry	Calcium	0.1	100.0

จากตารางข้างต้นจะเห็นได้ว่าแต่ละกลุ่มมีกฎความสัมพันธ์ของการซื้อสินค้าที่แตกต่างกันไป และมีบางกฎที่เหมือนกัน เช่น เมื่อลูกค้าซื้อ ผลิตภัณฑ์ Calcium และ CARBO-X แล้วมีโอกาสจะซื้อ L-CARNITINE ด้วย หรือเมื่อลูกค้าซื้อ ผลิตภัณฑ์ Bilberry และ Golden oil แล้วมีโอกาสจะซื้อ FishOil ด้วยเช่นกัน ดังนั้น โปรโมชันที่เหมาะสมในการใช้เป็นอันดับแรกคือ Calcium, CARBO-X => L-CARNITINE เนื่องจากเป็น กฎความสัมพันธ์ที่เกิดขึ้นในทุกกลุ่ม ยกเว้น Welcome และมีค่า Confidence ที่ค่อนข้างสูงในแต่ละกลุ่ม

4.2 แนวทางการนำไปใช้

จากผลการวิจัยสามารถนำมาเป็นแนวทางในการลดต้นทุนของค่าใช้จ่ายในการนำเสนอหรือดึงดูดลูกค้ารายใหม่ และทำให้ทราบว่าลูกค้าที่ซื้อผลิตภัณฑ์หนึ่งแล้วมีแนวโน้มจะซื้อผลิตภัณฑ์ใดต่อบ้าง เพื่อที่การนำเสนอผลิตภัณฑ์ไปเสนอขายจะได้ตรงกับความต้องการของลูกค้ามากที่สุด และทำให้เกิดการซื้ออย่างต่อเนื่องต่อไป ดังนี้

1. จัดทำโปรโมชันสินค้า Calcium, CARBO-X และ L-CARNITINE ซึ่งมีแนวทางดังนี้
 - 1.1 จัด Set 3 ราคาพิเศษ
 - 1.2 ส่วนลดสำหรับการซื้อ L-Carnitine เมื่อลูกค้าซื้อ Calcium ร่วมกับ CARBO-X
 - 1.3 ส่ง SMS แจ้ง Promotion ผลิตภัณฑ์ L-Carnitine ให้กับลูกค้าที่เคยมีประวัติการซื้อ Calcium และ CARBO-X พร้อมแจ้งประโยชน์ของการทานร่วมกัน เนื่องจากเป็นช่องทางเดียวที่ติดต่อกับลูกค้าได้
2. จัดทำการเทรนพนักงานให้พนักงานเสนอขาย L-Carnitine เมื่อมีลูกค้ามาซื้อ Calcium และ CARBO-X
3. วางสินค้าทั้ง 3 ไว้บนชั้นสินค้าเดียวกัน
4. ลดจำนวนวันล่าสุดที่มาซื้อสินค้าโดยใช้ Loyalty program เช่น แลกคะแนน หรือการให้คูปองส่วนลดโดยกำหนดระยะเวลาการให้คูปอง

4.3 ปัญหาและอุปสรรค

ปัญหาและอุปสรรคที่ค้นพบจากการการวิจัยนี้ มีดังนี้

(1) เนื่องจากฐานข้อมูลที่ใช้ในการทำให้เหมือนข้อมูลนั้นไม่ได้มีการออกแบบเพื่อใช้ในการทำให้เหมือนข้อมูลโครงสร้างของฐานข้อมูลและเพิ่มข้อมูลอาจไม่เหมาะสมหรือครบถ้วนเพียงพอ ดังนั้นทำให้ต้องมีการวิเคราะห์ การแทนค่าข้อมูล และการเปลี่ยนรูปข้อมูล ดังนั้นถ้าต้องการทำให้เหมือนข้อมูลอย่างจริงจังควรมีการพัฒนากระบวนการเก็บข้อมูลเพื่อให้ข้อมูลที่เก็บมีความถูกต้องและเหมาะสมกับการทำให้เหมือนข้อมูลในอนาคต เพราะปัจจุบันทางองค์กรไม่เคยมีการนำข้อมูลมาทำให้เหมือนข้อมูลมาก่อน

(2) แต่ละขั้นตอนของกระบวนการทำให้เหมือนข้อมูล เช่นการเตรียมข้อมูล การวิเคราะห์ข้อมูลจะต้องอาศัยความร่วมมือจากผู้มีอำนาจในการตัดสินใจในการให้ข้อมูล ซึ่งข้อมูลบางอย่างไม่สามารถนำออกมาวิเคราะห์ให้ได้

(3) แบบจำลองที่ได้ยังไม่สามารถตอบสนองผู้ใช้งานได้ทันทีเนื่องจากการทำให้เหมือนข้อมูลแต่ละครั้งจะต้องทำการปรับปรุงระบบเพื่อปรับตัวแบบจำลองใหม่เพื่อให้รองรับต่อการเปลี่ยนแปลงของพฤติกรรมของลูกค้า

4.4 ข้อเสนอแนะสำหรับงานวิจัยต่อเนื่อง

ข้อเสนอแนะสำหรับงานวิจัยต่อเนื่อง มีดังนี้

(1) แนะนำผู้บริหารเก็บข้อมูลลูกค้าเพิ่มเติม เช่น เพศ, วัน เดือน ปีเกิด, ที่อยู่ หรือ อีเมล เพื่อวิเคราะห์ข้อมูลลูกค้าได้มากขึ้น และสามารถนำผลการวิเคราะห์ข้อมูลไปทำการตลาดได้อย่างแม่นยำมากขึ้น

(2) การศึกษาคำนี้ไม่มีข้อมูลเชิงลึกของกลุ่มลูกค้า ทำให้ไม่ได้มีการวิเคราะห์ถึง เพศ อายุ หรือที่อยู่ ซึ่งเป็นจุดที่น่าสนใจว่ากลุ่มคน เพศ อายุ หรือสถานที่อยู่ใด เหมาะกับการขายผลิตภัณฑ์ชนิดใด

(3) ควรมีการนำข้อมูลใหม่มาวิเคราะห์เพื่อหาความสัมพันธ์ใหม่ให้ทันสมัยอยู่ตลอด และเพื่อเปรียบเทียบกับความสัมพันธ์เดิม

(4) แนะนำให้ทำให้เหมือนข้อมูลเพื่อการพยากรณ์ยอดขายตามฤดูกาล เพื่อทำนายการขายสินค้า และการสั่งซื้อหรือสำรองสินค้า

บรรณานุกรม

พิณทิรา ต้นเถียร. (2552). อุปนิสัยการใช้ยาเองในคนไทย. *วารสารโลหิตวิทยาและเวชศาสตร์บริการโลหิต*, 19(4), 309-316.

นริศรัทธา ตูลาผลม. (2559). เจาะตลาดอาหารเสริมความงามกับความท้าทายในอนาคต. สืบค้นเมื่อ 6 มิถุนายน 2562, จาก <https://www.scbeic.com/th/detail/product/2858>.

ศูนย์วิจัยกสิกรไทย. (2558). ธุรกิจร้านขายยา (Pharmacy หรือ Drug Store). สืบค้นเมื่อ 6 มิถุนายน 2562, จาก <https://thaipublica.org/wpcontent/uploads/2016/06.pdf>.

ศูนย์วิจัยกสิกรไทย. (2560). เกาะกระแสสินค้าสุขภาพ สร้างโอกาสธุรกิจ. สืบค้นเมื่อ 6 มิถุนายน 2562, จาก <https://kasikornbank.com/th/business/sme/KSMEKnowledge/article/KSMEAnalysis/Documents/HealthyProduct.pdf>.

สำนักงาน. (2562). สถิติใบอนุญาตประกอบธุรกิจเกี่ยวกับยา ประจำปี พ.ศ. 2562. สืบค้นเมื่อ 6 มิถุนายน 2562, จาก <http://www.fda.moph.go.th/sites/drug/SitePages/Statistic.aspx>.

อำนาจ มณีศรีวงศ์กุล. (2541). การแบ่งกลุ่มจังหวัดโดยใช้การวิเคราะห์ตัวแปรเชิงพหุ. *วารสารวิจัย มข.*, 3(2), 46-53.

เอกสิทธิ์ พัชรวงศ์ศักดิ์. (2557). *การวิเคราะห์ข้อมูล ด้วยเทคนิคดาต้าไมนิ่งเบื้องต้น*. กรุงเทพฯ: เอเชีย ดิจิตอลการพิมพ์จำกัด.

Bandura, A. (1999). *Social Cognitive Theory Of Personalit*. New York: Guilford Publications.

- Bharati, R. (2010). Data Mining Techniques and Applications. *Indian Journal of Computer Science and Engineering*, 1(40), 301-305.
- Birant, D. (2011). *Knowledge-Oriented Applications in Data Mining*. Rijeka: IN-TECH.
- Chen, M., Han, J., & Yu, P.S. (1996). Data mining: an overview from a database perspective. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 866-883.
- Dangeti, P. (2017). *Statistics for Machine Learning*. Birmingham: Packt Publishing.
- Han, J. K. (2012). *Data Mining Concepts and Techniques*. Waltham: Morgan Kaufmann.
- Jensen, K. (2012). The phases of the CRISP-DM process. Retrieved June 6, 2019, from <https://sites.google.com/site/kennethagregaardjensen/crisp-dm>.
- Kotler, P. (1997). *Marketing Management: Analysis, Planning, Implementation, and Control. 9th Edition*. Upper Saddle River: Prentice Hall.
- Tampakis, P., Pelekis, N., Doukeridis, C., & Theodoridis, Y. (2019). Scalable Distributed Subtrajectory Clustering. *arXiv preprint arXiv:1906.06956*, 1-18.
- Tan, P., Steinbach, M., Karpatne, A. & Kumar, V. (2019). *Introduction to Data Mining. 2nd Edition*. Harlow: Pearson.