

บทที่ 2

แนวคิด ทฤษฎี และวรรณกรรมที่เกี่ยวข้อง

บทนี้เป็น การนำเสนอแนวคิด ทฤษฎี และวรรณกรรมที่เกี่ยวข้องกับการเปรียบเทียบโมเดล ในการวิเคราะห์ข้อมูลการยืมหนังสือในห้องสมุดของมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ โดยใช้เทคนิคใน Association Rule ตลอดจนการนำเสนอผลการวิเคราะห์ในรูปแบบที่สามารถเผยแพร่ผ่านทางเว็บไซต์ จากการศึกษาพบว่า มีเอกสารเชิงแนวคิด และโครงการที่เกี่ยวข้องกับการวิจัยในประเด็นดังกล่าว ซึ่งสามารถสรุปสาระสำคัญได้ดังนี้

2.1 แนวคิด

- 2.1.1 แนวคิดเกี่ยวกับเหมืองข้อมูล (Data mining)
- 2.1.2 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)
- 2.1.3 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytic)
- 2.1.4 แนวคิดเกี่ยวกับการเปรียบเทียบประสิทธิภาพโมเดล
- 2.1.5 แนวคิดเกี่ยวกับกฎความสัมพันธ์ (Association Rules)
- 2.1.6 แนวคิดเกี่ยวกับกฎความสัมพันธ์ ด้วย Apriori Algorithm
- 2.1.7 แนวคิดเกี่ยวกับกฎความสัมพันธ์ ด้วย FP-Growth

2.2 ทฤษฎี

- 2.2.1 ทฤษฎีเกี่ยวกับการจัดการข้อมูลขนาดใหญ่
- 2.2.2 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์
- 2.2.3 ทฤษฎีเกี่ยวกับชุดคำสั่ง HTML
- 2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS
- 2.2.5 ทฤษฎีเกี่ยวกับชุดคำสั่ง PHP
- 2.2.6 ทฤษฎีเกี่ยวกับกฎความสัมพันธ์ (Association Rules)
- 2.2.7 ทฤษฎีเกี่ยวกับการออกแบบ Dashboard

2.3 วรรณกรรมที่เกี่ยวข้อง

2.4 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับเหมืองข้อมูล (Data mining)

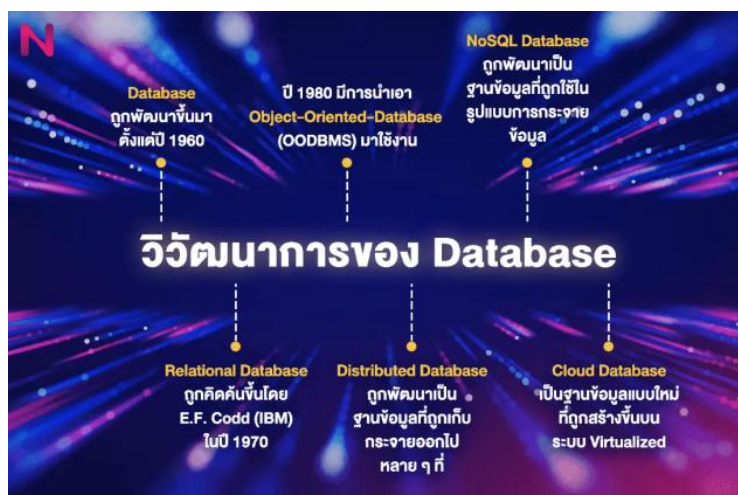
เทคโนโลยีการทำเหมืองข้อมูลมีต้นกำเนิดมาจากการประมวลผลไฟล์ในช่วงทศวรรษที่ 1960 ก่อนจะพัฒนามาเป็นการจัดเก็บข้อมูลในรูปแบบตารางและการใช้ภาษาคิวรีในการเรียกค้นข้อมูล เมื่อปริมาณข้อมูลเพิ่มขึ้นอย่างรวดเร็ว ทำให้เกิดความต้องการเครื่องมือในการวิเคราะห์ข้อมูลที่มีประสิทธิภาพสูง การทำเหมืองข้อมูลช่วยให้สามารถค้นพบรูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลขนาดใหญ่ ซึ่งนำไปสู่การตัดสินใจทางธุรกิจที่ชาญฉลาดยิ่งขึ้น

ในยุคดิจิทัลที่ข้อมูลมีความสำคัญอย่างยิ่ง การทำความเข้าใจพฤติกรรมของผู้ใช้เว็บไซต์เป็นสิ่งจำเป็นสำหรับองค์กรที่ต้องการสร้างความได้เปรียบในการแข่งขัน การวิเคราะห์ข้อมูลจากเว็บไซต์ช่วยให้สามารถปรับปรุงประสบการณ์ผู้ใช้ เพิ่มยอดขาย และพัฒนาผลิตภัณฑ์ใหม่ที่ตรงกับความต้องการของลูกค้าได้อย่างมีประสิทธิภาพ

การทำเหมืองข้อมูลหากผนวกเข้ากับระบบการจัดการความรู้ที่ได้จากฐานข้อมูล จะทำให้ได้ระบบการค้นหาข้อมูลและความรู้ หรือเรียกว่า (Knowledge and Data Discovery Management System : KDDMS) เป็นการทำเหมืองข้อมูลที่ใช้เทคนิคการกลั่นกรอง วิเคราะห์ ความรู้ที่ซ่อนอยู่ในฐานข้อมูลที่มีปริมาณมาก เพื่อได้ข้อมูลที่ต้องการ และมีประโยชน์สามารถนำข้อมูลที่ได้บริหารจัดการได้อย่างมีประสิทธิภาพ โดยผลลัพธ์มีอยู่หลายแบบ เช่น รูปแบบ (Pattern) ต้นแบบ (Model) กฎ (Rule) การทำนายหรือการคาดการณ์ล่วงหน้า (Prediction) ทั้งผลลัพธ์ที่ได้มาขึ้นอยู่กับการทำเหมืองข้อมูล (ภาณุสรณ์ ใจกลม และนายสมพงษ์ กาดาศาย, 2563)

2.1.1.1 วิวัฒนาการของการทำเหมืองข้อมูล

การพัฒนาเทคโนโลยีด้านฐานข้อมูล ได้มีวิวัฒนาการเรื่อยมาตั้งแต่อดีตจนถึงปัจจุบัน เพราะข้อมูลเป็นสิ่งสำคัญในการนำไปใช้ประโยชน์ในด้านต่าง ๆ ดังนั้นจึงได้มีการพัฒนา และปรับปรุงกระบวนการต่าง ๆ เพื่อรองรับการเพิ่มขึ้นของข้อมูลที่มีมากขึ้นเรื่อย ๆ จนกระทั่งสามารถเก็บรวบรวม และประมวลผลข้อมูลที่มีอยู่อย่างมหาศาลได้อย่างมีประสิทธิภาพ วิวัฒนาการของเทคโนโลยีด้านฐานข้อมูล



ภาพที่ 2.1 วิวัฒนาการเทคโนโลยีฐานข้อมูล

ที่มา : <https://nipa.cloud/th/blog/database-clouddatabase>

วิวัฒนาการเทคโนโลยีฐานข้อมูล เริ่มตั้งแต่ปี ค.ศ. 1960 โดยเทคโนโลยีฐานข้อมูลเริ่มจากการประมวลผลไฟล์ข้อมูลพื้นฐาน (File Processing) จากนั้นจึงมีการค้นคว้า และพัฒนาระบบฐานข้อมูลมาอย่างต่อเนื่อง จนกระทั่งเป็นระบบการเก็บข้อมูล การสร้างฐานข้อมูล (Database) และระบบจัดการฐานข้อมูล (Database Management System Language: DBMS)

ในปี ค.ศ. 1970 นักวิจัยและพัฒนาได้มีการเปลี่ยนแปลงการพัฒนาฐานข้อมูลแบบลำดับชั้น (Hierarchical Database) และฐานข้อมูลแบบเครือข่าย (Network Database) ไปสู่การพัฒนาการเก็บข้อมูลในรูปแบบตาราง (Relational Database System) โดยมีการสร้างเครื่องมือต่าง ๆ ที่ช่วยอำนวยความสะดวกในการจัดการกับข้อมูล และคิดค้นภาษาที่ใช้ในการเรียกข้อมูล (Query Language) เพื่ออำนวยความสะดวกในการเข้าถึงฐานข้อมูล

ในช่วงปี ค.ศ. 1980 เทคโนโลยีฐานข้อมูลได้เริ่มมีการปรับปรุงและพัฒนาในการหาระบบจัดการที่มีศักยภาพมากขึ้น ทำให้เกิดระบบการจัดการฐานข้อมูลที่มีประสิทธิภาพ เช่น ระบบการจัดการฐานข้อมูลเชิงวัตถุ (Objective-Oriented Database Management System) และระบบการจัดการฐานข้อมูลความสัมพันธ์ของวัตถุ (Object Relation Database Management System) จากนั้นเริ่มมีการพัฒนาความก้าวหน้าในเทคโนโลยีฮาร์ดแวร์ (Hardware) ซึ่งนำไปสู่การจัดเก็บข้อมูลจำนวนมากที่มีความซับซ้อนได้อย่างมีประสิทธิภาพเพิ่มขึ้น

ในปี ค.ศ. 1999-ปัจจุบัน สามารถจัดเก็บข้อมูลได้ในหลายรูปแบบแตกต่างกัน ทั้งระบบปฏิบัติการหรือการจัดเก็บฐานข้อมูล ซึ่งเป็นการนำข้อมูลทั้งหมดมารวมและจัดเก็บไว้

ในรูปแบบเดียวกัน เรียกว่า “คลังข้อมูล (Data Warehouse)” เพื่อความสะดวกในการจัดการต่อไป นอกจากนี้ยังมีการพัฒนาเทคนิคการวิเคราะห์ข้อมูลในหลาย ๆ มิติ เช่น เทคโนโลยีคลังข้อมูล การทำความสะอาดข้อมูล (Data Cleansing) การบูรณาการข้อมูล (Data Integration) และการประมวลผลเชิงวิเคราะห์ออนไลน์ (Online Analytical Processing: OLAP) รวมไปถึงการพัฒนาใช้ดาต้าไมนิ่ง (Data Mining) เพื่อนำข้อมูลจากฐานข้อมูลมาวิเคราะห์และประมวลผล โดยการสร้างแบบจำลอง และความสัมพันธ์ทางสถิติ ปัจจุบันการทำเหมืองข้อมูลเปรียบเสมือนอีกวิวัฒนาการหนึ่งในการจัดเก็บ และตีความหมายข้อมูลจากเดิมที่มีการจัดเก็บข้อมูลอย่างง่าย ๆ มาสู่การจัดเก็บในรูปฐานข้อมูลที่สามารถดึงค่าสารสนเทศของข้อมูลมาใช้ จนถึงการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล (พงศกร สุกันยา, 2565)

2.1.1.2 เทคโนโลยีที่เกี่ยวข้องกับเหมืองข้อมูล

การทำเหมืองข้อมูลเป็นการรวมเอาศาสตร์ความรู้ และเทคโนโลยีต่าง ๆ ที่หลากหลายเข้าไว้ด้วยกัน โดยมีรายละเอียดดังนี้

1) เทคโนโลยีฐานข้อมูล เนื่องจากฐานข้อมูลและคลังข้อมูลเป็นแหล่งที่เก็บรวบรวมข้อมูล เมื่อข้อมูลถูกสะสมไว้ในช่วงเวลาที่มีปริมาณมากพอให้สามารถนำมาค้นหาความรู้ได้ การทำเหมืองข้อมูลจึงมักมีความสัมพันธ์กับเทคโนโลยีฐานข้อมูล (Database Technology) ในลักษณะของกระบวนการค้นหาความรู้ที่ถูกเก็บไว้ในฐานข้อมูล หรือที่นิยมเรียกว่า (Knowledge Discovery in Databases: KDD)

2) สถิติศาสตร์ เทคนิคการวิเคราะห์ข้อมูลด้วยสถิติเป็นวิธีพื้นฐานที่นิยมใช้ในการวิเคราะห์ข้อมูลมานาน ทั้งในลักษณะของสถิติเชิงพรรณนา (Descriptive Statistics) เพื่อบรรยายถึงลักษณะของข้อมูล เช่น ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน หรือสถิติเชิงอนุมาน (Inference Statistics) ที่หาค่าความสัมพันธ์ระหว่างตัวแปร ได้แก่ การวิเคราะห์สหสัมพันธ์ การวิเคราะห์ความแปรปรวน เป็นต้น

3) การเรียนรู้ของเครื่อง (Machine Learning) การทำให้เครื่องคอมพิวเตอร์สามารถเรียนรู้การทำงานได้ด้วยตัวเอง โดยไม่ต้องพึ่งพาการสอนหรือทำตามคำสั่งของมนุษย์ คอมพิวเตอร์จะสามารถเรียนรู้เพื่อค้นหารูปแบบการแก้ไขปัญหา และการสร้างกฎจากข้อมูลโดยสามารถปรับปรุงการทำงานให้ดีขึ้นด้วยตัวเองจากการเรียนรู้ และประสบการณ์

การเรียนรู้ของเครื่องอาศัยอัลกอริทึมที่สามารถแบ่งออกเป็น 3 ประเภท ประกอบด้วย 1) การเรียนรู้แบบมีผู้สอน (Supervised Learning) 2) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) และ 3) การเรียนรู้แบบเสริมแรง (Reinforcement Learning)

4) การรู้จำรูปแบบ (Pattern Recognition) เป็นเทคนิคที่ใช้ในการจำแนกข้อมูล โดยอาศัยความรู้จดจำและเปรียบเทียบ ความเหมือน-ความต่างระหว่างลักษณะของรูปแบบที่จดจำกับรูปแบบอื่น

5) วิชาพลโลเซชัน (Visualization) การทำวิชาพลโลเซชัน เป็นการสร้างภาพ แผนผังหรือภาพเคลื่อนไหว ใช้ในการสื่อสารแทนข้อความ สำหรับข้อมูลที่มีปริมาณมหาศาลหากนำมาเสนอให้เห็นภาพในลักษณะที่สามารถวิเคราะห์หรือแปลความหมายได้สะดวก ก็อาจทำให้ค้นพบความสัมพันธ์ภายในข้อมูลที่ซ่อนอยู่ได้ดีขึ้น เช่น การนำชุดข้อมูลมาแสดงผลด้วยการพล็อตแบบกระจาย (Scatter Plot) เพื่อหาการเกาะกลุ่มกันของข้อมูล การแสดงส่วนร่วมกัน (Intersection) ของข้อมูล ทำให้สามารถแบ่งกลุ่มที่คล้ายคลึงกันภายในกลุ่ม หรือมีส่วนใดที่มีความคล้ายคลึงกับกลุ่มอื่นหรือมีกลุ่มใดบ้างที่แตกต่างกับกลุ่มอื่นอย่างนัยสำคัญ

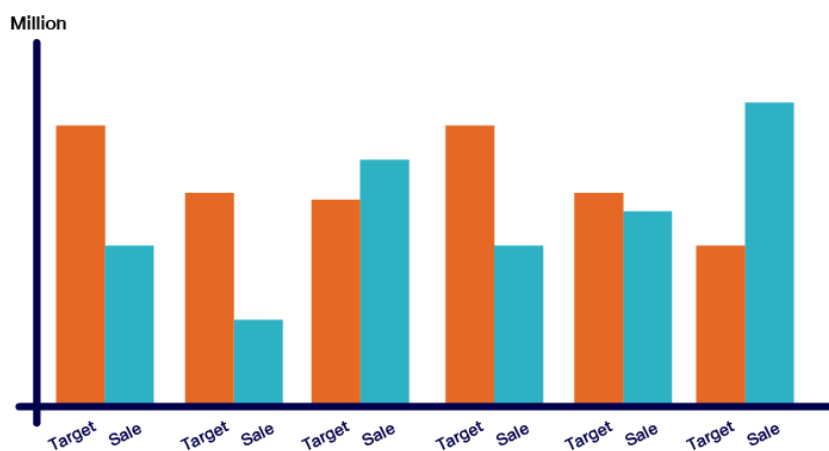
6) ศาสตร์ความรู้ด้านอื่น ๆ (Other Discipline) ได้แก่ การคำนวณประสิทธิภาพสูง (High-Performance Computing) เป็นต้นการทำเหมืองข้อมูลมักไม่ได้อาศัยเพียงเทคนิคใดเทคนิคหนึ่ง แต่มักเป็นการรวมเทคนิคต่าง ๆ ที่กล่าวนำเข้าไว้ด้วยกัน เช่น การนำข้อมูลจากฐานข้อมูลมาทำเหมืองข้อมูลโดยใช้การวิเคราะห์เชิงสถิติ แล้วนำเสนอผลโดยการสร้างภาพนามธรรม เป็นต้น (ปราโมทย์ ลือนาม, 2561)

2.1.2 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

Data Visualization หรือ Information Visualization คือ การถ่ายทอดข้อมูลในเชิงปริมาณที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ ซึ่งคำว่า “ประสิทธิภาพ” ในที่นี้หมายถึงมีความชัดเจน (Clarity), มีความแม่นยำ (Precision), และมีประสิทธิภาพ (Efficiency) หากไม่มีการทำ Data Visualization แล้ว อาจทำให้เราไม่สามารถค้นพบนัยยะของข้อมูลในแง่ของแนวโน้ม, รูปแบบพฤติกรรม, และความสัมพันธ์เชื่อมโยงได้

การเลือกรูปแบบ Visualization ให้เหมาะสมกับข้อมูล ในปัจจุบันเป็นยุคเทคโนโลยีเข้าถึงทุกคน ทำให้การรับรู้ข่าวสารข้อมูลต่าง ๆ เป็นไปได้ง่าย และรวดเร็วมากขึ้น คนที่นำเสนอข้อมูลจึงต้องนำเสนอข้อมูลให้น่าสนใจ เข้าใจง่าย และรวดเร็ว เพื่อให้การนำเสนอข้อมูลน่าสนใจ และมีประสิทธิภาพสูงสุด ซึ่งเลือกแผนภูมิชนิดนิยมที่ใช้กันอย่างแพร่หลาย ดังนี้

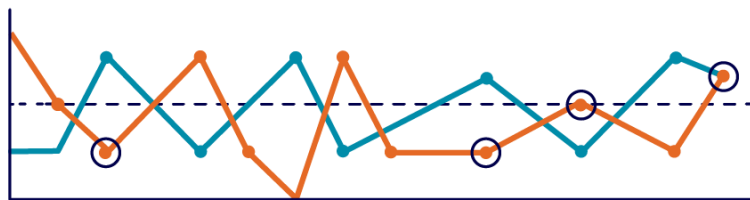
1) แผนภูมิแท่ง (Bar Charts) เป็นแผนภูมิที่ประกอบด้วยแกนนอน แกนตั้งที่นิยมแสดงออกมาในรูปแบบสี่เหลี่ยมที่สามารถบอกความสูงได้ เหมาะสำหรับการเปรียบเทียบ จำนวนของข้อมูลในแต่ละชุดเช่น รายรับในแต่ละเดือน, ยอดขายที่ขายได้จริงเปรียบเทียบกับ เป้าหมายยอดขายที่ตั้งไว้ เป็นต้นซึ่งแผนภูมิแท่งยังสามารถแบ่งออกเป็นหลายประเภท แผนภูมิแท่งแบบจัดกลุ่ม แผนภูมิแท่งแบบจัดกลุ่ม เหมาะสมกับการนำเสนอข้อมูลที่มีข้อมูล ย่อย ๆ อยู่ภายใต้ข้อมูลใหญ่เป็นการเน้นให้เห็นข้อมูลย่อยนั้น ๆ แผนภูมิแท่งแบบวางซ้อนกัน แผนภูมิแท่งแบบวางซ้อนกัน เหมาะสมกับการนำเสนอข้อมูลที่ต้องการให้เห็นข้อมูลย่อยในแต่ละ ข้อมูลใหญ่ และยังแสดงให้เห็นสัดส่วนของข้อมูลย่อยต่าง ๆ เหล่านั้นได้ด้วย ใช้แผนภูมินี้เมื่อ มีชุดข้อมูล หลายชุด และต้องการเน้นผลรวมทั้งหมด



ภาพที่ 2.2 แผนภูมิแท่ง (Bar Charts)

ที่มา : <https://bit.ly/3aXjpLg>

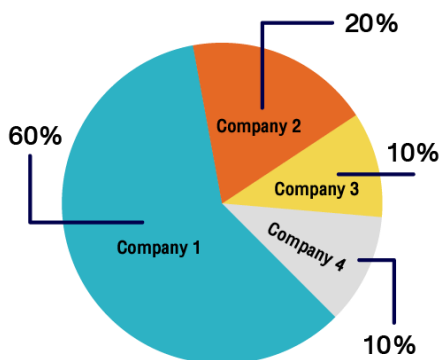
2) แผนภูมิเส้น (Line Charts) แผนภูมิเส้น มีลักษณะคล้ายแผนภูมิแท่ง ซึ่งประกอบด้วยแกนตั้งและนอน เพียงแต่เปลี่ยนจากแท่งข้อมูลเป็นจุดบนแผนภูมินั้นเอง แผนภูมิประเภทนี้เหมาะกับการนำเสนอข้อมูลตัวเลขที่มีความสัมพันธ์กับข้อมูลที่มีลักษณะเป็นช่วงใช้แสดงให้เห็นการเปลี่ยนแปลงที่เกิดขึ้นตามเวลาเพื่อดูแนวโน้ม รวมถึงสามารถใช้ พยากรณ์แนวโน้มในอนาคตได้เช่น ข้อมูลของยอดขายในแต่ละปี หรือไตรมาส และนำมาวิเคราะห์เพื่อดูแนวโน้ม เป็นต้น



ภาพที่ 2.3 แผนภูมิเส้น (Line Charts)

ที่มา : <https://bit.ly/3aXjpLg>

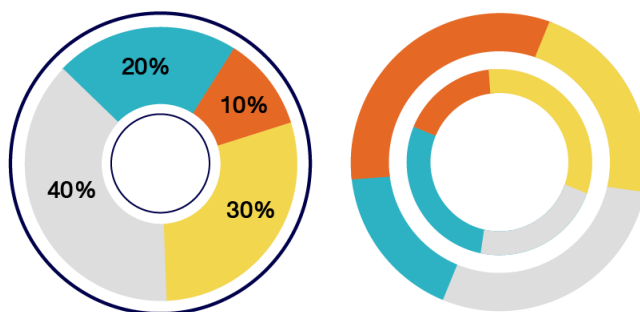
3) แผนภูมिवงกลม (Pie Charts) แผนภูมिवงกลมเหมาะกับการนำเสนอข้อมูลที่มีส่วนประกอบย่อยที่รวมกันเป็นส่วนใหญ่ มีการแบ่งส่วนให้ดูง่าย และสวยงามแต่ในทางกลับกันอาจจะดูยากในเรื่องของการประมาณขนาดของแต่ละชิ้น ยิ่งถ้ามีจำนวนชิ้นมาก จะยิ่งแยกยาก เพราะต้องใช้หลายสี ในการนำเสนอข้อมูล เช่น ส่วนแบ่งทางการตลาด (Market Share), ข้อมูลแสดงส่วนผสมต่าง ๆ เป็นต้น



ภาพที่ 2.4 แผนภูมिवงกลม (Pie Charts)

ที่มา : <https://bit.ly/3aXjpLg>

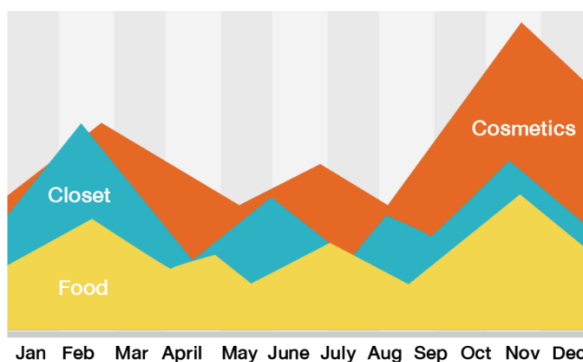
4) แผนภูมิโดนัท (Doughnut Charts) แผนภูมิโดนัทมีหลักการออกแบบเช่นเดียวกับแผนภูมिवงกลมแต่สามารถแสดงชุดข้อมูลได้มากกว่า one ชุด โดยนำเสนอข้อมูลเป็นวงกลมซ้อนกันหลาย ๆ ชั้น



ภาพที่ 2.5 แผนภูมิโดนัท (Doughnut Charts)

ที่มา : <https://bit.ly/3aXjpLg>

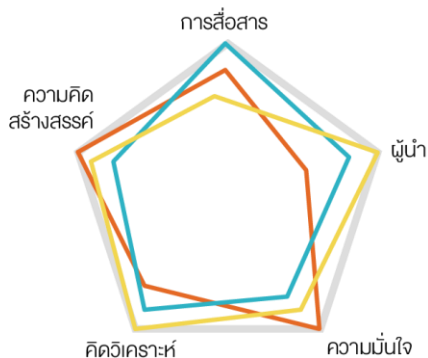
5) แผนภูมิพื้นที่ (Area Charts) มีหน้าตาคล้ายแผนภูมิเส้น แต่มีการแรเงาพื้นที่ใต้เส้นข้อมูล หรือระหว่าง two เส้นเพื่อแสดงให้เห็นปริมาณความแตกต่างระหว่างเส้น เหมาะ สำหรับเน้นความสำคัญของการเปลี่ยนแปลงในช่วงเวลาแสดงให้เห็นผลรวมของความแตกต่าง ระหว่างข้อมูล



ภาพที่ 2.6 แผนภูมิพื้นที่ (Area Charts)

ที่มา : <https://bit.ly/3aXjpLg>

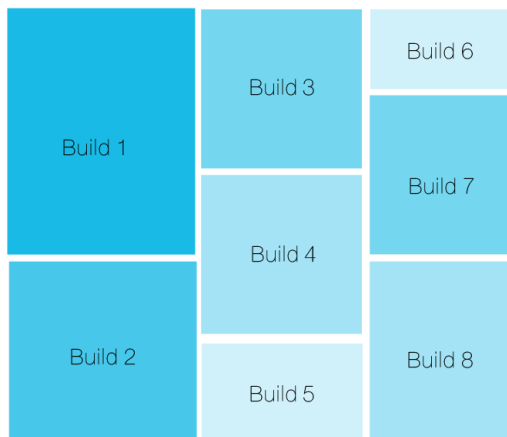
6) แผนภูมิเรดาร์ (Radar Charts) มีลักษณะคล้ายแผนภูมิเส้นที่มีการแสดงผลแบบวงกลมจำนวนเหลี่ยมของเรดาร์เท่ากับจำนวนหัวข้อของข้อมูล แผนภูมินี้ไม่ได้บอกถึงความต่อเนื่องของข้อมูลแต่เหมาะสำหรับการนำเสนอข้อมูลเป็นหัวข้อ แล้วนำมาวิเคราะห์หาจุดอ่อน จุดแข็งของข้อมูล



ภาพที่ 2.7 แผนภูมิเรดาร์ (Radar Charts)

ที่มา : <https://bit.ly/3aXjpLg>

7) แผนภูมิต้นไม้ (Tree Maps) คือการนำเสนอข้อมูลแบบแสดงให้เห็นพื้นที่ แสดงผลได้ในแบบลำดับชั้น เหมือนแบบโครงสร้างต้นไม้ อาจจะนำเสนอข้อมูลที่ต้องการให้เห็น ถึงเขตพื้นที่แสดงพื้นที่สีที่แตกต่างกันได้

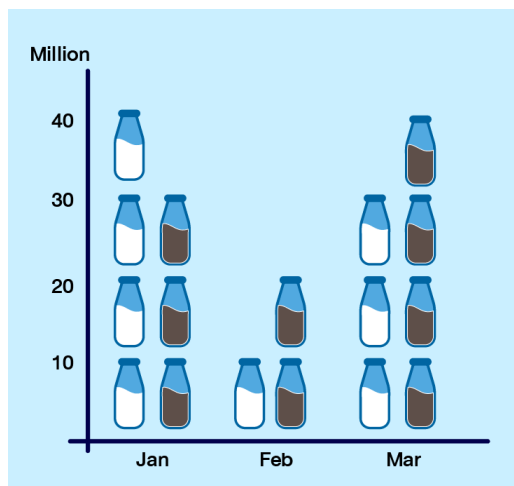


ภาพที่ 2.8 แผนภูมิต้นไม้ (Tree Maps)

ที่มา : <https://bit.ly/3aXjpLg>

8) แผนภูมิรูปภาพ (Picture Graph) เป็นแผนภูมิที่ประกอบไปด้วยแกนนอน และ แกนตั้ง แต่เลือกใช้รูปภาพ หรือไอคอนแทนจำนวนของสิ่งของนั้น ๆ เช่น การแสดงผลจำนวนของนมที่ขายได้ในแต่ละเดือน โดยนำเสนอทั้งนมรสจืด รสช็อกโกแลต เปรียบเทียบใน แต่ละเดือนซึ่งมีการนำเสนอลักษณะคล้ายกับกราฟแท่ง แต่เปลี่ยนจากแท่งเป็นรูปภาพของ นม 2 รสชาติแทน ก็ทำให้การนำเสนอข้อมูลน่าสนใจมากขึ้นซึ่งแนวทางการนำเสนอข้อมูล ลักษณะนี้ต้องอาศัยความ

คุ่นชินของคนดู เพื่อแทนสัญลักษณ์ภาพลงไป เช่น เมื่อพูดถึงจำนวนคน อาจแทนด้วยภาพไอคอนคนหรือเมื่อพูดถึงจำนวนเงิน ควรแทนภาพเป็นเหรียญเงิน หรือแบงค์แทน ก็จะทำให้คนดูเข้าใจง่ายจากสัญลักษณ์ภาพที่คุ้นเคยอยู่แล้ว และยังดึงดูดความสนใจได้มากกว่าการใช้กราฟแท่งสี่เหลี่ยม (จิราพรพรรณ กันทะมูล และพัชรินทร์ ชัยพรหม, 2563)



ภาพที่ 2.9 แผนภูมิรูปภาพ (Picture Graph)

ที่มา : <https://bit.ly/3aXjpLg>

2.1.3 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytic)

เป็นศาสตร์ในการวิเคราะห์และประมวลผลข้อมูลด้วยวิธีทางคณิตศาสตร์ สถิติ และการวิเคราะห์ข้อมูลที่มีอยู่ในปัจจุบัน อดีต และทำนายอนาคต เพื่อนำผลนั้นมาปรับใช้ให้เกิดประโยชน์ ความรู้เรื่องดังกล่าวสำคัญยิ่งขึ้น เมื่อเทคโนโลยีในปัจจุบันมีการเก็บข้อมูลมหาศาลในระดับ Big Data มากเกินกว่าคนทั่วไปจะสามารถใช้งานได้เต็มที่ Data Analytics จึงเหมาะที่จะเป็นเครื่องมือสำหรับธุรกิจ (Business Intelligence) ทั้งขนาดใหญ่และขนาดเล็ก

2.1.3.1 ความหมายการวิเคราะห์ข้อมูล (Data analytic)

Data Analytics นั้นเป็น Business Intelligence อย่างหนึ่ง ซึ่งเป็นศาสตร์ของการใช้ข้อมูลต่าง ๆ จากที่ต่าง ๆ มาร่วมวิเคราะห์รวมกัน เพื่อปรับปรุงธุรกิจ หรือการตลาด ซึ่ง Data Analytics นั้นเป็น Business Analytics อย่างหนึ่ง โดย Business Analytics เป็นวิธีการที่ใช้ข้อมูลเทคโนโลยีด้านสารสนเทศ เข้ามาทำการวิเคราะห์ข้อมูล เพื่อช่วยในการดำเนินธุรกิจ ซึ่ง Business Analytics แบ่งเป็น three แบบคือ

1) Descriptive Analytics เป็นรูปแบบการใช้ข้อมูลแบบพื้นฐานที่สุด โดยเน้นการ อธิบายว่ากำลังเกิดขึ้น หรืออาจจะเกิดอะไรขึ้น สามารถอธิบายถึงสาเหตุการเกิดต่าง ๆ ได้ว่าทำไม ซึ่ง Descriptive Analytics ตัวอย่างคือรายงานธุรกิจ รายงานด้านการทำ Campaign หรือ โฆษณา หรือรายงานผลดำเนินงานที่ผ่านมา เป็นข้อมูลพื้นฐานที่แสดงผลในการทำกิจกรรมต่าง ๆ ให้เราได้รับทราบ

2) Predictive Analytics เป็นรูปแบบการใช้ข้อมูลที่มีความซับซ้อนขึ้นมา โดยจะ เป็นการ “พยากรณ์” หรือ “ทำนาย” สิ่งที่กำลังเกิดขึ้น โดยใช้ข้อมูลในอดีต ร่วมกับโมเดล ทาง คณิตศาสตร์ต่าง ๆ หรือร่วมกับการทำ Data Mining นอกจากนี้ Predictive Analytics ยังทำให้ เรา สามารถวิเคราะห์หาโอกาสและความเสี่ยงต่าง ๆ ที่จะเกิดขึ้นในอนาคตได้

3) Prescriptive Analytics เป็นรูปแบบการวิเคราะห์ข้อมูล ที่มีความซับซ้อน และ ยากที่สุด เพราะไม่เพียงพยากรณ์หรือทำนายว่าอะไรจะเกิดขึ้น แต่ยังให้คำแนะนำในทางเลือก ต่าง ๆ และผลแต่ละทางเลือกว่าจะมี Pros & Cons อย่างไร โมเดลของ Prescriptive Analytics นั้น จะสามารถปรับเปลี่ยนได้ตามข้อมูลที่เพิ่มเติมเข้ามามากขึ้น และ Prescriptive Analytics นี้ยัง เป็น การใช้ข้อมูลที่มากที่สุด และเกี่ยวพันกับเรื่อง Big Data เป็นอย่างมาก (กัญญา อับปัมญญา และ สุชาดา ดาจันทร์, 2563)

2.1.4 แนวคิดเกี่ยวกับการเปรียบเทียบประสิทธิภาพอัลกอริทึม

เทคนิคการสร้างกฎความสัมพันธ์มุ่งค้นหากฎและรูปแบบไอเทมเซตที่เกิดบ่อยใน ฐานข้อมูล มีหลายอัลกอริทึมที่นิยมใช้ เช่น อัลกอริทึม Apriori และอัลกอริทึม FP-Growth

2.1.4.1 อัลกอริทึม Apriori ใช้วิธีวนตรวจสอบฐานข้อมูลหลายรอบ โดยสร้างชุด ตัวแทนที่ละระดับเพื่อค้นหาไอเทมเซตที่เกิดบ่อยและเปรียบเทียบกับค่าสนับสนุนขั้นต่ำ ชุดไอเทม เซตที่มีค่า Support เท่ากับหรือมากกว่าจะถือว่าเป็นเกิดบ่อย ส่วนที่ต่ำกว่าจะถูกตัดทิ้ง จากนั้นจะสร้าง ชุดตัวแทนใหม่จากไอเทมเซตที่เกิดบ่อย ข้อดีคือเข้าใจง่าย แต่ทำงานช้ากับฐานข้อมูลขนาดใหญ่ เพราะต้องตรวจสอบหลายรอบ

2.1.4.2 อัลกอริทึม FP-Growth ใช้หลักการสร้างต้นไม้ (FP-tree) และการแบ่งแยก และพิชิต (Divide-and-conquer) เพื่อประมวลผลข้อมูลย่อยอย่างรวดเร็ว ในขั้นแรกจะสร้างราก ต้นไม้เป็นค่าว่าง (null) แล้วจัดเรียงข้อมูลโดยแทนค่าตัวแปรด้วยโหนดซึ่งเทคนิคนี้ไม่สร้างชุดข้อมูล Candidate itemsets ทำให้ได้กฎความสัมพันธ์เร็ว แต่ข้อมูลต้องอยู่ในรูปแบบไบนารี (กรีชบดินทร์ พิวหอม, 2567)

2.1.5 ทฤษฎีเกี่ยวกับกฎความสัมพันธ์ (Association Rules)

กฎความสัมพันธ์ (Association Rules) เป็นวิธีทำเหมืองข้อมูลวิธีหนึ่งที่ได้รับค่านิยม เพื่อใช้ในการค้นหาความสัมพันธ์ของข้อมูลที่สนใจ โดยจะนำข้อมูลความสัมพันธ์ที่ได้ ไปประกอบการตัดสินใจในเรื่องต่างๆ เช่น การวิเคราะห์ข้อมูลการซื้อสินค้าหรือบริการของลูกค้าว่ามี การซื้อรูปแบบใด เช่น ถ้าซื้อ สินค้าอันนี้แล้วจะซื้อสินค้าใดควบคู่กันไปด้วย ซึ่งข้อมูลที่ได้จะทำให้ สามารถดำเนินการจัดวางสินค้า หรือบริการให้อยู่บริเวณใกล้เคียงกันเพื่อความสะดวกของลูกค้า ทำให้ลูกค้าเกิดความสะดวกสบายและอยากที่จะมาซื้อสินค้าหรือบริการของที่ร้าน และทางผู้ขาย ก็สามารถวางแผนการตลาดให้เหมาะสมได้ยิ่งขึ้น เทคนิคในการค้นหาความสัมพันธ์จะอยู่ใน รูปแบบของเงื่อนไขและผลลัพธ์ที่เกิดขึ้น ดังนี้

$$A (\text{เงื่อนไข}) \rightarrow B (\text{ผลลัพธ์})$$

ซึ่งเงื่อนไขและผลลัพธ์ที่มีความสัมพันธ์กันจะถูกลบทิ้งก็ต่อเมื่อมีค่าสนับสนุน (Support) และมี ค่าความเชื่อมั่น (Confidence) มากกว่าค่าขั้นต่ำที่กำหนดไว้ โดยค่าสนับสนุน และค่าความเชื่อมั่น สามารถคำนวณได้ดังนี้

ค่าสนับสนุน (Support) หาได้จากนำจำนวนข้อมูลที่มีเงื่อนไขและผลลัพธ์ตรงตามกฎ ($A \rightarrow B$)หารด้วย จำนวนข้อมูลรายการทั้งหมด (R) ซึ่งจะอยู่ในรูปสมการดังนี้

$$\text{Support } (A \rightarrow B) = P(A \cup B) = (A \rightarrow B) / R$$

ค่าความเชื่อมั่น (Confidence) หาได้จากการนำจำนวนข้อมูลที่มีเงื่อนไขและผลลัพธ์ตรงตามกฎ ($A \rightarrow B$) หารด้วยจำนวนข้อมูลที่เป็นเงื่อนไข (A) ซึ่งจะอยู่ในรูปสมการดังนี้

$$\text{Confidence } (A \rightarrow B) = P(A|B) = (A \rightarrow B) / A$$

ตัวอย่างการหาค่าสนับสนุนและค่าความเชื่อมั่นจากข้อมูลการซื้อกับข้าวของลูกค้า ทั้งหมด 4 คน โดยมีข้อมูลดังตารางที่ 1 และนำข้อมูลมาหาความถี่ในการซื้อกับข้าวร่วมกัน ดัง ตารางที่ 2

ตารางที่ 2.1 ตารางแสดงข้อมูลการขายกับข้าว

ลูกค้า	กับข้าว
1	ปลาร้าทอด, แกงซี่เหล็ก, ไก่ทอด, ผัดปลาตาก
2	กะเพราหมู, ไก่ทอด
3	ปลาร้าทอด, กะเพราหมู, ผัดปลาตาก
4	ปลาร้าทอด, ผัดปลาตาก, แกงซี่เหล็ก

ตารางที่ 2.2 ตารางแสดงข้อมูลการเกิดเหตุการณ์ร่วมของการซื้อกับข้าว

	ปลาร้าทอด	แกงซี่เหล็ก ผัด	ไก่ทอด	กะเพราหมู	ปลาตาก
ปลาร้าทอด	3	2	1	1	3
แกงซี่เหล็ก	2	2	1	0	2
ไก่ทอด	1	1	2	1	1
กะเพราหมู	1	0	1	2	1
ผัดปลาตาก	3	2	1	1	3

สมมุติให้กฎที่ 1 คือ ปลาร้าทอด → แกงซี่เหล็ก

$$\text{ค่าสนับสนุน} = 2/4 = 0.4 = 50\%$$

เลข 2 มาจากเหตุการณ์ทั้งหมดที่ลูกค้าซื้อปลาร้าทอดและแกงซี่เหล็กด้วย

เลข 4 มาจากจำนวนรายการซื้อทั้งหมด

ค่าความเชื่อมั่น = $2/3 = 0.67 = 67\%$ เลข 2 มาจากเหตุการณ์ทั้งหมดที่ลูกค้าซื้อปลาร้าทอดแล้วจะซื้อแกงซี่เหล็กด้วย เลข 3 มาจากจำนวนลูกค้าที่ซื้อปลาร้าทอดทั้งหมด

สมมุติให้กฎที่ 2 คือ แกงซี่เหล็ก → ปลาร้าทอด

$$\text{ค่าสนับสนุน} = 2/4 = 0.5 = 50\%$$

เลข 2 มาจากเหตุการณ์ทั้งหมดที่ลูกค้าซื้อแกงซี่เหล็กและปลาร้าทอดด้วย

เลข 4 มาจากจำนวนรายการซื้อทั้งหมด

$$\text{ค่าความเชื่อมั่น} = 2/2 = 1 = 100\%$$

เลข 2 มาจากเหตุการณ์ทั้งหมดที่ลูกค้าซื้อแกงซี่เหล็กและปลาร้าทอดด้วย

เลข 2 มาจากจำนวนลูกค้าที่ซื้อแกงซี่เหล็กทั้งหมด (ชลิยา ใจเย็น อะอิตะ, สิริกุล กวมทรัพย์, ชลาลัย โชคดีศรีจันทร์ และสุนิสา ริมเจริญ, 2563)

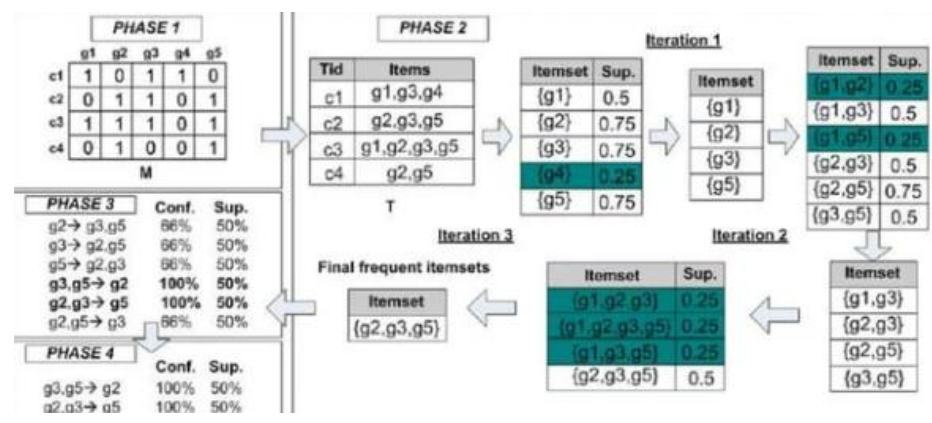
2.1.6 แนวคิดเกี่ยวกับกฎความสัมพันธ์ด้วย Apriori Algorithm

ความสัมพันธ์ (Association) โดยหลักการทำงาน คือการหาความสัมพันธ์ของข้อมูลภายในกลุ่มข้อมูล เพื่อใช้ลักษณะของข้อมูลหนึ่งในการบอกถึงลักษณะที่จะเกิดขึ้นกับอีกตัวหนึ่ง ซึ่งอาจจะเป็นการหาความสัมพันธ์ของข้อมูลในกลุ่มเดียวกัน ตัวอย่างเช่น การระบุว่าในกลุ่มของนักเรียนที่สมัครเรียนสาขาวิชาพยาบาลศาสตร์แล้วจะเป็นเพศหญิงนั้น มีโอกาสเกิดขึ้นร่วมกัน ผลลัพธ์ที่ได้จากการค้นหาว่าสัมพันธ์นั้นสามารถนำไปใช้ในการวิเคราะห์สนับสนุนหรือพัฒนางานที่เกี่ยวข้องกับฐานข้อมูลนั้น ได้อย่างมีประสิทธิภาพมากยิ่งขึ้นซึ่งวิธีที่ได้รับความนิยม และเป็นที่ยอมรับ คือการค้นหาในกลุ่มข้อมูลที่ปรากฏข้อมูลร่วมกันบ่อย

การค้นหากฎความสัมพันธ์ของข้อมูลในฐานข้อมูลได้พัฒนาขึ้นครั้งแรกโดยนักวิจัยจากศูนย์วิจัย IBM (International Business Machines Corporation) ประเทศสหรัฐอเมริกา จุดประสงค์เพื่อค้นหาความสัมพันธ์ที่น่าสนใจ ซึ่งซ่อนอยู่ในข้อมูลการซื้อสินค้าของลูกค้าว่าจะซื้อสินค้าใดบ้างร่วมกันในตะกร้ารถเข็นในห้างสรรพสินค้า (Market Basket Analysis) เพื่อให้ทราบถึงพฤติกรรมการซื้อสินค้าของลูกค้า เช่น เมื่อลูกค้าซื้อนมแล้วจะซื้อขนมปังด้วยการค้นหา กฎความสัมพันธ์มีขั้นตอนที่สำคัญที่สุดคือการค้นหาในกลุ่มข้อมูลที่ปรากฏร่วมกันบ่อย ซึ่งจะมีขั้นตอนกระบวนการการทำงานนานที่สุด ดังนั้นในการเลือกขั้นตอนวิธีในการค้นหาความสัมพันธ์จึงเป็นสิ่งสำคัญเพื่อให้เหมาะสมกับลักษณะของข้อมูล เพื่อให้ลดระยะเวลา และเมื่อที่หน่วยความจำในการประมวลผลการทำงาน กระบวนการค้นหาความสัมพันธ์มีขั้นตอนการค้นหาข้อมูลที่ปรากฏร่วมกันบ่อยทั้งหมด กลุ่มข้อมูลเหล่านั้นจะต้องมีค่าสนับสนุนมากกว่าหรือเท่ากับค่าสนับสนุนขั้นต่ำที่ผู้ใช้กำหนดจึงจะถือว่าเป็นกลุ่มข้อมูลที่ปรากฏร่วมกันบ่อย และขั้นตอนการนำกลุ่มข้อมูลที่ปรากฏบ่อยมาสร้างกฎความสัมพันธ์ ซึ่งกฎความสัมพันธ์จะเป็นที่ยอมรับได้หากตรวจสอบค่าความเชื่อมั่นของกฎนั้นมีค่ามากกว่า หรือเท่ากับค่าความเชื่อมั่นขั้นต่ำที่ผู้ใช้กำหนด

ขั้นตอนวิธี Apriori (Apriori Algorithm) เป็นขั้นตอนวิธีที่ได้รับการยอมรับและเป็นที่ยอมรับในการค้นหาความสัมพันธ์โดยเซตที่มีความถี่มากกว่าค่าสนับสนุนขั้นต่ำ เซตย่อยจะมีความถี่ที่มากกว่า และสามารถตัดเซตรายการที่มีความถี่ต่ำออก ขั้นตอนวิธีในการค้นหาเซตรายการที่ปรากฏร่วมกันบ่อย และนำมาสร้างกฎความสัมพันธ์ข้อดีของขั้นตอนวิธี Apriori คือสามารถทำงานได้ดีหากกำหนดค่าสนับสนุนขั้นต่ำมีค่ามาก มีขนาดของฐานข้อมูลขนาดเล็ก และมี

จำนวนของเซตรายการน้อย ส่วนข้อเสียคือต้องอ่านข้อมูลจากฐานข้อมูลหลายครั้ง (กิตติชัย ถวิลไทย และภวนาถ จิตวีเชื้อ, 2564)



ภาพที่ 2.10 ขั้นตอนวิธี Apriori

ที่มา : <https://n9.cl/6m92th>

2.1.7 แนวคิดเกี่ยวกับกฎความสัมพันธ์ ด้วย FP-Growth

FP Growth (Frequent Pattern Growth) เป็นอัลกอริทึมที่ใช้ในการค้นหาแพทเทิร์นที่เกิดขึ้นบ่อยในชุดข้อมูลขนาดใหญ่ โดยอัลกอริทึมนี้ช่วยในการค้นหาและวิเคราะห์ข้อมูลที่มีลักษณะซ้ำซาก เช่น การซื้อสินค้าที่เกิดขึ้นในตลาด เพื่อตัดสินใจทางธุรกิจอย่างมีประสิทธิภาพ FP Growth ใช้หลักการของโครงสร้างต้นไม้ที่เรียกว่า FP-tree ซึ่งช่วยลดจำนวนการคำนวณและทำให้การค้นหาข้อมูลที่มีบ่อยเร็วขึ้น นอกจากนี้ยังช่วยลดปริมาณข้อมูลที่ต้องจัดการในกระบวนการทำเหมืองข้อมูล (Data Mining) ด้วยการทำงานที่มีประสิทธิภาพและรวดเร็วกว่าอัลกอริทึมอื่น ๆ เช่น Apriori

2.1.7.1 หลักการทำงานของ FP-Growth

FP-Growth (Frequent Pattern Growth) เป็นอัลกอริทึมที่ใช้ในการค้นหาลวดลายที่เกิดขึ้นบ่อยในชุดข้อมูลขนาดใหญ่ โดยอัลกอริทึมนี้ถูกพัฒนาเพื่อให้สามารถทำงานได้อย่างมีประสิทธิภาพมากกว่าอัลกอริทึมที่ใช้วิธีการสร้างชุดข้อมูลรวม (Apriori) ซึ่งมีข้อเสียเรื่องการต้องการการคำนวณและการจัดเก็บข้อมูลจำนวนมาก หลักการทำงานของ FP-Growth สามารถแบ่งออกเป็นขั้นตอนหลักดังนี้: การสร้าง FP-Tree ขั้นตอนแรกคือการสร้างต้นไม้ FP (Frequent Pattern Tree) ซึ่งเป็นโครงสร้างข้อมูลที่ใช้ในการเก็บข้อมูลเกี่ยวกับลวดลายที่เกิดขึ้นบ่อย ในขั้นตอนนี้ อัลกอริทึมจะทำการสแกนชุดข้อมูลเพื่อคำนวณความถี่ของรายการแต่ละรายการ

จากนั้นจะกรองเอาเฉพาะรายการที่เกิดขึ้นบ่อยออกมาและสร้างต้นไม้ FP โดยการจัดระเบียบรายการเหล่านี้ตามลำดับความถี่การค้นหาลวดลายที่เกิดขึ้นบ่อย: หลังจากที่เราสร้าง FP-Tree เรียบร้อยแล้ว อัลกอริทึมจะใช้ FP-Tree ในการค้นหาลวดลายที่เกิดขึ้นบ่อยโดยใช้วิธีการเรียกใช้ FP-Growth Algorithm ซึ่งจะค้นหาและรวบรวมลวดลายที่เกิดขึ้นบ่อยจากต้นไม้ FP โดยการแบ่งต้นไม้ออกเป็นหลาย ๆ ส่วนและดำเนินการค้นหาในแต่ละส่วนการสร้างลวดลายสุดท้าย: เมื่อได้ลวดลายที่เกิดขึ้นบ่อยจาก FP-Tree แล้ว ขั้นตอนสุดท้ายคือการสร้างลวดลายสุดท้ายที่สามารถใช้ในการวิเคราะห์ข้อมูลเพิ่มเติมหรือในการทำการคาดการณ์ FP-Growth เป็นอัลกอริทึมที่มีข้อดีอย่างมากในการจัดการกับชุดข้อมูลขนาดใหญ่ เนื่องจากมันสามารถลดปริมาณการคำนวณและการจัดเก็บข้อมูลได้อย่างมีประสิทธิภาพ โดยไม่ต้องทำการสร้างชุดข้อมูลรวมทั้งหมดเหมือนกับอัลกอริทึมแบบเดิม

2.1.7.2 การประยุกต์ใช้งาน FP-Growth

1) การวิเคราะห์ตลาดและการซื้อสินค้า: FP-Growth ใช้ในการวิเคราะห์พฤติกรรมการซื้อของลูกค้า โดยสามารถค้นหาความสัมพันธ์ระหว่างสินค้าที่ลูกค้ามักจะซื้อพร้อมกัน เช่น ถ้าลูกค้าซื้อกาแฟ อาจจะมีแนวโน้มที่จะซื้อครีมบำรุงผิวด้วย การใช้ FP-Growth ช่วยให้ธุรกิจสามารถจัดโปรโมชั่นหรือวางแผนการจัดเรียงสินค้าในร้านค้าได้อย่างมีประสิทธิภาพ

2) การวิเคราะห์ข้อมูลธุรกิจ: ในการจัดการข้อมูลธุรกิจ เช่น ข้อมูลการขายหรือการผลิต FP-Growth สามารถช่วยในการค้นหาลวดลายที่เกิดขึ้นซ้ำในกระบวนการธุรกิจ ซึ่งอาจช่วยในการปรับปรุงประสิทธิภาพการทำงาน หรือการวางแผนทรัพยากรให้เหมาะสม

3) การวิเคราะห์ข้อมูลเว็บ: FP-Growth สามารถนำมาใช้ในการวิเคราะห์ข้อมูลการเข้าชมเว็บไซต์ เช่น การค้นหาความสัมพันธ์ระหว่างหน้าที่ผู้ใช้เข้าชมและพฤติกรรมการใช้เว็บไซต์ ซึ่งช่วยในการออกแบบเว็บไซต์ให้มีประสิทธิภาพมากขึ้นและตอบสนองความต้องการของผู้ใช้ได้ดียิ่งขึ้น

2.1.7.3 ข้อดีของ FP-Growth

1) ประสิทธิภาพสูง FP-Growth ใช้โครงสร้างต้นไม้ (FP-tree) ที่ช่วยในการจัดเก็บข้อมูลในรูปแบบที่ประหยัดพื้นที่และสามารถเข้าถึงข้อมูลได้อย่างรวดเร็ว

2) การประมวลผลที่รวดเร็ว FP-Growth สามารถค้นหาลวดลายการเกิดซ้ำได้โดยไม่ต้องทำการคำนวณซ้ำซ้อนหลายรอบเหมือนอัลกอริทึมแบบดั้งเดิม (schooltha, 2567)

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการจัดการข้อมูลขนาดใหญ่

2.2.1.1 ข้อมูลขนาดใหญ่ (Big Data)

Big Data หรือข้อมูลขนาดใหญ่ มีบทบาทสำคัญในหลายอุตสาหกรรม รวมถึงในวงการการศึกษาและห้องสมุดที่เก็บรวบรวมข้อมูลจำนวนมากจากกิจกรรมต่างๆ ของผู้ใช้ เช่นการยืมคืนหนังสือ การสืบค้นข้อมูลในฐานข้อมูลออนไลน์ และการใช้งานทรัพยากรดิจิทัลต่างๆ ด้วยปริมาณข้อมูลที่มากขึ้น การจัดการและวิเคราะห์ข้อมูลเหล่านี้กลายเป็นงานที่ซับซ้อนและต้องการเทคโนโลยีที่สามารถจัดการข้อมูลขนาดใหญ่ได้อย่างมีประสิทธิภาพ ในการวิจัยที่เกี่ยวข้องกับการเปรียบเทียบโมเดลเพื่อวิเคราะห์ข้อมูลการยืมหนังสือของห้องสมุดมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ การนำทฤษฎีและเทคนิคของการจัดการข้อมูลขนาดใหญ่มาใช้จะช่วยให้สามารถสร้างระบบที่มีประสิทธิภาพสำหรับการวิเคราะห์ข้อมูลที่มีจำนวนมากและซับซ้อนได้อย่างมีประสิทธิภาพ บทความนี้จะอธิบายทฤษฎีพื้นฐานเกี่ยวกับการจัดการข้อมูลขนาดใหญ่ รวมถึงการประยุกต์ใช้ในกรณีศึกษาการเปรียบเทียบโมเดลวิเคราะห์ข้อมูลการยืมหนังสือในห้องสมุดของมหาวิทยาลัย คำว่า Big Data มักจะสื่อถึงชุดข้อมูลที่มีขนาดใหญ่ ซึ่งอาจจะแบ่งย่อยออกเป็นข้อมูล 3 รูปแบบคือ

1) ข้อมูลเชิงโครงสร้าง (Structured data) คือตารางข้อมูลที่มีการจัดเรียงอย่างมีรูปแบบชัดเจนและเป็นระเบียบ สามารถนำมาใช้วิเคราะห์ได้เลย เช่น ข้อมูลที่เก็บไว้ในโปรแกรม spreadsheet อย่าง Microsoft Excel

2) ข้อมูลกึ่งโครงสร้าง (Semi-structured data) คือข้อมูลที่ถูกจัดเก็บอย่างมีรูปแบบในระดับหนึ่งและข้อมูลที่สามารถค้นหา (Search) หรือแท็ก (Tag) ได้ เช่น เว็บไซต์ที่มีการระบุชื่อเพจ คำสำคัญในเพจ และวันที่อัปเดตข้อมูล

3) ข้อมูลที่ไม่มีโครงสร้างแน่นอน (Unstructured data) มักจะหมายถึงข้อความยาวๆ ข้อความเสียง รูปภาพ และวิดีโอ

2.2.1.2 การจัดการข้อมูลขนาดใหญ่

การรวบรวมข้อมูลมาจัดเก็บ (Storage) เป็นการรวบรวมข้อมูลของจากหลากหลายทั้งที่มา และการใช้งานที่แตกต่างกันอย่างมากมาย ซึ่งกลไกและเทคโนโลยีแบบดั้งเดิม ETL (Extract Transform and Load) ไม่สามารถทำได้ ซึ่ง Big Data หรือข้อมูลขนาดใหญ่ต้องการเทคนิค วิธีการ และเทคโนโลยีใหม่ในการรวบรวมข้อมูลขนาดเทราไบต์ และเพตาไบต์ ในการ

รวบรวมข้อมูลนั้นต้องมีการประมวลผลจัดรูปแบบให้เหมาะสมสำหรับการใช้ในการวิเคราะห์หรือใช้งานสำหรับธุรกิจหรือวัตถุประสงค์นั้นๆ

1) การประมวลผล (Processing) ข้อมูลขนาดใหญ่ หรือ Big Data นั้นมีความต้องการสถานที่จัดเก็บขนาดใหญ่ การจัดเก็บข้อมูลขนาดใหญ่จะเป็นชนิดใดก็ได้ไม่ว่าจะเป็นแบบ On-premises หรือแบบ Cloud ขึ้นกับความต้องการหรือความสะดวกในการใช้ ซึ่งเราสามารถเลือกใช้และประเมินผลได้เช่นเดียวกัน บางครั้งก็มีความจำเป็นที่ต้องจัดเก็บไว้ใกล้กับแหล่งข้อมูล หรือข้อมูลบางอันต้องการความยืดหยุ่นสูง และไม่ต้องการบริหารจัดการก็ใช้เป็นแบบ Cloud ซึ่งกำลังเป็นที่นิยมกันเป็นอย่างมาก

2) การวิเคราะห์ (Analyst) การวิเคราะห์ข้อมูลทำให้เกิดความกระจ่าง และชัดเจนในชุดข้อมูลที่มีอยู่ในการสำรวจข้อมูลยังทำให้เกิดค้นพบสิ่งใหม่ แชร์สิ่งที่ค้นพบใหม่ ๆ ต่อผู้อื่น สร้างรูปแบบจำลองข้อมูลด้วยการเรียนรู้ของเครื่องจักร และปัญญาประดิษฐ์ AI และนำข้อมูลเหล่านั้นไปใช้งาน (กิตติชัย ถวิลไทย และภวนาถ จิตรีเชื้อ, 2564)

2.2.2 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

2.2.2.1 หลักในการออกแบบเว็บไซต์

เว็บไซต์เป็นสื่อที่ได้รับความนิยมอย่างมากบนอินเทอร์เน็ต ซึ่งเว็บไซต์เป็นสื่อที่อยู่ในความควบคุมของผู้ใช้โดยสมบูรณ์ กล่าวคือ ผู้ใช้สามารถตัดสินใจเลือกได้ว่า จะดู เว็บไซต์ใดและจะไม่เลือกดูเว็บไซต์ใด ได้ตามต้องการ จึงทำให้ผู้ใช้ไม่มีความอดทนต่ออุปสรรค และปัญหาที่เกิดจากการออกแบบเว็บไซต์ผิดพลาดถ้าผู้ใช้เห็นว่าเว็บที่กำลังดูอยู่นั้นไม่มี ประโยชน์ต่อตัวเขาหรือไม่เข้าใจว่าเว็บไซต์นี้จะใช้งานอย่างไร เขาก็สามารถที่จะเปลี่ยนไปดู เว็บไซต์อื่น ๆ ได้อย่างรวดเร็ว เนื่องจากในปัจจุบันมีเว็บไซต์อยู่มากมาย และยังมีเว็บไซต์ที่เกิดขึ้นใหม่ ๆ ทุกวัน ผู้ใช้จึงมีทางเลือกมากขึ้น และสามารถเปรียบเทียบคุณภาพของเว็บไซต์ต่าง ๆ ได้เองเว็บไซต์ที่ได้รับการออกแบบอย่างสวยงาม มีการใช้งานที่สะดวก ย่อมได้รับความนิยมจากผู้ใช้มากกว่าเว็บไซต์ที่ดู ลับสนวนววย มีข้อมูลมากมายแต่หาอะไรไม่เจอ นอกจากนี้ยังใช้เวลาในการแสดงผลแต่ละหน้านานเกินไป ซึ่งปัญหาเหล่านี้ล้วนเป็นผลมาจากการออกแบบเว็บไซต์ไม่ดีทั้งสิ้น ดังนั้น การออกแบบเว็บไซต์จึงเป็นกระบวนการสำคัญในการสร้างเว็บไซต์ให้ประทับใจผู้ใช้ ทำให้เขาอยากกลับมา เว็บไซต์เดิมอีกในอนาคต ซึ่งนอกจากต้องพัฒนาเว็บไซต์ที่ดีมีประโยชน์แล้วยังต้องคำนึงถึงการ แข่งขันกับเว็บไซต์อื่น ๆ อีกด้วย การออกแบบเว็บไซต์ที่ดีจึงต้องควรคำนึงถึงหลักการดังนี้

1) ความเรียบง่าย เข้าใจง่าย การออกแบบเว็บไซต์ที่ดี จะต้องเน้นที่ความเรียบง่ายเป็นหลัก โดยเลือกนำเสนอเฉพาะสิ่งที่ต้องการนำเสนอจริง ๆ ในรูปแบบที่หลากหลาย โดยอาจจะเป็น สีสัน กราฟิก ภาพเคลื่อนไหวหรือตัวอักษร ที่สำคัญจะต้องมีการนำเสนอที่ไม่ดูรกหน้าเว็บจนเกินไป เพื่อไม่ให้เกิดความรู้สึกรกบกรวนสายตา หรือสร้างความเบื่อหน่าย นำรำคาญให้กับผู้ที่เข้าชมเว็บไซต์

2) ความสม่ำเสมอ ไม่สับสน ควรออกแบบเว็บไซต์ด้วยความสม่ำเสมอ คือจะต้องมีรูปแบบ กราฟิก โทนมสี และการตกแต่งต่าง ๆ ให้แต่ละหน้าบนเว็บไซต์มีความคล้ายคลึงกัน และเป็นแนวเดียวกันไปตลอดทั้งเว็บไซต์ ดังตัวอย่างเว็บไซต์ทั่ว ๆ ไปที่จะสังเกตเห็นได้ว่าทุกหน้าของเว็บไซต์นั้นจะเน้นการตกแต่งในรูปแบบเดียวกันทั้งหมด ต่างก็แค่การนำเสนอของแต่ละหน้าเท่านั้น

3) สร้างความโดดเด่น เป็นเอกลักษณ์ การออกแบบเว็บไซต์เพื่อให้สามารถสื่อถึงจุดประสงค์ในการนำเสนอเว็บได้ดี จะต้องมีการสร้างความเป็นเอกลักษณ์ และจุดเด่นให้กับเว็บไซต์ เพื่อให้สามารถสะท้อนถึงลักษณะขององค์กรได้มากที่สุด โดยการสร้างเอกลักษณ์ดังกล่าวนั้น อาจใช้ชุดสี รูปภาพ ตัวอักษร หรือกราฟิก นอกจากนี้ก็ต้องขึ้นอยู่กับว่าเป็นเว็บไซต์แบบทางการหรือไม่ เพื่อจะได้ออกแบบได้อย่างเหมาะสมที่สุด

4) เนื้อหาต้องดี ครบถ้วน เนื้อหาเป็นสิ่งสำคัญที่สุดของการสร้างเว็บไซต์ เพราะสิ่งที่ทำให้ผู้คน เกิดความสนใจ และหมั่นติดตามเว็บไซต์เหล่านั้นอยู่เสมอ ก็คือเนื้อหาที่มีความสมบูรณ์ และน่าสนใจ นอกจากนี้จะต้องมีการปรับปรุงพัฒนาเนื้อหาบนเว็บให้มีความทันสมัยอยู่เสมอ รวมถึงข้อมูลต้องมีความถูกต้องที่สุด

5) ระบบเนวิเกชัน ใช้งานง่าย ระบบเนวิเกชัน เป็นเสมือนป้ายบอกทางเพื่อให้ผู้ใช้งานไม่เกิดความสับสนในขณะที่ใช้งานเว็บไซต์ ซึ่งการออกแบบเนวิเกชันก็จะต้องเน้นที่ความเรียบง่าย ใช้งาน สะดวก และมีความเข้าใจได้ง่าย ที่สำคัญจะต้องมีตำแหน่งการวางที่สม่ำเสมอเพื่อให้ดูเป็นแนวทางเดียวกัน ทำให้ผู้ใช้งานหรือผู้ชมรู้สึกประทับใจ และจดจำเว็บไซต์ได้ง่ายขึ้น ส่วนใครที่มีการนำกราฟิกมาใช้ในระบบเนวิเกชัน ก็จะต้องเลือกกราฟิกที่สามารถสื่อความหมายได้ดี เช่นกัน

6) คุณภาพของเว็บไซต์ เว็บไซต์ที่ดีจะต้องมีคุณภาพ ทั้งสิ่งที่ปรากฏให้เห็นบนเว็บไซต์ ไม่ว่าจะเป็นกราฟิก ชนิดตัวอักษร รูปภาพหรือสีสันทันที่ใช้ เนื้อหาที่นำมาแสดงผล ซึ่งหากเว็บไซต์มีคุณภาพก็จะสร้างความน่าเชื่อถือ และเป็นจุดเด่นที่ทำให้ผู้คนส่วนใหญ่เกิดความสนใจได้ดี เพราะฉะนั้นห้ามละเลยในส่วนของคุณภาพเด็ดขาด

7) ความสะดวกในการเข้าใช้งาน เว็บไซต์ควรให้ความสะดวกสบายแก่ผู้ใช้งานได้ดี คือจะต้องมีการแสดงผลได้ในทุกระบบปฏิบัติการ ไม่ว่าจะเป็นเว็บเบราว์เซอร์ คอมพิวเตอร์ โน้ตบุ๊ก หรือบนโทรศัพท์มือถือ ที่สำคัญจะต้องมีความละเอียดของการแสดงผล และสามารถใช้งานได้โดยไม่มีปัญหาด้วย (จิราพรพรณ กันทะมูล และพัชรินทร์ ชัยพรหม, 2563)

2.2.3 ทฤษฎีเกี่ยวกับชุดคำสั่ง HTML5

HTML5 เป็นภาษาโปรแกรมที่มีตัวย่อมาจาก Hyper Text Markup Language ถูกพัฒนาขึ้นมาเพื่อรองรับการทำงานของเว็บไซต์ และแอปพลิเคชันบนอินเทอร์เน็ต ในยุคปัจจุบัน HTML5 ได้รับการพัฒนาต่อยอดจากเวอร์ชันก่อนหน้า อย่าง HTML4.01 และ XHTML 1.0 เพื่อให้สามารถรองรับเทคโนโลยีใหม่ๆ และฟีเจอร์ที่มีการใช้ในยุคที่เว็บไซต์มีการตอบสนองและโต้ตอบกับผู้ใช้มากขึ้น เช่น การรองรับการเล่นสื่อมีเดีย (audio และ video) การจัดการกับข้อมูลที่ซับซ้อน และการสร้างแอปพลิเคชันบนเว็บแบบ Real-time อย่างไรก็ตามถือว่าคุณสมบัติ HTML5 เป็นส่วนหนึ่งในการแสดงผลข้อมูลผ่านเว็บไซต์ เป็นการบูรณาการการวิเคราะห์ข้อมูลด้วยโมเดลหลายประเภทเพื่อหาวิธีที่เหมาะสมที่สุด และนำผลลัพธ์มานำเสนอผ่านเว็บไซต์อย่างมีประสิทธิภาพ ซึ่งจะช่วยให้การบริหารจัดการและการให้บริการมีความคล่องตัวและตอบสนองต่อผู้ใช้ได้ดียิ่งขึ้น อีกทั้ง HTML5 ยังเน้นการทำงานร่วมกับ JavaScript อย่างมีประสิทธิภาพมากยิ่งขึ้น โดยมีการพัฒนามาตรฐานที่เรียกว่า WebSockets เพื่อสนับสนุนการสื่อสารแบบสองทางระหว่างเบราว์เซอร์ และเซิร์ฟเวอร์ ซึ่งเป็นสิ่งสำคัญในการพัฒนาแอปพลิเคชันแบบ Real-time เช่น ระบบแชต หรือแอปพลิเคชันที่ต้องการข้อมูลที่อัปเดตตลอดเวลา (ภาณุสรณ์ ใจกลม และสมพงษ์ กาดาสาย, 2564)

(Tag) เป็น คำสั่งหลักของ HTML แทบจะพูดได้ว่า ทุกสิ่งทุกอย่างของ HTML จะขึ้นอยู่กับ tag ทั้งนี้ ไม่ว่าจะเป็น การเน้นข้อความ การแสดงภาพประกอบ หรือการสร้างจุดเชื่อมโยง (link) โปรแกรมเบราว์เซอร์จะตีพิมพ์เอกสารออกมาในรูปแบบใด ก็โดยคำสั่ง tag ทั้งสิ้น

2.2.3.1 ไวยากรณ์ของ HTML จะแบ่งออกเป็น 2 ส่วน คือ

1) ส่วนของคำสั่ง (Tag) เป็นส่วนที่กำหนดรูปแบบของข้อความที่แสดง ซึ่งเราเรียกว่า Tag โดยจะอยู่ในเครื่องหมาย < ... >

2) ส่วนของบทความทั่ว ๆ ไป เป็นส่วนของข้อความที่เราต้องการแสดงผล ตัวอย่างการใช้งานภาษา HTML



ภาพที่ 2.11 โครงสร้างพื้นฐานของ HTML

ที่มา : <https://kru-it.com/html/html-structure>

ในความง่ายของภาษา HTML นั้นเพราะภาษานี้ไม่มีโครงสร้างใด ๆ มากำหนด นอกจาก โครงสร้างพื้นฐานเท่านั้น หรือแม้แต่จะไม่มีโครงสร้างพื้นฐานอยู่ โปรแกรมที่เขียนขึ้นมา นั้นก็สามารถทำงานได้เสมือนมีโครงสร้างทั้งนี้ก็เป็นเพราะว่าตัวโปรแกรมเว็บเบราว์เซอร์จะมองเห็น ทุกสิ่งทุกอย่างใน โปรแกรม HTML เป็นส่วนเนื้อหาทั้งสิ้น ยกเว้นในส่วนหัวที่ต้องมีการกำหนดแยก ออกไปให้เห็นชัดเท่านั้น จะเขียนคำสั่ง หรือข้อความที่ต้องการให้แสดงอย่างไรก็ได้ เป็นเสมือนพิมพ์ งานเอกสารทั่ว ๆ ไป เพียง แต่ทำตำแหน่งใดมีการทำตำแหน่งพิเศษขึ้นมา เว็บเบราว์เซอร์ถึงจะ แสดงผลออกมาตามที่ถูกกำหนด โดยใช้คำสั่งให้ตรงกับรหัสที่กำหนดเท่านั้น

หลังจากมีการพิมพ์โปรแกรมนี้เสร็จเรียบร้อยแล้ว ให้บันทึกเป็นไฟล์ที่มีนามสกุล .html จากนั้นให้เรียกโปรแกรมเว็บเบราว์เซอร์ขึ้นมาทำการทดสอบข้อมูลที่เรารสร้างจะถูกนำ ออกมาแสดงที่จอภาพ ถ้าไม่เขียนอะไรติดบนจอภาพก็จะแสดงผลตามนั้น ถ้าเรามีการปรับปรุง แก้ไขข้อมูลในโปรแกรมเดิมให้อยู่ในรูปของโปรแกรมใหม่ก็จำเป็นต้องโหลดโปรแกรมขึ้นมาใหม่ เพียงแต่เลื่อนเมาส์ไปคลิกที่ปุ่ม Refresh โปรแกรมก็จะทำการประมวลผล และแสดงผลออกมาใหม่ ในคำสั่ง HTML ส่วน ใหญ่ใช้ตัวเปิดเป็นเครื่องหมายน้อยกว่า < ตามด้วยคำสั่ง และปิดท้ายด้วย เครื่องหมายมากกว่า > และมีตัวปิดที่มีรูปแบบเหมือนตัวเปิดเสมอ เพียงแต่จะมีเครื่องหมาย / อยู่ หน้าคำสั่งนั้นๆ เช่น คำสั่งจะมีเป็นคำสั่งปิดเมื่อใดที่ผู้เขียนลืมหรือพิมพ์คำสั่งผิด จะส่งผลให้การ ทำงานของโปรแกรมผิดพลาดทันที คำสั่งเริ่มต้นสำหรับ HTML

คำสั่งหรือ Tag ที่ใช้ในภาษา HTML ประกอบไปด้วยเครื่องหมายน้อยกว่า < ตามด้วยชื่อคำสั่ง และปิดท้ายด้วยเครื่องหมายมากกว่า > เป็นส่วนที่ทำหน้าที่ตกแต่งข้อความ เพื่อการแสดงผลข้อมูลโดยทั่วไปคำสั่งของ HTML ส่วนใหญ่จะอยู่เป็นคู่มีเพียงบางคำสั่งเท่านั้น ที่มีรูปแบบคำสั่งอยู่เพียงตัวเดียว ในแต่ละคำสั่งจะมีคำสั่งเปิด และปิดคำสั่งปิดของแต่ละคำสั่งจะมีรูปแบบเหมือนคำสั่งเปิด เพียงแต่จะเพิ่ม / (Slash) นำหน้าคำสั่งปิดให้ดูแตกต่างเท่านั้น และในคำสั่งเปิดบางคำสั่ง อาจมีส่วนขยายอื่นผสมอยู่ด้วยในการเขียนด้วยตัวอักษรเล็กหรือใหญ่ทั้งหมดหรือเขียนปนกันก็ได้ ไม่มีผลอะไร (วิชา หนูเรื่อง, ชัญญา กลางแก้ว และอิสริยาพร จรจรรย์, 2562)

2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS

CSS ย่อมาจาก Cascading Style Sheet มักเรียกโดยย่อว่า "สไตลชีต" คือภาษาที่ใช้เป็นส่วนของการจัดรูปแบบการแสดงผลเอกสาร โดยที่ CSS กำหนดกฎเกณฑ์ในการระบุรูปแบบ (หรือ "Style") ของเนื้อหาในเอกสาร อันได้แก่ สีของข้อความ สีพื้นหลัง ประเภทตัวอักษร และการจัดวางข้อความ ซึ่งการกำหนดรูปแบบ หรือ Style นี้ใช้หลักการของการแยกเนื้อหาเอกสาร HTML ออกจาก คำสั่งที่ใช้ในการจัดรูปแบบการแสดงผล กำหนดให้รูปแบบของการแสดงผลเอกสาร ไม่ขึ้นอยู่กับเนื้อหาของเอกสาร เพื่อให้ง่ายต่อการจัดรูปแบบการแสดงผลผลลัพธ์ของเอกสาร HTML โดยเฉพาะในกรณีที่มีการเปลี่ยนแปลงเนื้อหาเอกสารบ่อยครั้ง หรือต้องการควบคุมให้รูปแบบการแสดงผลเอกสาร HTML มีลักษณะของความสม่ำเสมอทั่วกันทุกหน้าเอกสารภายในเว็บไซต์เดียวกัน โดยกฎเกณฑ์ในการกำหนดรูปแบบ (Style) เอกสาร HTML ถูกเพิ่มเข้ามาครั้งแรกใน HTML 4.0 เมื่อปี พ.ศ. 2539 ใน รูปแบบของ CSS level 1 Recommendations ที่กำหนดโดยองค์กร World Wide Web Consortium หรือ W3C CSS กับ HTML / XHTML นั้นทำหน้าที่คนละอย่างกัน โดย HTML / XHTML นั้นจะทำหน้าที่ในการวางโครงสร้างเอกสารอย่างเป็นทางการแบบถูกต้อง เข้าใจง่าย ไม่เกี่ยวข้องกับการแสดงผลส่วน CSS จะทำหน้าที่ในการตกแต่งเอกสารให้สวยงาม เรียกได้ว่า HTML / XHTML คือส่วน coding ส่วน CSS คือส่วน design (จิราวุธ วารินทร์, 2555)

```

h1 {
  font-family: courier, courier-new, serif;
  font-size: 20pt;
  color: blue;
  border-bottom: 2px solid blue;
}
p {
  font-family: arial, verdana, sans-serif;
  font-size: 12pt;
  color: #6B6BD7;
}
.red_txt {
  color: red;
}

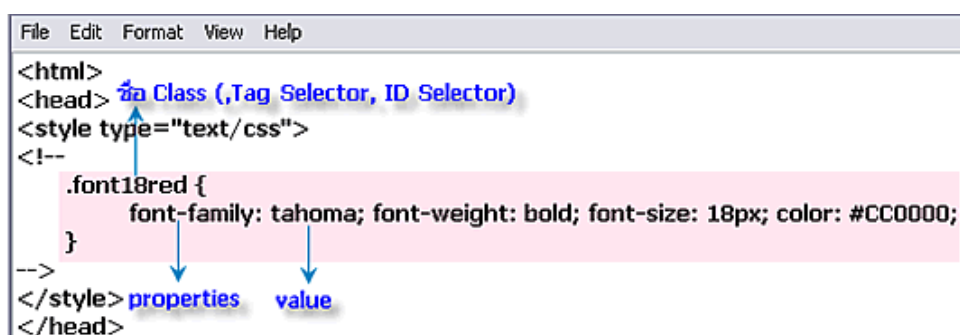
```

ภาพที่ 2.12 ตัวอย่างชุดคำสั่ง CSS

ที่มา : <https://rshweb.medium.com/what-is-css-c302f3aa089a>

2.2.4.1 ไวยากรณ์ของ CSS

- 1) CSS จะประกอบไปด้วย ชื่อ Class หรือชื่อ ID หรือชื่อ Tag HTML ใน CSS จะเรียกว่า Selector ที่ใช้สำหรับการอ้างถึง
- 2) มีชื่อ Properties หรือชื่อคุณสมบัติ และตามด้วย Value เพื่อกำหนดค่าให้แก่คุณสมบัตินั้น และระหว่างชื่อ Properties กับ Value จะถูกคั่นด้วยเครื่องหมาย colon (:)
- 3) คุณสมบัติภายในจะมีได้มากกว่า 1 คุณสมบัติตามที่เรากำหนด และ properties ในแต่ละ properties จะถูกคั่นด้วยเครื่องหมาย semi-colon (;)
- 4) ต้องเขียนคุณสมบัติของ CSS นั้นไว้ภายในเครื่องหมายปีกกา {...} เท่านั้น (กิตติชัย ฤทธิไทย และภูวนาถ จิตรวีเชื้อ, 2564)



ภาพที่ 2.13 ไวยากรณ์ของ CSS

ที่มา : <https://www.dwthai.com>

2.2.5 ทฤษฎีเกี่ยวกับชุดคำสั่ง PHP

ในช่วงแรกภาษาที่นิยมใช้งานบนระบบเครือข่าย คือ ภาษา HTML (Hypertext Markup Language) แต่ภาษา HTML มีลักษณะเป็น Static คือ ภาษาที่มีลักษณะของข้อมูลคงที่ ซึ่งไม่เพียงพอต่อความต้องการในปัจจุบันที่นิยมใช้ระบบเครือข่าย Internet เป็นศูนย์กลางในการติดต่อระหว่างกัน ทำให้ต้องการใช้เว็บไซต์ที่มีลักษณะเป็นแบบ Dynamic คือ เว็บไซต์ที่ข้อมูลสามารถเปลี่ยนแปลงได้โดยอัตโนมัติตามเงื่อนไขต่าง ๆ ที่ผู้เขียนเว็บไซต์เป็นผู้กำหนด และการควบคุมการทำงานเหล่านี้จะกระทำโดยโปรแกรมภาษาสคริปต์ เช่น ภาษา PHP ซึ่งเป็น ภาษาหนึ่งที่ได้รับค่านิยมเป็นอย่างมากในปัจจุบัน

PHP เป็นภาษาสคริปต์ฝั่งเซิร์ฟเวอร์ ที่ใช้ในการพัฒนาเว็บไซต์แบบคงที่หรือเว็บไซต์ไดนามิกหรือแอปพลิเคชันเว็บ PHP ย่อมาจาก Hypertext Pre-processor ซึ่งก่อนหน้านี้ย่อมาจาก Personal Home Pages สคริปต์ PHP สามารถตีความได้บนเซิร์ฟเวอร์ที่ติดตั้ง PHP เท่านั้น คอมพิวเตอร์ไคลเอนต์ที่เข้าถึงสคริปต์ PHP ต้องใช้เว็บเบราว์เซอร์เท่านั้น ไฟล์ PHP มีแท้ก PHP และลงท้ายด้วยนามสกุล “.php”

ภาษาสคริปต์คือ ชุดคำสั่งการเขียนโปรแกรมที่ได้รับการตีความขณะรันไทม์ ภาษาสคริปต์คือภาษาที่ตีความสคริปต์ขณะรันไทม์ โดยปกติสคริปต์จะฝังอยู่ในสภาพแวดล้อมซอฟต์แวร์อื่นโดยปกติแล้ว วัตถุประสงค์ของสคริปต์คือเพื่อปรับปรุงประสิทธิภาพหรือทำงานประจำสำหรับแอปพลิเคชันสคริปต์ฝั่งเซิร์ฟเวอร์ถูกตีความบนเซิร์ฟเวอร์ในขณะที่แอปพลิเคชันไคลเอนต์ตีความสคริปต์ฝั่งไคลเอนต์ PHP เป็นสคริปต์ฝั่งเซิร์ฟเวอร์ที่ถูกตีความบนเซิร์ฟเวอร์ในขณะนั้น [Java ต้นฉบับ](#) เป็นตัวอย่างของสคริปต์ฝั่งไคลเอนต์ที่ถูกตีความโดยเบราว์เซอร์ไคลเอนต์ ทั้ง PHP และ Java สามารถฝังสคริปต์เข้าในหน้า HTML ได้

ภาษาโปรแกรม	ภาษาสคริปต์
มีคุณสมบัติทั้งหมดที่จำเป็นในการพัฒนาแอปพลิเคชันที่สมบูรณ์	ส่วนใหญ่ใช้สำหรับงานประจำ
โค้ดจะต้องได้รับการคอมไพล์ก่อนจึงจะสามารถดำเนินการได้	โดยปกติโค้ดจะถูกดำเนินการโดยไม่ต้องคอมไพล์
ไม่จำเป็นต้องฝังเป็นภาษาอื่น	มักจะฝังอยู่ในสภาพแวดล้อมซอฟต์แวร์อื่นๆ

ภาพที่ 2.14 ภาษาโปรแกรมเทียบกับภาษาสคริปต์

ที่มา : <https://www.guru99.com/th/what-is-php-first-php-program.html>

PHP หมายถึง หน้าแรกส่วนบุคคลแต่ตอนนี้ย่อมาจาก backronym PHP แบบเรียก
ซ้ำ Hypertext Preprocessor โค้ด PHP อาจฝังอยู่ในโค้ด HTML หรือใช้ร่วมกับระบบเทมเพลตเว็บ
ระบบจัดการเนื้อหาเว็บ และกรอบงานเว็บต่าง ๆ ได้

```
<?php
    echo 'Hello World';
?>
```

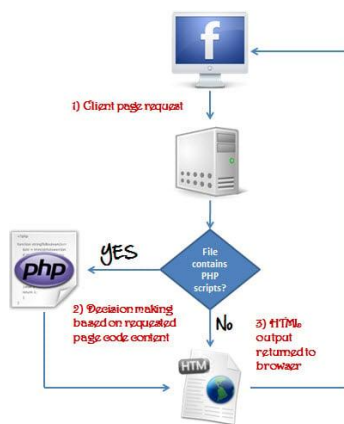
ภาพที่ 2.15 ไวยากรณ์ PHP

ที่มา : <https://www.guru99.com/th/what-is-php-first-php-program.html>

ไฟล์ PHP ยังสามารถมีแท็ก เช่น HTML และสคริปต์ฝังไคลเอ็นต์ เช่น Javaต้นฉบับ

- HTML เป็นข้อได้เปรียบเพิ่มเติม เมื่อเรียนรู้ภาษา PHP คุณสามารถเรียนรู้
PHP ได้โดยไม่ต้องรู้ HTML แต่ขอแนะนำว่าคุณควรมีความรู้พื้นฐานของ HTML เสียก่อน
- ระบบการจัดการฐานข้อมูล DBMS สำหรับแอปพลิเคชันที่ขับเคลื่อนด้วย
ฐานข้อมูล
- คุณจะต้องมีหัวข้อขั้นสูง เช่น แอปพลิเคชันเชิงโต้ตอบและบริการบน
เว็บ JavaScriptและ XML

แผนผังกระแสข้อมูลด้านล่างนี้แสดงให้เห็นสถาปัตยกรรมพื้นฐานของแอปพลิเคชัน
เว็บ PHP และวิธีการที่เซิร์ฟเวอร์จัดการคำขอต่าง ๆ

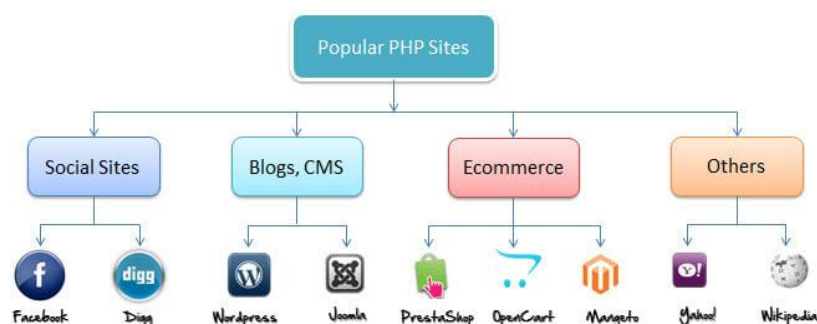


ภาพที่ 2.16 แผนผังกระแสข้อมูล

ที่มา : <https://www.guru99.com/th/what-is-php-first-php-program.html>

ทำไมต้องใช้ PHP คุณคงเคยได้ยินมาบ้างแล้ว [การเขียนโปรแกรมภาษา](#)ข้างนอกนั้น คุณอาจสงสัยว่าทำไมเราถึงต้องการใช้ PHP เป็นพืชต่อการเขียนโปรแกรมเว็บด้านล่างนี้คือเหตุผลที่น่าสนใจบางประการ

- 1) PHP คือ โอเพ่นซอร์สและฟรี
- 2) เส้นโค้งการเรียนรู้ระยะสั้นเมื่อเทียบกับภาษาอื่น เช่น JSP, ASP เป็นต้น
- 3) เอกสารชุมชนขนาดใหญ่
- 4) เซิร์ฟเวอร์เว็บโฮสติ้งส่วนใหญ่รองรับ PHP เป็นค่าเริ่มต้น ไม่เหมือนกับภาษาอื่นๆ เช่น ASP ที่ต้องใช้ IIS ทำให้ PHP เป็นตัวเลือกที่คุ้มค่า
- 5) PHP ได้รับการอัปเดตเป็นประจำเพื่อให้ทันกับแนวโน้มเทคโนโลยีล่าสุด
- 6) ประโยชน์อื่นๆ ที่คุณได้รับจาก PHP ก็คือมันเป็นภาษาสคริปต์ฝั่งเซิร์ฟเวอร์- ซึ่งหมายความว่า คุณจะไม่ต้องติดตั้งบนเซิร์ฟเวอร์เท่านั้น และคอมพิวเตอร์ไคลเอนต์ที่ร้องขอทรัพยากรจากเซิร์ฟเวอร์ไม่จำเป็นต้องติดตั้ง PHP แค่เว็บเบราว์เซอร์ก็เพียงพอแล้ว
- 7) PHP มีในการสร้างการสนับสนุนการทำงานร่วมกัน MySQL- ไม่ได้หมายความว่า คุณไม่สามารถใช้ PHP กับระบบจัดการฐานข้อมูลอื่นได้ คุณยังคงสามารถใช้ PHP กับได้
- 8) PHP คือการข้ามแพลตฟอร์ม ซึ่งหมายความว่า คุณสามารถปรับใช้แอปพลิเคชันของคุณบนระบบปฏิบัติการต่างๆ มากมาย เช่น Windows, Linux, Mac OS เป็นต้น



ภาพที่ 2.17 แผนภาพด้านล่างแสดงเว็บไซต์ยอดนิยมบางส่วนที่ใช้ PHP

ที่มา : <https://www.guru99.com/th/what-is-php-first-php-program.html>

PHP กับ Asp.Net กับ JSP กับ CFML **ASP** – หน้าเซิร์ฟเวอร์ที่ใช้งานอยู่ **JSP** – Java Server Pages, CFML – ภาษา Cold Fusion Markup รูปด้านล่างนี้เปรียบเทียบภาษาสคริปต์ฝั่งเซิร์ฟเวอร์ต่างๆ กับ PHP

คุณสมบัติ	PHP	ASP	JSP	ซีเอฟเอ็มแอล
โครงการเรียนรู้	Short	นานกว่า PHP	นานกว่า PHP	นานกว่า PHP
เว็บโฮสติ้ง	รองรับโดยเซิร์ฟเวอร์โฮสติ้งเกือบทั้งหมด	ต้องการเซิร์ฟเวอร์เฉพาะ	ค่อนข้างสนับสนุน.	ต้องการเซิร์ฟเวอร์เฉพาะ
โอเพ่นซอร์ส	ใช่	ไม่	ใช่	ทั้งเชิงพาณิชย์และโอเพ่นซอร์ส
การสนับสนุนบริการเว็บ	สร้างขึ้นใน	ใช้กรอบงาน.NET	ใช้เพิ่มในไลบรารี	สร้างขึ้นใน
บูรณาการกับ HTML	สะดวกสบาย	ค่อนข้างซับซ้อน	ค่อนข้างซับซ้อน	สะดวกสบาย
MySQL สนับสนุน	พื้นเมือง	ต้องการไดรเวอร์ของบุคคลที่สาม	ต้องการไดรเวอร์ของบุคคลที่สาม	เวอร์ชันปัจจุบันมีการสนับสนุนแบบเนทีฟเวอร์ชันเก่าใช้ ODBC
ขยายได้อย่างง่ายดายด้วยภาษาอื่น	ใช่	ไม่	ขยายการใช้งาน Java ชั้นเรียนและห้องสมุด	ใช่

ภาพที่ 2.18 การเปรียบเทียบ PHP กับ Asp.Net กับ JSP กับ CFML

ที่มา : <https://www.guru99.com/th/what-is-php-first-php-program.html>

PHP ได้รับการออกแบบมาเพื่อทำงานกับ HTML ดังนั้นจึงสามารถฝังลงในโค้ด HTML ได้



ภาพที่ 2.19 แสดงการฝังโค้ด HTML

ที่มา : <https://www.guru99.com/th/what-is-php-first-php-program.html>

คุณสามารถสร้างไฟล์ PHP ได้โดยไม่ต้องมีแท็ก html และเรียกว่าไฟล์ Pure PHP เซิร์ฟเวอร์ตีความโค้ด PHP และแสดงผลผลลัพธ์เป็นโค้ด HTML ไปยังเว็บเบราว์เซอร์ เพื่อให้เซิร์ฟเวอร์ระบุโค้ด PHP จากโค้ด HTML เราจะต้องใส่โค้ด PHP ไว้ในแท็ก PHP เสมอ

แท็ก PHP เริ่มต้นด้วยสัญลักษณ์น้อยกว่า ตามด้วยเครื่องหมายคำถาม จากนั้นคำว่า “php” PHP เป็นภาษาที่คำนึงถึงตัวพิมพ์เล็กและตัวพิมพ์ใหญ่ “VAR” ไม่เหมือนกับ “var” แท็ก PHP นั้นไม่คำนึงถึงตัวพิมพ์เล็ก-ใหญ่ แต่ขอแนะนำอย่างยิ่งให้ใช้ตัวอักษรพิมพ์เล็ก

เราจะอ้างถึงบรรทัดโค้ด PHP เป็นคำสั่ง PHP ลงท้ายด้วยเครื่องหมายอัฒภาค (;) หากคุณมีเพียงคำสั่งเดียว คุณสามารถละเครื่องหมายอัฒภาคได้ หากคุณมีมากกว่าหนึ่งคำสั่ง คุณต้องจบแต่ละบรรทัดด้วยเครื่องหมายอัฒภาค เพื่อความสอดคล้องกัน ขอแนะนำให้คุณลงท้ายคำสั่งของคุณด้วยเครื่องหมายอัฒภาคเสมอ สคริปต์ PHP ดำเนินการบนเซิร์ฟเวอร์ ผลลัพธ์จะถูกส่งกลับในรูปแบบ HTML

สรุป PHP ย่อมาจาก Hypertext pre-processor PHP เป็นภาษาสคริปต์ฝั่งเซิร์ฟเวอร์ ซึ่งหมายความว่ามีการดำเนินการบนเซิร์ฟเวอร์ แอปพลิเคชันไคลเอนต์ไม่จำเป็นต้องติดตั้ง PHP ไฟล์ PHP จะถูกบันทึกด้วยนามสกุลไฟล์ “.php” และโค้ดการพัฒนา PHP อยู่ในแท็ก PHP เป็นไอเฟนซอร์สและข้ามแพลตฟอร์ม (พีโอน่า บราวน์, 2567)

2.2.6 ทฤษฎีเกี่ยวกับกฎความสัมพันธ์ association rule

เป็นเทคนิคที่ใช้ในการค้นหาความสัมพันธ์ระหว่างรายการข้อมูลหรือกลุ่มข้อมูลในชุดข้อมูลขนาดใหญ่ โดยกฎที่ได้จะเป็นรูปแบบของเงื่อนไขหรือข้อสรุปที่บ่งชี้ว่ารายการหนึ่งจะสัมพันธ์กับอีกรายการหนึ่ง เช่นในกรณีของข้อมูลการยืมหนังสือ การค้นหาความสัมพันธ์อาจพบว่าหากนักศึกษายืมหนังสือเกี่ยวกับ "การจัดการธุรกิจ" พวกเขาอาจจะยืมหนังสือเกี่ยวกับ "การบริหารทรัพยากรมนุษย์" ด้วย หนึ่งในเทคนิคที่ได้รับความนิยมในการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์ระหว่างรายการต่างๆ คือ กฎความสัมพันธ์ (Association Rule Mining) โดยเทคนิคนี้ช่วยให้นักวิเคราะห์สามารถค้นหาความสัมพันธ์ระหว่างข้อมูลหลายรายการได้อย่างมีประสิทธิภาพ

Association Rule (กฎความสัมพันธ์) เป็นเทคนิคที่นิยมนำมาใช้ในงานวิจัย สำหรับค้นหาความสัมพันธ์ของข้อมูลในฐานข้อมูลขนาดใหญ่ โดยมีวัตถุประสงค์เพื่อค้นหาความสัมพันธ์ที่ซ่อนอยู่ในข้อมูล ตัวอย่างการนำกฎความสัมพันธ์นี้ไปใช้ ได้แก่ การวิเคราะห์ การซื้อสินค้าของลูกค้า (Market Basket Analysis) ซึ่งวิธีนี้จะหาความสัมพันธ์ของสินค้าที่ลูกค้า มักจะซื้อพร้อมกัน ข้อมูลการซื้อสินค้าของลูกค้าเหล่านี้สามารถนำมาช่วยในการวางแผนทางการตลาด เช่น การจัดวางสินค้าที่มักจะถูกซื้อด้วยกันไว้ใกล้ๆกัน หรือการจัดโปรโมชั่นให้กับสินค้า เป็นต้น (สายชล สมนบูรณ์ทอง, 2558)

2.2.7 ทฤษฎีเกี่ยวกับการออกแบบ Dashboard

แดชบอร์ด (Dashboard) เป็นหน้าจอแสดงผลแบบหน้าเดียวที่นำเสนอข้อมูลในรูปแบบเชิงภาพ เพื่อถ่ายทอดสารสนเทศที่สำคัญภายในขอบเขตที่จำกัดของหน้าหนึ่งหน้า แดชบอร์ดที่ได้รับการออกแบบอย่างมีประสิทธิภาพจะประกอบด้วยองค์ประกอบหลักที่สำคัญที่สุดของข้อมูล ช่วยให้สามารถตรวจสอบและวิเคราะห์ข้อมูลทางธุรกิจ ค้นหาคำตอบ และติดตามตัวชี้วัดที่สำคัญ (Key Metrics) ได้อย่างรวดเร็ว การแสดงผลบนแดชบอร์ดอาจมาจากชุดข้อมูลพื้นฐานเพียงชุดเดียวหรือหลายชุดข้อมูล รวมถึงสามารถดึงข้อมูลจากรายงานเดี่ยวหรือหลายรายงาน ทั้งนี้ แดชบอร์ดยังสามารถผสานรวมข้อมูลจากแหล่งภายในองค์กรและระบบคลาวด์เพื่อให้มุมมองข้อมูลแบบองค์รวมโดยไม่จำกัดแหล่งที่มาของข้อมูล นอกจากนี้ แดชบอร์ดยังสามารถโต้ตอบกับผู้ใช้ได้ โดยมีการอัปเดตองค์ประกอบต่าง ๆ เช่น ไทล์ (Tiles) หรือกราฟ เมื่อข้อมูลพื้นฐานมีการเปลี่ยนแปลง

สำหรับการสร้างแดชบอร์ด มีเครื่องมือที่สามารถนำมาใช้ได้หลากหลาย หากเป็นซอฟต์แวร์พื้นฐานที่มีอยู่แล้ว เช่น Microsoft Excel สามารถใช้ฟังก์ชัน Pivot Table และ Pivot Chart ในการแสดงผลข้อมูล อย่างไรก็ตาม หากต้องจัดการข้อมูลที่มีปริมาณมาก อาจจำเป็นต้องใช้เครื่องมือด้าน Business Intelligence (BI) เพื่อเพิ่มประสิทธิภาพในการวิเคราะห์และนำเสนอข้อมูล เช่น Power BI Desktop, Tableau, SimpleKPI หรือ InfoCaptor ซึ่งช่วยให้การวิเคราะห์ข้อมูลเป็นไปอย่างมีประสิทธิภาพ และสามารถนำเสนอข้อมูลเชิงลึกได้อย่างชัดเจน (วรวิษ วิทยะกิจ และอนุชา รินทาวุธ, 2562)

2.2.7.1 แดชบอร์ด (Dashboard) สามารถแบ่งออกเป็น 3 ประเภทใหญ่ ๆ คือ

1) Strategic dashboard จะเป็นข้อมูลที่ใช้สำหรับประกอบการวางแผนด้านกลยุทธ์ การวางแผนที่มีการกำหนดวิสัยทัศน์ มีการกำหนดเป้าหมายระยะยาวที่แน่ชัด มีการวิเคราะห์อนาคตและคิดเชิงการแข่งขันที่ต้องการระบบการทำงานที่มีความสามารถในการปรับตัวสูง สำหรับการทำงานในสิ่งแวดล้อมที่มีการเปลี่ยนแปลงอย่างรวดเร็ว เพื่อให้ทันกับการเปลี่ยนแปลงที่เกิดขึ้น ต้องการระบบการทำงานที่คล่องตัว ต้องการดำเนินงานมีประสิทธิภาพสูงในการนำสู่เป้าหมายในอนาคต

2) Analytical Dashboard เป็นแดชบอร์ดที่ใช้สำหรับการวิเคราะห์ข้อมูล โดยนำข้อมูลที่ได้จากการปฏิบัติงานและรวบรวมผ่านเครื่องมือวัดต่าง ๆ (Instruments) มาจัดการและจำแนกกลุ่ม จัดประเภท คำนวณค่า สรุปผล และนำเสนอในรูปแบบที่เหมาะสม เพื่อสนับสนุนการตัดสินใจและการวิเคราะห์เชิงลึก

3) Operational Dashboard เป็นแดชบอร์ดที่ใช้สำหรับการสนับสนุนการดำเนินงานขององค์กร โดยนำเสนอข้อมูลสรุปที่เกี่ยวข้องทั้งจากแหล่งข้อมูลภายในและภายนอกองค์กร ซึ่งเป็นข้อมูลสำคัญที่ใช้เป็นพื้นฐานในการวางแผน และบริหารจัดการองค์กรอย่างมีประสิทธิภาพ

2.2.7.2 หลักการออกแบบแดชบอร์ด 4 ขั้นตอน

1) Users การทำความเข้าใจผู้ใช้เป็นขั้นตอนแรกและสำคัญที่สุดของการออกแบบแดชบอร์ด โดยใช้หลักการของ Design Thinking เพื่อค้นหา Insight ของผู้ใช้และตอบคำถามว่าผู้ใช้งานต้องการนำแดชบอร์ดไปใช้เพื่อวัตถุประสงค์ใด เช่น การตัดสินใจทางธุรกิจ การติดตามผลการดำเนินงาน หรือการแจ้งเตือนเหตุการณ์ผิดปกติ การเก็บ Requirement จากผู้ใช้เป็นสิ่งสำคัญเพื่อระบุว่าผู้ใช้งานต้องการดูข้อมูลประเภทใด และต้องการการแสดงผลในรูปแบบใด เช่น กราฟ ตาราง หรือแผนภูมิ หากผู้พัฒนาเข้าใจรูปแบบการใช้งานของผู้ใช้ได้อย่างถูกต้อง จะสามารถออกแบบแดชบอร์ดที่ตรงกับความต้องการได้มากขึ้น

2) Content เมื่อเข้าใจวัตถุประสงค์ของผู้ใช้แล้ว จำเป็นต้องวิเคราะห์เนื้อหาหรือข้อมูลที่ต้องการนำเสนอ ซึ่งประกอบด้วย measures (ตัวเลขหรือค่าที่สนใจ เช่น ยอดขาย จำนวนลูกค้า) และ dimensions (มุมมองในการวิเคราะห์ข้อมูล เช่น ช่วงเวลา ประเภทสินค้า หรือพื้นที่) การเลือก measures และ dimensions อย่างเหมาะสมจะช่วยให้การวัดผลมีประสิทธิภาพและสอดคล้องกับเป้าหมายการใช้งาน

3) Presentation การเลือกประเภทของกราฟให้เหมาะสมกับลักษณะของข้อมูลและวัตถุประสงค์การนำเสนอเป็นอีกขั้นตอนสำคัญ เพื่อให้แดชบอร์ดสามารถสื่อสารข้อมูลได้อย่างครบถ้วนและชัดเจน

4) Navigation การจัดวางองค์ประกอบของแดชบอร์ดต้องคำนึงถึงความสัมพันธ์ของข้อมูล โดยกราฟที่มีเนื้อหาเกี่ยวข้องกันควรจัดวางไว้ใกล้กันเพื่อลดความสับสนของผู้ใช้ นอกจากนี้ ควรใช้แนวคิด visual hierarchy ในการออกแบบ โดยนำเสนอข้อมูลจากภาพรวมก่อนลงไปสู่รายละเอียด การแสดงผลข้อมูลในแดชบอร์ดมักอยู่ในรูปแบบของ tiles ซึ่งสามารถปักหมุดไปยังแดชบอร์ดหรือเชื่อมโยงไปยังรายงานที่เกี่ยวข้องเพื่อการวิเคราะห์เพิ่มเติม

2.3 วรรณกรรมที่เกี่ยวข้อง

วิเศษฐ์ นันทะศรี และจรัญ แสงราช (2564) ได้ทำการวิจัยเรื่อง การใช้เทคนิคเหมืองข้อมูล เพื่อวิเคราะห์พฤติกรรมการใช้บริการห้องสมุดของนักศึกษา วิทยาลัยการศึกษาคณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏเลย

งานวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์พฤติกรรมการเข้าใช้บริการห้องสมุดของนักศึกษาโดยใช้เทคนิคเหมืองข้อมูล ความสัมพันธ์ โดยใช้ข้อมูลการเข้าใช้บริการห้องสมุดของนักศึกษาระดับปริญญาตรี ภาคปกติ มหาวิทยาลัยราชภัฏเลย ปีการศึกษา 2559 โดยใช้เทคนิคเหมืองข้อมูลเพื่อวิเคราะห์พฤติกรรมการใช้บริการห้องสมุดของนักศึกษา

ผลงานวิจัยนี้สรุปได้ว่าพฤติกรรมการเข้าใช้ห้องสมุด ส่งผลต่อผลการเรียนของนักศึกษา ดังนั้นมหาวิทยาลัย ควรจะต้องคิดกิจกรรมใหม่ ๆ เพิ่มเพื่อที่จะได้ส่งเสริมให้นักศึกษาที่มีผลการเรียนเฉลี่ยต่ำให้สนใจเข้าห้องสมุดเพิ่มมากขึ้น เพื่อเป็นการกระตุ้นให้นักศึกษาเกิดการใฝ่หาความรู้ ในการพัฒนาตนเองอย่างต่อเนื่อง

ธนพร เพ็ญขจร (2564) ได้ทำการวิจัยเรื่อง การพัฒนาบริการแนะนำหนังสือโดยประยุกต์ใช้เทคนิคเหมืองข้อมูลและเทคโนโลยีความเป็นจริงเสริม

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความสัมพันธ์ของหนังสือตามพฤติกรรมการยืมของผู้ใช้โดยใช้เทคนิคเหมืองข้อมูล และเพื่อพัฒนาบริการแนะนำหนังสือตามพฤติกรรมการยืมหนังสือ โดยประยุกต์ใช้เทคโนโลยีความเป็นจริงเสริม เครื่องมือที่ใช้ในการวิจัย คือ 1) บริการแนะนำหนังสือโดยใช้เทคโนโลยีความเป็นจริงเสริม 2) แบบประเมินความเหมาะสมของบริการแนะนำหนังสือ โดยใช้เทคโนโลยีความเป็นจริงเสริม

ผลการวิจัยพบว่า การวิเคราะห์ความสัมพันธ์ของหนังสือตามพฤติกรรมการยืมของผู้ใช้โดยใช้เทคนิคเหมืองข้อมูล มีความสัมพันธ์ของหนังสือที่ผู้ใช้มักจะยืมพร้อมกัน จำนวน 48 กลุ่ม และแบ่งออกได้เป็น 6 รูปแบบ ส่วนผลการประเมินความเหมาะสมของบริการแนะนำหนังสือ โดยใช้เทคโนโลยีความเป็นจริงเสริม พบว่า ความเหมาะสมในภาพรวมอยู่ในระดับมาก โดยทั้ง 3 ด้าน คือด้านการออกแบบ ด้านการใช้งาน และด้านเนื้อหา มีความเหมาะสมอยู่ในระดับมาก

พงศกร สุกันยา (2565) ได้ทำการวิจัยเรื่อง พัฒนารูปแบบการแนะนำหนังสือด้วยเทคนิคเหมืองข้อมูล วิทยาลัยการศึกษาคณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏเลย การแพทย์ วิทยาลัยการศึกษาคณะศึกษาศาสตร์ มหาวิทยาลัยราชภัฏเลย

การวิจัยนี้เป็นการวิจัยเชิงทดลอง มีวัตถุประสงค์เพื่อ 1) วิเคราะห์หากฎความสัมพันธ์ของการใช้หนังสือ ของห้องสมุดศูนย์การแพทย์ปัญญานันทภิกขุ ชลประทาน มหาวิทยาลัยศรีนครินทรวิโรฒ ด้วยเทคนิคเหมืองข้อมูล และ 2) นำกฎความสัมพันธ์ที่ได้จากเทคนิคเหมืองข้อมูลมาประยุกต์ใช้สำหรับการแนะนำหนังสือของห้องสมุดศูนย์การแพทย์ปัญญานันทภิกขุ ชลประทาน มหาวิทยาลัยศรีนครินทรวิโรฒ และประเมินประสิทธิภาพการใช้งาน โดยแบ่งวิธีการดำเนินการวิจัย 2 ระยะ

ผลการวิจัยระยะที่ 1 การวิเคราะห์หากฎความสัมพันธ์ของการใช้หนังสือของห้องสมุดศูนย์การแพทย์ปัญญานันทภิกขุ ชลประทาน มหาวิทยาลัยศรีนครินทรวิโรฒ ด้วยเทคนิคเหมืองข้อมูลพบว่า มีรายการหนังสือที่มีโอกาสยืมพร้อมกัน มีจำนวน 11 รายการ ด้วยค่าความเชื่อมั่นมากกว่าหรือ เท่ากับร้อยละ 50

ผลการวิจัยระยะที่ 2 การนำกฎความสัมพันธ์ที่ได้จากเทคนิคเหมืองข้อมูลมาประยุกต์ใช้สำหรับการแนะนำหนังสือของห้องสมุดศูนย์การแพทย์ปัญญานันทภิกขุ ชลประทาน มหาวิทยาลัยศรีนครินทรวิโรฒ และประเมินประสิทธิภาพการใช้งานพบว่า การพัฒนาระบบแนะนำหนังสือห้องสมุดศูนย์การแพทย์ปัญญานันทภิกขุ ชลประทาน มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นต้นแบบสำหรับการแนะนำหนังสือ โดยใช้โปรแกรมจำลองเครื่องคอมพิวเตอร์เป็น Web Server (Apache) โปรแกรมจัดการฐานข้อมูล Mysql (phpMyAdmin) สำหรับบริหารฐานข้อมูลที่พัฒนาโดย PHP เพื่อใช้เชื่อมต่อไปยังฐานข้อมูล และCodeIgniter Framework สำหรับการพัฒนาเว็บไซต์ ส่วนผลการประเมินประสิทธิภาพการใช้งานระบบแนะนำหนังสือพบว่า กลุ่มตัวอย่างมีความพึงพอใจการใช้ระบบแนะนำหนังสือภาพรวมอยู่ในระดับมาก

อนันตญา ใจดี (2563) ได้ทำการวิจัยเรื่อง การประยุกต์การเรียนรู้ด้วยกฎความสัมพันธ์ และการวิเคราะห์องค์ประกอบหลักเพื่อระบุปัจจัยที่มีความสัมพันธ์ต่อการตัดสินใจส่งจ่ายยาปฏิชีวนะอย่างสมเหตุผลสำหรับการติดเชื้อทางเดินหายใจส่วนบนในผู้ป่วยนอก

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีความสัมพันธ์ต่อการตัดสินใจส่งจ่ายยาปฏิชีวนะอย่างสมเหตุผลสำหรับการติดเชื้อทางเดินหายใจส่วนบนในผู้ป่วยนอกของโรงพยาบาลอินทร์บุรี จังหวัดสิงห์บุรี และนำผลที่ได้จากการศึกษาเสนอแนะแนวทางการให้ความรู้เพื่อป้องกันการส่งจ่ายยาปฏิชีวนะอย่างไม่สมเหตุผลสำหรับการติดเชื้อทางเดินหายใจส่วนบนในผู้ป่วยนอก โดยวิเคราะห์หาปัจจัยที่มีความสัมพันธ์ต่อการตัดสินใจส่งจ่ายยาปฏิชีวนะอย่างสมเหตุผลด้วย 2 วิธี ได้แก่ วิธีที่ 1 วิธีสถิติไคสแควร์ร่วมกับวิธีวิเคราะห์องค์ประกอบหลักสร้างเป็นแบบจำลอง

พยากรณ์ที่ 1 และวิธีที่ 2 วิธีวิเคราะห์องค์ประกอบหลักเพียงวิธีเดียวสร้างเป็นแบบจำลอง พยากรณ์ที่ 2 จากนั้นวิเคราะห์หารูปแบบความสัมพันธ์ของปัจจัยที่มีความสัมพันธ์ต่อการตัดสินใจสั่งจ่ายยาปฏิชีวนะอย่างสมเหตุผลด้วย Apriori algorithm

ผลจากการศึกษาพบว่า ปัจจัยที่ได้จากวิธีวิเคราะห์วิธีที่ 1 จำนวน 27 ปัจจัย วิเคราะห์หารูปแบบความสัมพันธ์ด้วย Apriori ได้จำนวน 18 รูปแบบ ส่วนปัจจัย ที่ได้จากวิธีวิเคราะห์วิธีที่ 2 จำนวน 50 ปัจจัย วิเคราะห์หารูปแบบความสัมพันธ์ด้วย Apriori ได้ จำนวน 2 รูปแบบ แบบจำลองพยากรณ์การตัดสินใจสั่งจ่ายยาปฏิชีวนะอย่างสมเหตุผลทั้ง 2 แบบจำลองสามารถทำนายการตัดสินใจสั่งจ่ายยาปฏิชีวนะอย่างสมเหตุผลได้ถูกต้องแม่นยำสูงใกล้เคียงกัน

เบญจมาวรรณ ใจสูงเนิน (2564) ได้ทำการวิจัยเรื่อง ระบบแนะนำหนังสืออ่านเพิ่มเติม (Supplementary books recommendation system)

งานวิจัยนี้มีวัตถุประสงค์เพื่อแนะนำหนังสือโดยพิจารณาจากความสอดคล้องของหนังสือที่เพิ่มเติมจากการศึกษาในชั้นเรียน ซึ่งเป็นหนังสือที่เกี่ยวข้องกับคำค้นที่เฉพาะเจาะจงเป็นหลัก โดยจะมุ่งเน้นที่การแนะนำหนังสืออ่านเพิ่มเติมจากการศึกษาในมหาวิทยาลัยผ่านการพิจารณาคำอธิบายรายวิชา และประมวลรายวิชา โดยรูปแบบการแนะนำหนังสืออ่านเพิ่มเติมจะเป็นการสร้างรายการของหนังสือที่มีความเกี่ยวข้องกับคำค้นหรือกลุ่มของคำค้นที่ผู้อ่านกำหนด โดยหนังสือในรายการที่ถูกระบุจะถูกระบุเรียงลำดับตามความเกี่ยวข้อง (เรียงลำดับจากความเกี่ยวข้องกับคำค้นมากที่สุดไปน้อยที่สุด) ซึ่งรายการแนะนำหนังสือที่ได้จะช่วยให้ผู้อ่านสามารถตัดสินใจที่จะเลือกหนังสืออ่านเพิ่มเติมได้ง่าย และสะดวกยิ่งขึ้น เพื่อส่งเสริมการอ่าน และการเรียนรู้ด้วยตนเอง อันนำมาซึ่งการพัฒนาตนเอง

ผลการวิจัยพบว่าระบบแนะนำหนังสือที่พัฒนาขึ้นสามารถช่วยเพิ่มประสิทธิภาพในการเลือกหนังสืออ่านเพิ่มเติมของผู้ใช้ได้เป็นอย่างดี โดยเฉลี่ยแล้วผู้ใช้ให้คะแนนความพึงพอใจต่อระบบที่ค่าเฉลี่ย ($\bar{X}=4.45$) ซึ่งแสดงให้เห็นว่าระบบสามารถตอบสนองความต้องการของผู้ใช้ได้เป็นอย่างดี นอกจากนี้ระบบยังได้รับการทดสอบกับกลุ่มผู้ใช้หลายกลุ่มพบว่ามีความแม่นยำในการแนะนำหนังสือสูง และช่วยเพิ่มโอกาสในการค้นพบหนังสือที่ผู้ใช้อาจสนใจแต่ไม่เคยรู้จักมาก่อน ข้อมูลเพิ่มเติมจากงานวิจัยนี้แสดงให้เห็นว่า การใช้เทคนิคการกรองข้อมูลและการวิเคราะห์พฤติกรรมผู้ใช้ในการแนะนำหนังสือสามารถช่วยเสริมสร้างประสบการณ์การอ่านที่ดีขึ้น และทำให้ผู้ใช้ได้รับความรู้และเนื้อหาที่ตรงกับความสนใจได้อย่างมีประสิทธิภาพ

อับดุลเลาะ บากา, สุลัยมาน เกอโสะ, อิสมาแอ ล่าเตะเกะ, อรรถพล อุดุลยศาสน์, พรรณี แพงทิพย์ และฟูโดละห์ ดือมอง (2565) ได้ทำการวิจัยเรื่อง การหากฎความสัมพันธ์ด้วยขั้นตอนวิธี เอพี-โกรธ เพื่อวิเคราะห์พฤติกรรมกรรมการซื้อสินค้าของลูกค้าร้านค้าปลีก กรณีศึกษาร้านค้าที่ทรูซ์ อปจังหวัดปัตตานี

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ศึกษารูปแบบพฤติกรรมกรรมการซื้อสินค้าของลูกค้าร้านค้าปลีกโดยใช้ขั้นตอนวิธีเอพี-โกรธ 2) วิเคราะห์พฤติกรรมกรรมการซื้อสินค้าของลูกค้าร้านค้าปลีกเพื่อการส่งเสริมการตลาด และ 3) ศึกษาผลการทดลองใช้รูปแบบพฤติกรรมกรรมการซื้อสินค้าของลูกค้าร้านค้าปลีกสำหรับการส่งเสริมการตลาด กลุ่มตัวอย่างที่ใช้ในการศึกษาในครั้งนี้ คือชุดข้อมูลรายการขายรายวันจำนวน 1,377,853 รายการ จากฐานข้อมูลของร้านดีทรูซ์ออปจังหวัดปัตตานี ตั้งแต่ปี พ.ศ. 2562 - 2564 มีทั้งหมด 43 คุณลักษณะ การออกแบบงานวิจัยนี้เป็นไปตามระเบียบวิธีการ CRISP-DM และวิเคราะห์ข้อมูลผ่านโปรแกรม RapidMiner Studio เวอร์ชัน 9 โดยใช้ขั้นตอนวิธีของเอพี-โกรธ ในการค้นหากลุ่มข้อมูลที่ปรากฏร่วมกันบ่อย

ผลการวิจัยพบว่า 1) การศึกษารูปแบบพฤติกรรมกรรมการซื้อสินค้าจากฐานข้อมูลร้านค้าปลีก พบว่าได้จำนวนรูปแบบพฤติกรรมกรรมการซื้อสินค้าทั้งหมด 13 กฎความสัมพันธ์ เมื่อกำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 0.01 และค่าความเชื่อมั่นขั้นต่ำเท่ากับ 0.20 2) ผลการวิเคราะห์พฤติกรรมกรรมการซื้อสินค้าของลูกค้าร้านค้าปลีกเพื่อการส่งเสริมการตลาด พบว่ามีเพียง 4 กฎความสัมพันธ์ เท่านั้นที่มีระดับความเชื่อมั่นสูงกว่าเกณฑ์ร้อยละ 30 และ 3) การศึกษาผลการทดลองใช้รูปแบบพฤติกรรมกรรมการซื้อสินค้าของลูกค้าร้านค้าปลีก สำหรับการจัดรายการส่งเสริมการขาย และการปรับปรุงการออกแบบชั้นวางสำหรับร้านค้าเพื่อเพิ่มยอดขาย พบว่ายอดขายเฉลี่ยเพิ่มขึ้นร้อยละ 5 เมื่อเทียบกับช่วง 3 เดือนที่ผ่านมา

กริชบดินทร์ ผิวหอม (2567) ได้ทำการวิจัยเรื่อง การเปรียบเทียบประสิทธิภาพอัลกอริทึม Apriori และ FP-Growth ด้วยชุดข้อมูลร้านขายของชำ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาและเปรียบเทียบประสิทธิภาพของอัลกอริทึม Apriori และ FP Growth ในการค้นหาไอเทมเซตที่เกิดขึ้นบ่อย (Frequent itemsets) ซึ่งเป็นขั้นตอนสำคัญในกระบวนการทำเหมือง ข้อมูลกฎความสัมพันธ์ (Association rules mining) และ 2) เพื่อเปรียบเทียบความเร็ว และการใช้หน่วยความจำในการค้นหาไอเทมเซตที่เกิดขึ้นบ่อย

ผลการวิจัยพบว่า 1) แต่ละอัลกอริทึมที่ใช้ในการค้นหาไอเทมเซตที่เกิดขึ้นบ่อยมีข้อดีและข้อเสียที่แตกต่างกัน สำหรับชุดข้อมูลและค่าสนับสนุนที่แตกต่างกัน และ 2) การเปรียบเทียบ

ความเร็วและการใช้หน่วยความจำ ชี้ให้เห็นว่าอัลกอริทึม FP-Growth ทำงานได้เร็วกว่าอัลกอริทึม Apriori และใช้หน่วยความจำน้อยกว่า เนื่องจากอัลกอริทึม Apriori ต้องสแกนฐานข้อมูลหลายรอบ และมีการสร้างชุดข้อมูลตัวแทน (Candidate itemsets) ทำให้ใช้เวลานาน ส่วนอัลกอริทึม FP-Growth เข้าสแกนฐานข้อมูลแค่ 2 รอบ และไม่มีการสร้างชุดตัวแทน ทำให้ทำงานได้ดีกว่า

ชิตนภัส มาศชาย, นุชนารถ ศรีษะโหนด, วิสูตร มาศชาย และวิภาวรรณ บัวทอง (2564) ได้ทำการวิจัยเรื่อง การหาความสัมพันธ์การประเมินสมรรถภาพทางกายของนักเรียน โดยใช้เทคนิคการทำเหมืองข้อมูล กรณีศึกษาโรงเรียนราชประชานุเคราะห์

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์หาความสัมพันธ์การประเมินสมรรถภาพทางกาย ในการดำเนินการวิจัยได้เลือกกลุ่มตัวอย่างนักเรียนโรงเรียนราชประชานุเคราะห์ 1 จำนวน 812 คน โดยใช้หลักการการทำเหมืองข้อมูลแบบกฎความสัมพันธ์ (Association Rule) โดยใช้ขั้นตอนวิธีแบบ Apriori และ FP-Growth และการทำเหมืองแบบจำแนกประเภทข้อมูล (Classification) โดยเลือกการเปรียบเทียบข้อมูลด้วยขั้นตอนวิธี 2 วิธี คือ ต้นไม้ตัดสินใจ (Decision Tree) แบบ J48 และป่าสุ่ม (Random Forest) ซึ่งสร้างตัวแบบและหาความสัมพันธ์ด้วยโปรแกรม Weka 3.8.5

ผลการวิจัยพบว่า การทำเหมืองข้อมูลแบบจำแนกประเภทข้อมูลของการทดสอบสมรรถภาพทางกายของนักเรียนโดยใช้ต้นไม้ตัดสินใจ และการประเมินแบบจำลองด้วยวิธี Percentage split ร้อยละ 70 มีประสิทธิภาพมากที่สุด โดยได้ค่าความแม่นยำ ร้อยละ 100 ค่าความไว ร้อยละ 100 และค่าวัดประสิทธิภาพโดยรวม ร้อยละ 100 และการทำเหมืองข้อมูลแบบกฎความสัมพันธ์ แบบ Apriori และ FP Growth ได้กำหนดค่า Minimum Support เท่ากับ 0.1 และค่า Minimum Confidence เท่ากับ 0.9 ได้กฎความสัมพันธ์ จำนวน 5 ข้อ และ 4 ข้อตามลำดับ ซึ่งกฎที่ได้ทั้ง 2 แบบมีความสัมพันธ์กัน ทั้งยังพบว่ากฎที่ได้จากกฎความสัมพันธ์ และแบบจำแนกประเภทข้อมูลให้ผลลัพธ์คล้ายกัน และไปในทิศทางเดียวกัน

สุภาพรพรณ ดงมณีพรณ (2564) ได้ทำการวิจัยเรื่อง การหาความสัมพันธ์จากฐานข้อมูลการซื้อขายผลิตภัณฑ์เสริมอาหารยี่ห้อมายเฮลท์ ของลูกค้าโดยใช้อัลกอริทึมเอพพี-โกรธ และการแบ่งกลุ่มลูกค้าตามพฤติกรรมกรรมการซื้อสินค้ามายเฮลท์ โดยเทคนิคอาร์เอฟเอ็มด้วยโปรแกรม RapidMiner กรณีศึกษาของร้านยาเซนแห่งหนึ่ง ค้นหากฎความสัมพันธ์ด้วยอัลกอริทึม FP-Growth และจัดกลุ่มลูกค้าตามพฤติกรรมกรรมการซื้อสินค้าด้วยเทคนิค RFM รวมทั้งการค้นหากฎความสัมพันธ์ของการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH ของแต่ละกลุ่มลูกค้า และ

พฤติกรรมกรรมการซื้อสินค้า โดยใช้วิธีการทำเหมืองข้อมูล ด้วยข้อมูลของลูกค้าที่ซื้อสินค้า MYHEALTH ภายใต้กรอบ CRISP-DM

งานวิจัยนี้จัดทำเหมืองข้อมูลกับเพื่อค้นหาความสัมพันธ์จากข้อมูลการซื้อสินค้าของลูกค้า 93,278 แถว 2 คอลัมน์ และการค้นหาการแบ่งกลุ่มลูกค้าด้วยข้อมูลการซื้อสินค้า 63,860 แถว 5 คอลัมน์

ผลการวิจัยแสดงให้เห็นถึงความสัมพันธ์ของการซื้อผลิตภัณฑ์เสริมอาหารยี่ห้อ MYHEALTH จำนวน 7 กฎ ตัวอย่างหนึ่งของความสัมพันธ์ คือ จากรายการซื้อผลิตภัณฑ์หนึ่งพันรายการพบว่ามี 1 รายการ หากลูกค้าซื้อ L-Carnitine และRoyal Jelly แล้วมีโอกาสที่จะซื้อ Carbo-X รวมด้วย ถึงร้อยละ 71.4 และสามารถแบ่งกลุ่มลูกค้าเป็น 6 กลุ่ม ได้แก่ Blue Diamond, Diamond, Platinum, Gold, Silver, Welcome กล่าวคือ 1) ลูกค้ากลุ่ม Blue Diamond เมื่อซื้อผลิตภัณฑ์ CALCIUM และCARBO-X แล้วมีโอกาสจะซื้อผลิตภัณฑ์ L-CARNITINE รวมด้วย 2) ลูกค้ากลุ่ม Diamond เมื่อซื้อผลิตภัณฑ์ CARBO-X และRoyal Jelly แล้วมีโอกาสจะซื้อผลิตภัณฑ์ L-CARNITINE รวมด้วย 3) ลูกค้ากลุ่ม Platinum เมื่อซื้อผลิตภัณฑ์ Fish Oil และBilberry แล้วมีโอกาสจะซื้อผลิตภัณฑ์ Golden oil รวมด้วย 4) ลูกค้ากลุ่ม Gold เมื่อซื้อผลิตภัณฑ์ L-CARNITINE และ Acerola Cherry แล้วมีโอกาสจะซื้อผลิตภัณฑ์ CARBO-X รวมด้วย 5) ลูกค้ากลุ่ม Silver เมื่อซื้อผลิตภัณฑ์ Calcium และCARBO-X มีโอกาสจะซื้อผลิตภัณฑ์ L-CARNITINE รวมด้วย 6) ลูกค้ากลุ่ม Welcome เมื่อซื้อผลิตภัณฑ์ Rice Bran Oil และCARBO-X แล้วมีโอกาสจะซื้อผลิตภัณฑ์ L-CARNITINE รวมด้วย ซึ่งความสัมพันธ์เหล่านี้สามารถช่วยในการจัดทำแผนการตลาด เพื่อตอบสนองความต้องการของลูกค้า และส่งเสริมให้เกิดการซื้ออย่างต่อเนื่องของลูกค้าได้ นอกจากนี้บริษัทยังสามารถสร้างแผนจัดตำแหน่งสินค้า เพื่อเพิ่มยอดขายให้กับธุรกิจ รวมถึงการแบ่งกลุ่มลูกค้าตามลักษณะเฉพาะของพวกเขาเพื่อส่งเสริมพฤติกรรมการซื้อของลูกค้า

2.4 บทสรุป

จากการศึกษาวรรณกรรมที่เกี่ยวข้อง พบว่าการวิจัยเหล่านี้แสดงให้เห็นถึงความสำคัญของการเลือกอัลกอริทึมที่เหมาะสมในกระบวนการทำเหมืองข้อมูลความสัมพันธ์ มีความสำคัญเนื่องจากการใช้วิธีที่มีประสิทธิภาพจะช่วยเพิ่มความแม่นยำและประสิทธิภาพในการค้นหาไอเท็มเซตที่เกิดขึ้นบ่อย การเปรียบเทียบอัลกอริทึมต่าง ๆ จะพิจารณาจากปัจจัยหลายประการ เช่น ความเร็วในการประมวลผล การใช้ทรัพยากร และความเหมาะสมของข้อมูลที่น่ามาวิเคราะห์ การ

เลือกใช้อัลกอริทึมที่เหมาะสมจะช่วยให้กระบวนการทำเหมืองข้อมูลมีประสิทธิภาพสูงสุด นอกจากนี้ การเลือกใช้อัลกอริทึมที่เหมาะสมแล้ว การตั้งค่าพารามิเตอร์ที่สอดคล้องกับลักษณะของข้อมูลก็มีความสำคัญต่อประสิทธิภาพของกระบวนการทำเหมืองข้อมูล โดยพารามิเตอร์ที่ต้องพิจารณามีหลายปัจจัย เช่น ค่าความถี่ต่ำสุดที่ใช้กำหนดว่าไอเท็มเซตใดควรถูกนำมาพิจารณา ค่าความเชื่อมั่นของกฎความสัมพันธ์ที่สะท้อนถึงความน่าเชื่อถือของกฎที่ค้นพบ และค่าตัวชี้วัดอื่น ๆ ที่ช่วยให้สามารถระบุความสัมพันธ์ที่มีความหมายได้มากกว่าการเกิดขึ้นโดยบังเอิญ นอกจากนี้ ควรมีการกำหนดจำนวนไอเท็มสูงสุดที่สามารถรวมกันในกฎความสัมพันธ์ เพื่อป้องกันความซับซ้อนที่มากเกินไปและลดการใช้ทรัพยากรที่ไม่จำเป็น