

บทที่ 2

แนวคิด ทฤษฎี และวรรณกรรมที่เกี่ยวข้อง

บทนี้เป็นคำแนะนำเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้องกับการวิเคราะห์พฤติกรรมผู้บริโภคของลูกค้า โดยใช้เทคนิค RFM Analysis และ Machine Learning เพื่อจัดกลุ่มและวิเคราะห์การซื้อซ้ำของลูกค้า โดยนำเสนอข้อมูลการวิเคราะห์ผ่านทางเว็บไซต์ ซึ่งมีข้อมูลและทฤษฎีที่เกี่ยวข้องต่างๆ ดังนี้

2.1 แนวคิด

- 2.1.1 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)
- 2.1.2 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytic)
- 2.1.3 แนวคิดเกี่ยวกับการทำความสะอาดข้อมูล (Data Cleaning)
- 2.1.4 แนวคิดเกี่ยวกับกฎความสัมพันธ์ (Association Rules)
- 2.1.5 แนวคิดการจับกลุ่มข้อมูลด้วย RFM (Recency, Frequency, Monetary)
- 2.1.6 แนวคิดการจับกลุ่มข้อมูลด้วย K-Means
- 2.1.7 แนวคิดเกี่ยวกับกฎความสัมพันธ์ ด้วย FP-Growth
- 2.1.8 แนวคิดแนวคิดการจับกลุ่มข้อมูลด้วย Decision Tree

2.2 ทฤษฎี

- 2.2.1 ทฤษฎีเกี่ยวกับเหมืองข้อมูล (Data mining)
- 2.2.2 ทฤษฎีเกี่ยวกับการสร้างเว็บไซต์
- 2.2.3 ทฤษฎีเกี่ยวกับ Machine Learning
- 2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่ง HTML
- 2.2.5 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS

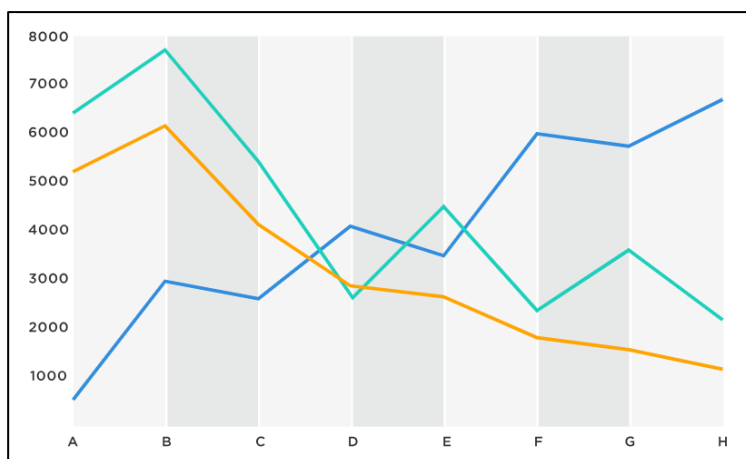
2.3 วรรณกรรมที่เกี่ยวข้อง

2.4 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization) การแสดงข้อมูลที่ผ่านการวิเคราะห์และประมวลผลในรูปแบบที่เข้าใจง่ายและสื่อสารได้ชัดเจน เช่น แผนภูมิ แผนที่ อินโฟกราฟิก หรือรูปภาพ การนำเสนอเหล่านี้ช่วยให้ผู้ใช้งานสามารถมองเห็นแนวโน้ม รูปแบบ และข้อมูลเชิงลึกได้อย่างรวดเร็ว องค์กรจึงสามารถนำข้อมูลไปใช้ในการตัดสินใจที่แม่นยำและมีประสิทธิภาพมากขึ้น เนื่องจากทุกองค์กรมีข้อมูลสำคัญและมีความซับซ้อน ซึ่งข้อมูลปริมาณมากอาจทำให้การวิเคราะห์เกิดความผิดพลาดได้ ด้วยเหตุผลนี้ทำให้องค์กรต้องมีการทำ Data Visualization เพื่อช่วยให้ผู้ใช้งานมองเห็นภาพรวมได้ชัดเจน มีส่วนช่วยในการตัดสินใจทั้งในด้านการบริหารและการวางแผนธุรกิจ โดยการนำ Data Visualization นั้น มีรูปแบบการแสดงผลหลายประเภทด้วยกัน ซึ่งมี 5 รูปแบบที่เป็นที่นิยมนำมาใช้ในการทำงาน (cheewin.j, 2568)

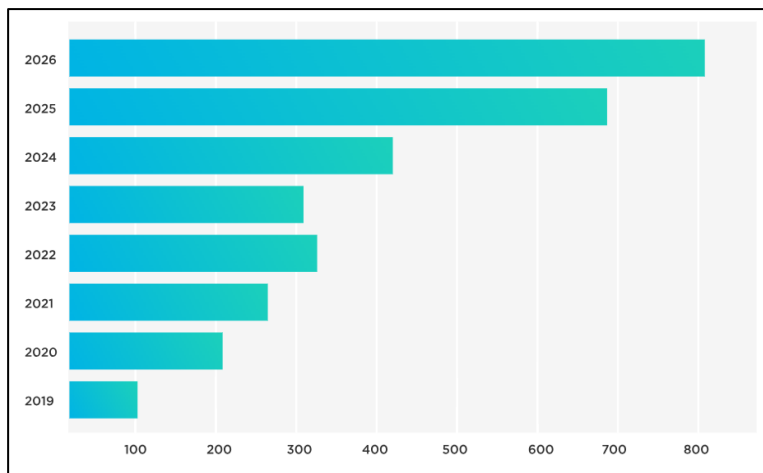
- 1) Line Chart คือแผนภูมิที่จะใช้เพื่อแสดงความสัมพันธ์ของข้อมูลที่เปลี่ยนแปลงตามช่วงเวลา เรียงตามแกน X และแกน Y โดยใช้จุดแทนข้อมูล และลากเส้นเชื่อมต่อดูจุดแต่ละจุดเพื่อให้เห็นการเปลี่ยนแปลง



ภาพที่ 2.1 แผนภูมิแบบเส้น

(ที่มา: <https://www.jaspersoft.com/articles/what-is-a-line-chart>)

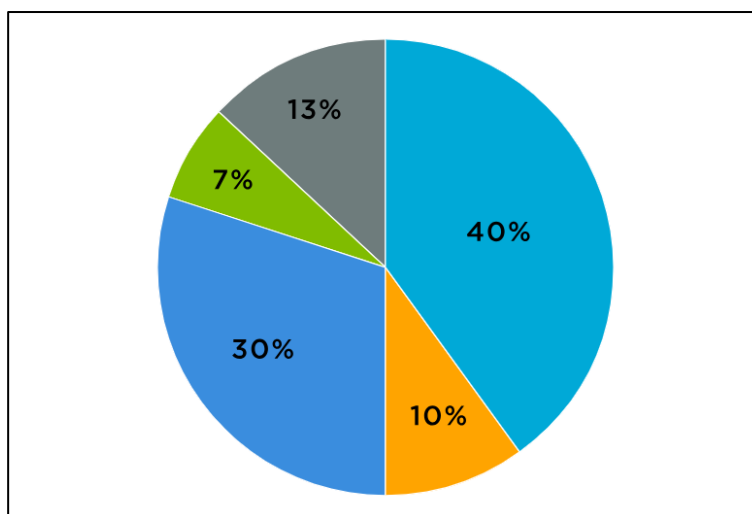
- 2) Bar Chart คือแผนภูมิที่จะใช้เพื่อแสดงเหมาะสำหรับการเปรียบเทียบข้อมูลระหว่างหมวดหมู่ต่างๆ โดยใช้แท่งสีเหลี่ยมในการแสดงข้อมูลแทนจุด และใช้ความสูงต่ำเพื่อแสดงให้เห็นความแตกต่าง



ภาพที่ 2.2 แผนภูมิแบบแท่ง

(ที่มา: <https://www.jaspersoft.com/articles/what-is-a-bar-chart>)

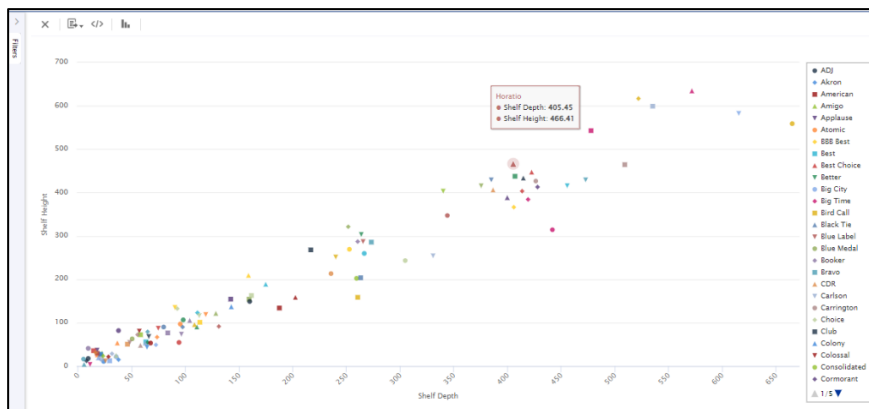
- 3) Pie Chart คือแผนภูมิที่อยู่ในรูปแบบวงกลม โดยจะมีการแบ่งวงกลมออกเป็นส่วนๆ ตามจำนวนข้อมูลที่ต้องการเปรียบเทียบ และใช้ตัวเลขมาแสดงปริมาณข้อมูลแต่ละส่วนเมื่อเทียบกับปริมาณข้อมูลทั้งหมด



ภาพที่ 2.3 แผนภูมิแบบวงกลม

(ที่มา: <https://www.jaspersoft.com/articles/what-is-a-pie-chart>)

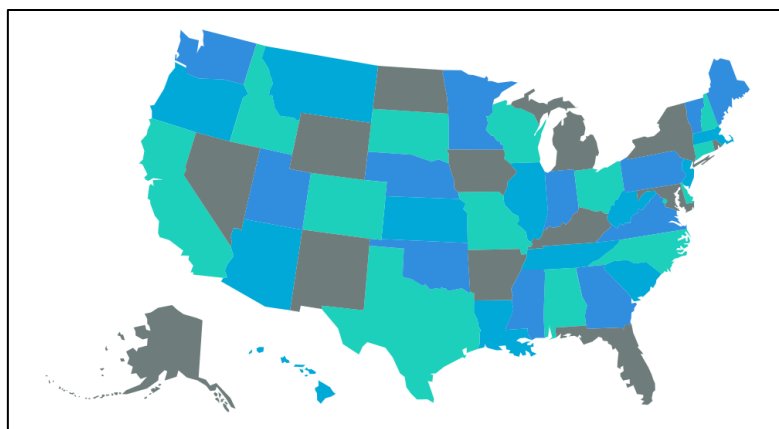
- 4) Scatter Plot คือกราฟที่แสดงความสัมพันธ์ระหว่างชุดข้อมูลสองชุด โดยใช้จุดแทนค่าข้อมูลบนแกน X และแกน Y จุดที่กระจายตัวบนกราฟช่วยให้เห็นรูปแบบหรือแนวโน้มความสัมพันธ์ระหว่างตัวแปร



ภาพที่ 2.4 กราฟแบบกระจาย

(ที่มา: <https://www.jaspersoft.com/articles/what-is-a-scatter-chart>)

- 5) Map คือกราฟแสดงข้อมูลที่เกี่ยวข้องกับตำแหน่งบนพื้นที่ โดยใช้แผนที่เป็นเครื่องมือในการแสดงผล และปรับแต่งพื้นที่เพื่อแสดงค่าหรือความแตกต่าง



ภาพที่ 2.5 กราฟแสดงแผนที่

(ที่มา: <https://www.jaspersoft.com/articles/what-is-a-map-chart>)

2.1.2 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytic) กระบวนการตรวจสอบแปลง และตีความข้อมูลที่มีอยู่ เพื่อนำมาใช้ในการค้นหาความรู้เชิงลึก (Insight) และสนับสนุนการตัดสินใจ โดยอาศัยเครื่องมือและเทคนิคทางสถิติ คณิตศาสตร์ และเทคโนโลยีสารสนเทศ เพื่อให้ข้อมูลดิบ (Raw Data) มีความหมายมากขึ้น และสามารถนำไปใช้ในการแก้ไขปัญหา การคาดการณ์แนวโน้ม รวมถึงการกำหนดกลยุทธ์ทางธุรกิจได้อย่างมีประสิทธิภาพ

2.1.2.1 การวิเคราะห์ข้อมูล (Data Analytics) คือการนำข้อมูลที่มีอยู่มาเข้าสู่กระบวนการวิเคราะห์ เพื่อนำข้อมูลไปใช้ประโยชน์ในด้านต่าง ๆ ที่ครอบคลุมในทุกอุตสาหกรรม ทั้งภาครัฐและภาคเอกชน ซึ่งจะต้องอาศัยเทคโนโลยีหรือซอฟต์แวร์เฉพาะด้านที่

เข้ามาช่วยให้การประมวลผลและวิเคราะห์ข้อมูลเป็นไปอย่างแม่นยำและนำไปใช้งานได้อย่างมีประสิทธิภาพ โดยสามารถแบ่งประเภทของการวิเคราะห์ข้อมูล ได้ 4 รูปแบบ ดังนี้

1) Descriptive Analytics คือการวิเคราะห์ข้อมูลแบบพื้นฐาน เพื่อแสดงผลที่เกิดขึ้น หรือกำลังจะเกิดขึ้น จากข้อมูลในอดีต ในลักษณะที่เข้าใจง่ายสามารถสร้างขึ้นได้ด้วยตนเอง เช่น รายงาน แผนภูมิ กราฟ ตาราง เป็นต้น ซึ่งจะช่วยให้เข้าใจการเปลี่ยนแปลงที่เกิดขึ้นกับองค์กรได้ดียิ่งขึ้น เช่น การเปลี่ยนแปลงของราคาสินค้ารายปี การเติบโตของยอดขายรายเดือน การเปรียบเทียบยอดขายในแต่ละสาขาหรือแต่ละช่องทาง การเปรียบเทียบจำนวนผู้ใช้งานเว็บไซต์ในแต่ละช่วงเวลา เป็นต้น ซึ่ง Descriptive Analytics คือการอธิบายสิ่งที่เกิดขึ้นตามวัตถุประสงค์และช่วงเวลาที่กำหนดนั่นเอง

2) Diagnostic Analytics คือการวิเคราะห์เชิงวินิจฉัย ซึ่งเป็นรูปแบบหนึ่งของการวิเคราะห์ขั้นสูงแบบเจาะลึก โดยการวิเคราะห์ข้อมูลที่ได้จากการทำ Descriptive Analytics เพื่อหาคำตอบว่าทำไมจึงเกิดสิ่งนั้น ๆ หรืออธิบายปัจจัยและตัวแปรที่เป็นสาเหตุของการเกิดสิ่งนั้น ๆ ขึ้น ซึ่งจะต้องอาศัยเทคนิคต่าง ๆ เข้ามาช่วย เช่น การทำ Data discovery หรือ Data mining เป็นต้น ตัวอย่างการวิเคราะห์ข้อมูลเชิงวินิจฉัย เช่น อธิบายสาเหตุที่ทำให้ยอดขายเพิ่มขึ้น จำนวน ผู้ใช้ที่เข้าชมเว็บไซต์เพิ่มขึ้น จำนวนลูกค้าที่เข้าใช้บริการที่หน้าร้านลดลง โปรโมชันที่ไม่ค่อยได้รับความนิยม เนื้อหาโฆษณาที่ได้ CTR% มากกว่าเนื้อหาอื่น ๆ หรือวิเคราะห์การทำงานของเครื่องคอมพิวเตอร์เพื่อตรวจจับความผิดปกติ เป็นต้น ข้อมูลเหล่านี้จะทำให้องค์กรรู้ถึงความต้องการของตลาด เข้าใจพฤติกรรมของลูกค้า รู้สาเหตุของปัญหาด้านเทคโนโลยี รวมถึงสามารถปรับปรุงวัฒนธรรมองค์กร เพื่อการทำงานที่ดีขึ้น

3) Predictive Analytics คือการวิเคราะห์ข้อมูลทั้งข้อมูลในอดีตและปัจจุบันออกมาในเชิงคาดการณ์ ทำนาย หรือการพยากรณ์ เพื่อหาแนวโน้มที่จะเกิดสิ่งต่าง ๆ ขึ้นตามวัตถุประสงค์ที่กำหนด โดยการสร้างแบบจำลองทางสถิติ บวกกับการนำเทคโนโลยีปัญญาประดิษฐ์มาใช้ ซึ่งสามารถสร้างประโยชน์ได้มากมายในหลายแง่มุม เช่น การคาดการณ์ความเสี่ยงและโอกาส ยอดขาย ภัยไซเบอร์ สภาพอากาศ การลงทุน หุ่น หรือผลการเลือกตั้ง เป็นต้น อย่างไรก็ตามการทำ Predictive Analytics ที่ถูกต้องและแม่นยำนั้น ขึ้นอยู่กับคุณภาพของข้อมูล ซึ่งเป็นสิ่งที่องค์กรควรให้ความสำคัญเป็นอันดับแรก โดยการเตรียมข้อมูลให้มีคุณภาพที่ดีและเหมาะสม ก่อนนำไปใช้วิเคราะห์เพื่อให้เกิดผลลัพธ์ที่มีประสิทธิภาพ ลดข้อผิดพลาด และเกิดประโยชน์อย่างแท้จริง

4) Prescriptive Analytics คือการวิเคราะห์แบบให้คำแนะนำ ซึ่งเป็นกระบวนการวิเคราะห์ที่มีความซับซ้อนมากที่สุด ต่อเนื่องจากการทำ Predictive Analytics กล่าวคือ เมื่อได้

ข้อมูลแนวโน้มที่จะเกิดบางสิ่งขึ้นแล้ว การทำ Prescriptive Analytics จะช่วยแนะนำแนวทางการดำเนินการในขั้นตอนต่อไปที่เหมาะสมที่สุด และวิเคราะห์ไปถึงผลที่จะเกิดขึ้นถ้าหากเลือกปฏิบัติตามแนวทางนั้น ๆ หรือแม้แต่แนะนำแนวทางในการรับมือและแก้ไขปัญหา การวิเคราะห์แบบให้คำแนะนำจึงถือเป็นเครื่องมือที่สำคัญอย่างมากสำหรับการตัดสินใจที่ซับซ้อนด้วยข้อมูล การวิเคราะห์แบบให้คำแนะนำคือทำงานร่วมกันระหว่าง Big data อัลกอริทึมของ Machine learning และเทคโนโลยี AI เพื่อช่วยในการวิเคราะห์ข้อมูลจำนวนมากที่มีความซับซ้อนเกินกว่าที่มนุษย์จะทำได้ ซึ่งการทำ Prescriptive Analytics ยังช่วยเพิ่มประสิทธิภาพในการตัดสินใจด้านต่าง ๆ ไม่ว่าจะเป็น ด้านการลงทุน ด้านการตลาด ด้านการตรวจจับการฉ้อโกง (Fraud Detection) ในอุตสาหกรรมธนาคาร ด้านการพัฒนาและปรับปรุงผลิตภัณฑ์ เป็นต้น ซึ่งนับว่าการวิเคราะห์ที่ทรงพลังเป็นอย่างมาก และเป็นแรงขับเคลื่อนสำคัญที่จะช่วยให้องค์กรสามารถบรรลุเป้าหมายทางธุรกิจได้อย่างรวดเร็วและมีประสิทธิภาพ

2.1.2.2 วิวัฒนาการของการวิเคราะห์ข้อมูล (Data Analytics Evolution)

การวิเคราะห์ข้อมูลแท้จริงแล้วเป็นเรื่องที่อยู่กับมนุษย์มาเป็นเวลานาน ซึ่งมีพื้นฐานมาจากการใช้ตัวเลขหรือสถิติ เพียงแต่ในสมัยก่อนการที่รัฐบาลหรือองค์กรจะบันทึกข้อมูล แล้วนำมาวิเคราะห์นั้น จะต้องใช้เวลานานและใช้ความพยายามอย่างสูง ตัวอย่างเช่น เมื่อย้อนกลับไปสมัยที่ยังไม่มีคอมพิวเตอร์ สำนักสำรวจสำมะโนประชากรของสหรัฐฯ ใช้เวลามากกว่า 7 ปี ในการประมวลผลข้อมูลให้เสร็จสมบูรณ์ โดยมีการแบ่งยุคสมัยดังนี้

1) แต่เมื่อยุคสมัยเปลี่ยนไป โลกเริ่มมีการพัฒนาความก้าวหน้าด้านเทคโนโลยี จนกระทั่งช่วงยุค 1940 – 1950 เป็นยุคที่องค์กรภาครัฐเริ่มใช้คอมพิวเตอร์ และเริ่มมีการวิเคราะห์ข้อมูลเกิดขึ้น รวมทั้งการวิเคราะห์ข้อมูลเริ่มถูกนำไปใช้ในองค์กรด้านงานวิจัยมากขึ้น โดยค่อย ๆ เริ่มจากการรวบรวมข้อมูลแบบออฟไลน์และวิเคราะห์ด้วยตนเอง

2) ต่อเนื่องสู่ช่วงปลายยุค 1960 Analytics เริ่มได้รับความสนใจมากขึ้นเมื่อคอมพิวเตอร์กลายเป็นระบบสนับสนุนการตัดสินใจในด้านต่าง ๆ ทำให้มีการพัฒนาเทคโนโลยีเกิดขึ้น เช่น องค์กรด้านไอทีที่มีชื่อเสียงอย่าง IBM ได้คิดค้นเทคโนโลยีการเก็บข้อมูล “Disk storage” เป็นจุดเริ่มต้นของการเข้าถึงข้อมูลและแบ่งปันข้อมูลระหว่างคอมพิวเตอร์ได้ รวมทั้งมีการพัฒนาระบบจัดการฐานข้อมูล (DBMS) อีกด้วย เป็นยุคแรกเริ่มของการนำเทคโนโลยีมาช่วยในการวิเคราะห์ข้อมูลอย่างเห็นได้ชัด

3) ต่อมาในยุค 1970 – 1980 ด้วยการพัฒนาของยุคสมัยและเทคโนโลยีที่รวดเร็ว บริษัทด้านเทคโนโลยี และสตาร์ทอัพหลายแห่ง ได้พัฒนานวัตกรรมที่เกี่ยวข้องกับ

ข้อมูลขึ้นมามากมาย ไม่ว่าจะเป็น Relational database, Data warehouse, Decision support systems (DSS), SQL รวมทั้ง ETL เป็นต้น

4) จนเข้าสู่ยุค 1990 การค้นหาออนไลน์ได้รับความนิยมมากขึ้น ทั้งในภาคธุรกิจและประชาชนทั่วไป โดยบริษัทยักษ์ใหญ่อย่าง Google, Amazon และ eBay เนื่องจากข้อมูลที่เพิ่มมากขึ้นเป็นจุดเริ่มต้นให้โลกเกิดคำว่า ข้อมูลมหาดศาล หรือ Big data ขึ้นมา และทำให้เกิดนวัตกรรมต่าง ๆ ตามมา ไม่ว่าจะเป็น Database marketing, Data mart, OLAP, Data mining และ Data visualization ซึ่งเข้ามาช่วยในวิเคราะห์ข้อมูลแนวใหม่ทางธุรกิจ ซึ่งทำให้ได้ข้อสรุปและคำแนะนำที่ดีขึ้นกว่าในยุคสมัยก่อน ๆ อย่างมาก ช่วยให้การทำธุรกิจมีประสิทธิภาพยิ่งขึ้น แต่ในขณะเดียวกัน เมื่อปริมาณข้อมูลมีการเพิ่มมากขึ้นอย่างรวดเร็ว หลายองค์กรจึงเริ่มพบปัญหาในการจัดเก็บและการประมวลผลข้อมูล

5) จนมาถึงยุคปี 2000 จนถึงปัจจุบัน ซึ่งเป็นยุคของการใช้สมาร์ทโฟน, แอปพลิเคชัน หรืออุปกรณ์ที่เชื่อมต่อต่าง ๆ ซึ่งเป็นยุคเกิดการแข่งขันทางธุรกิจที่สูงมาก เพื่อเอาชนะใจลูกค้า ทำให้เกิดแพลตฟอร์มที่เกี่ยวกับการประมวลผลการวิเคราะห์ออนไลน์เกิดขึ้น และเกิดข้อมูลที่ไม่มีโครงสร้างขึ้นมามากมาย ถือเป็นจุดเริ่มต้นของฐานข้อมูลรูปแบบใหม่ที่เรียกว่า NoSQL โดยในยุคนี้บริษัทไอทีทั่วโลก ได้มีการพัฒนาเทคโนโลยีใหม่ ๆ ออกมาเพื่อตอบสนองความต้องการของธุรกิจ ไม่ว่าจะเป็น NoSQL, Cloud storage, Cloud computing, AI และ ML เป็นต้น ซึ่งเทคโนโลยีเหล่านี้ถูกพัฒนามาเพื่อลดข้อจำกัดในการจัดเก็บข้อมูล ช่วยให้สามารถประมวลผลข้อมูลได้เร็วและแม่นยำขึ้น สามารถนำข้อมูลไปวิเคราะห์ขั้นสูงเพื่อหาข้อมูลเชิงลึกที่เป็นประโยชน์ต่อธุรกิจ ทำให้ความสามารถของการทำ Data analytics นั้น ก็มีการยกระดับขึ้นไปพร้อมกับการพัฒนาของเทคโนโลยีต่อเนื่องมาจนถึงปัจจุบัน (blendata, 2567)

2.1.3 แนวคิดเกี่ยวกับการทำความสะอาดข้อมูล (Data Cleaning)

การทำความสะอาดข้อมูล (Data Cleansing หรือ Data Cleaning) คือ กระบวนการตรวจสอบ แก้ไข หรือลบข้อมูลที่ไม่สมบูรณ์ ซ้ำซ้อน ผิดพลาด หรืออยู่ในรูปแบบที่ไม่เหมาะสม เพื่อให้ข้อมูลมีความถูกต้อง เป็นระเบียบ และสามารถนำไปใช้วิเคราะห์ได้อย่างมีประสิทธิภาพ (ditto, 2567)

2.1.3.1 ความสำคัญ

1) ช่วยให้ผลลัพธ์ของการวิเคราะห์แม่นยำขึ้น ลดความคลาดเคลื่อน (errors) จากข้อมูลที่ผิดหรือขาด

2) ทำให้ได้รายงาน (Insight / Report) ที่สามารถนำไปใช้สนับสนุนการตัดสินใจได้อย่างรวดเร็วและเชื่อถือได้

3) ป้องกันผลกระทบทางลบจากการตัดสินใจที่อาศัยข้อมูลที่ไม่สมบูรณ์

2.1.3.2 ลักษณะของข้อมูลที่ควรผ่านการทำความสะอาด

1) ข้อมูลที่อยู่ในหลายรูปแบบ (หลุดหลายไฟล์หรือหลายแหล่ง) ซึ่งอาจมีนามสกุลไฟล์ต่างกัน เช่น .pdf .doc .xls ฯลฯ

2) ข้อมูลที่ไม่ได้อยู่ในรูปแบบที่ต้องการ เช่น เป็นรูปภาพ (.jpg, .png) แต่ต้องการไฟล์ที่อ่านได้เป็นข้อความหรือไฟล์ structured (.csv, .json)

3) ข้อมูลที่ไม่ถูกต้อง เช่น มีการกรอกผิดพลาด ข้อมูลที่เกินความเป็นจริง หรือมีความผิดพลาดจากการป้อนข้อมูลคน

2.1.3.3 ขั้นตอนการทำ Data Cleansing

1) ลบข้อมูลที่ซ้ำซ้อน (Duplicate Removal)

2) ลบข้อมูลที่ไม่เกี่ยวข้องกับวัตถุประสงค์การวิเคราะห์ (Irrelevant Data Removal)

3) แทนที่ข้อมูลที่หายไปหรือผิดพลาด (Data Imputation / Replacement)

4) บำรุงรักษาข้อมูลให้คงอยู่ในสภาพดีอยู่เสมอ (Data Maintenance)

5) ตรวจสอบความถูกต้องของข้อมูลหลังทำความสะอาด (Validation)

2.1.4 เทคนิคเกี่ยวกับกฎความสัมพันธ์ (Association Rules)

กฎความสัมพันธ์ (Association Rules) หมายถึง การทำเหมืองข้อมูล เพื่อหาความสัมพันธ์ของเหมืองข้อมูลมักใช้ในธุรกิจการค้าปลีก (retailing business) เช่น ร้านค้าสะดวกซื้อ หรือ ซูเปอร์มาร์เก็ต เป็นการวิเคราะห์ตะกร้าตลาด (Market basket analysis) เพื่อศึกษาพฤติกรรมการซื้อสินค้าของลูกค้า และหาความสัมพันธ์ของสินค้าที่ลูกค้าซื้อ เพื่อนำผลลัพธ์ที่ได้จากความสัมพันธ์ มาใช้ในการจัดวางสินค้าบนชั้น เพื่อให้ลูกค้าสามารถหยิบซื้อสินค้าด้วยกันได้สะดวก หรือนำผลลัพธ์ที่ได้มาใช้ในการส่งเสริมการขายสินค้าหรือจัดทำแคตตาล็อกสินค้า ความสัมพันธ์ของสินค้าที่ลูกค้าซื้อจะแสดงในรูปแบบของกฎความสัมพันธ์ (Association Rule) ดังนี้

$A \rightarrow B[\text{Support, Confident}]$ โดยที่ A, B แทนรายการสินค้า

สูตรการหาความสัมพันธ์

เช่น Milk → Eggs [Support = 25% ,Confident=33.34%] หมายความว่า 25% ของทรานแซคชันทั้งหมด ลูกค้าจะซื้อนม (Milk) และซื้อไข่ (Eggs) พร้อมกัน และ 33.34% ของลูกค้าที่ซื้อนมแล้วจะซื้อไข่ด้วย

กฎความสัมพันธ์ที่สนใจหรือกฎความสัมพันธ์ที่แข็งแกร่ง (Strong Association Rules) คือ กฎความสัมพันธ์ที่มีค่าสนับสนุน (support) และค่าความเชื่อมั่น (confidence) ผ่านเกณฑ์ขั้นต่ำ (Minimum Threshold) ที่ผู้วิเคราะห์ข้อกำหนดขึ้นมา

กฎความสัมพันธ์ (Association Rule) คือการทำเหมืองข้อมูลโดยใช้กฎความสัมพันธ์ เพื่อหาความสัมพันธ์ของเหมืองข้อมูลมักใช้ในธุรกิจการค้าปลีก (retailing business) เช่น ร้านค้าสะดวกซื้อ หรือ ซูเปอร์มาร์เก็ต เป็นการวิเคราะห์ตะกร้าตลาด (Market basket analysis) เพื่อศึกษาพฤติกรรมการซื้อสินค้าของลูกค้า และหาความสัมพันธ์ของสินค้าที่ลูกค้าซื้อ เพื่อนำผลลัพธ์ที่ได้จากความสัมพันธ์ (Mindphp, 2562)

2.1.5 เทคนิคการจับกลุ่มข้อมูลด้วย RFM (Recency, Frequency, Monetary)

RFM Analysis เป็นการแบ่งกลุ่มลูกค้าโดยการนำ Transactions Data / Sales Data ไม่ว่าจะจาก Shopee, Lazada, Point of Sales (POS) มาวิเคราะห์ เพื่อแบ่ง Segmentation ของลูกค้าออกเป็นกลุ่มๆ ผ่าน 3 criteria หลักๆ ได้แก่

- R – Recency ลูกค้ายี่สิบล่าสุดเมื่อไหร่
- F – Frequency ความถี่ในการซื้อ
- M – Monetary ปริมาณการซื้อ หรือ Basket Size

และหลังจากนั้นจะกำหนดกลยุทธ์สำหรับลูกค้าแต่ละกลุ่มแตกต่างกัน

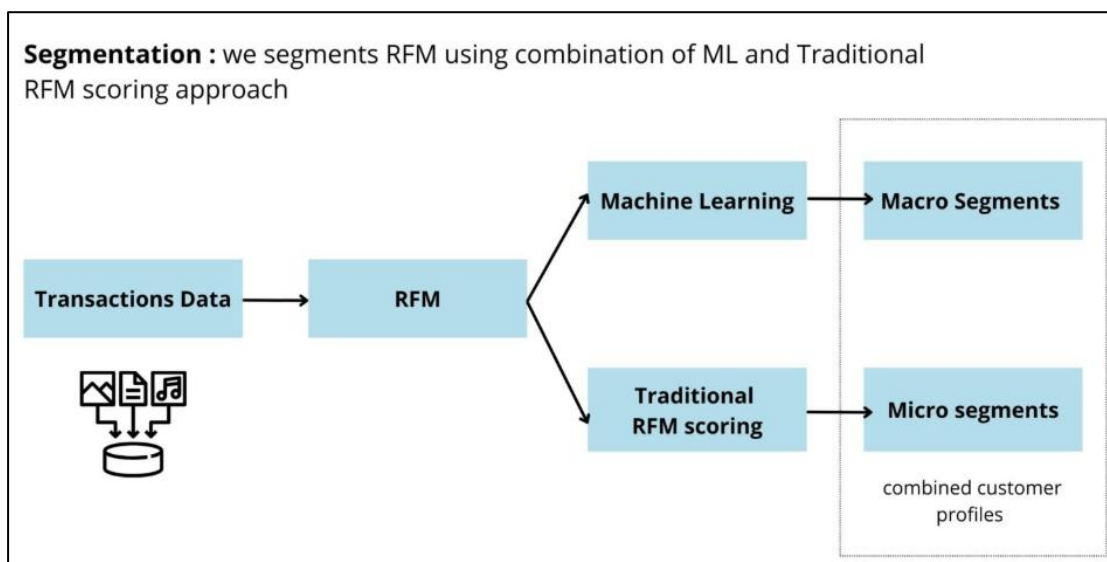
Segmentation หมายถึง การแบ่งลูกค้าออกเป็นกลุ่มๆ ตามความชอบพฤติกรรม ที่แตกต่างกัน ซึ่งสามารถแบ่งได้จากหลากหลาย Criteria เช่น เพศ อายุ Basket Size พฤติกรรมการใช้งานเว็บไซต์ ทักษะคิด เป็นต้น ยกตัวอย่างลูกค้ากลุ่มแรก เป็นกลุ่มลูกค้าที่ใช้จ่ายต่อครั้งสูง ใช้ของ Luxury เพื่อแสดงฐานะของตัวเอง ลูกค้ากลุ่มที่สอง เป็นลูกค้ากลุ่มที่ใช้จ่ายต่อครั้งปานกลาง เลือกซื้อของจากฟังก์ชันการใช้งานเท่านั้น ลูกค้ากลุ่มที่สาม เป็นลูกค้าที่ใช้จ่ายต่อครั้งต่ำ เลือกซื้อของเฉพาะช่วงโปรโมชั่น มีความอ่อนไหวต่อราคา

2.1.5.1 ขั้นตอนการวางกลยุทธ์ RFM Analysis เพื่อแบ่งกลุ่มลูกค้า

1) กำหนดเป้าหมาย และโจทย์ทางธุรกิจที่ต้องการทราบ ก่อนอื่น ธุรกิจต้องกำหนดเป้าหมายเป้าหมายของธุรกิจ เช่น การเพิ่มยอดขาย หรือ การรักษาลูกค้าเก่า โดยข้อมูลที่ธุรกิจมีจะเข้ามาช่วยตอบคำถามทางธุรกิจที่เกิดขึ้น ไม่ว่าจะเป็น

- สัดส่วนของลูกค้าขาจรและลูกค้าประจำเป็นอย่างไรบ้าง แต่ละกลุ่มสร้างรายได้เป็นสัดส่วนเท่าไร
- จากข้อมูล Transaction Data ทั้งหมด สามารถแบ่ง Segmentation ได้กี่กลุ่ม แต่ละกลุ่มมีโปรไฟล์แตกต่างกันยังไง (โปรไฟล์ เช่น ความถี่ในการซื้อ Customer Lifetime Value ปริมาณการซื้อ เป็นต้น)
- ควรจะวางกลยุทธ์การสื่อสารกับลูกค้าแต่ละกลุ่มยังไงให้เหมาะสม

2) เตรียมข้อมูล และเลือกวิธีการวิเคราะห์ข้อมูลให้ตอบโจทย์ เมื่อกำหนดเป้าหมายของธุรกิจแล้ว จะต้องเตรียมความพร้อมของข้อมูล Transaction Data ที่ธุรกิจมีจากช่องทางต่างๆ เช่น Point of sales , Marketplace เพื่อเตรียมการนำข้อมูล ดังกล่าวมาวิเคราะห์หา Segmentation ผ่านวิธีการ RFM analysis



ภาพที่ 2.6 แสดงภาพรวมของการนำข้อมูล Transactions Data มาสร้าง Segmentation ผ่านการใช้ RFM analysis

(ที่มา: <https://predictive.co.th/blog/rfm-analysis/>)

จากภาพด้านบน จะเป็นภาพรวมของการนำข้อมูล Transactions Data มาสร้าง Segmentation ผ่านการใช้ RFM analysis ด้วย 2 วิธีได้แก่

- Machine Learning เพื่อหา Macro segment
- Traditional RFM scoring เพื่อหา Micro segment

3) เริ่มต้นวิเคราะห์ข้อมูล เมื่อเราเตรียมความพร้อมของข้อมูลได้อย่างครบถ้วนแล้ว ในขั้นตอนนี้เราจะเริ่มต้นวิเคราะห์ข้อมูลพฤติกรรมของลูกค้า ผ่าน Transaction Data เพื่อแบ่งกลุ่มลูกค้า (Segmentation) ตามพฤติกรรมที่แตกต่างกัน โดยการนำ RFM analysis นั้นสามารถวิเคราะห์ข้อมูลได้ 2 วิธีหลักๆ ได้แก่

- Machine Learning เพื่อหา Macro segment ใช้หลักสถิติ clustering algorithms = K-Means เพื่อแบ่งกลุ่มลูกค้า เป็นลูกค้าขาจร และ ลูกค้าขาประจำ เพื่อดูสัดส่วนของลูกค้าแต่ละประเภท และสร้างโปรไฟล์ตาม Criteria ต่างๆ เช่น จำนวนลูกค้า , ยอดขายของลูกค้าแต่ละกลุ่ม (กลุ่มขาจร และ กลุ่มขาประจำ) , จำนวนวันนับจากการซื้อครั้งสุดท้าย เมื่อเราสร้างโปรไฟล์ของลูกค้าแล้ว เราสามารถนำข้อมูลมาวิเคราะห์ต่อได้ดังนี้

- หาสัดส่วนของลูกค้าขาจร และ ลูกค้าขาประจำ
- ลูกค้าเก่าสร้างรายได้เป็นที่เปอร์เซ็นต์ และ ลูกค้าใหม่สร้างรายได้เป็นที่เปอร์เซ็นต์
- หาเทรนด์ ของจำนวนลูกค้าแต่ละกลุ่มเพื่อเข้าใจเทรนด์การเพิ่มขึ้น – ลดลงของลูกค้าแต่ละกลุ่ม เพื่อเข้าใจภาพรวม และ วางกลยุทธ์แก้ไขได้ตรงจุด

- Traditional RFM scoring เพื่อหา Micro segment แบ่งกลุ่มลูกค้าออกได้เป็น Micro Segmentation ด้วยวิธีการสร้าง Rule Based Segmentation หรือการกำหนดคุณสมบัติของลูกค้าแต่ละกลุ่ม ซึ่งในขั้นตอนจะเป็นขั้นตอนที่ทำร่วมกับแบรนต์เพื่อกำหนดคุณสมบัติของลูกค้าแต่ละกลุ่มร่วมกัน

Segment	คำอธิบาย
Champions	เพิ่งซื้อไปไม่นานมานี้ ซื้อบ่อย และ ใช้จ่ายเยอะ
Loyal customers	ใช้จ่ายเยอะ และซื้อบ่อย
Need Attention	เหนือกว่าค่าเฉลี่ยทั้งหมด
Small basket size	ใช้จ่ายไม่เยอะ แต่ซื้อบ่อย และ เพิ่งซื้อไปไม่นาน
Potential Loyalists	เพิ่งซื้อไปไม่นาน ใช้จ่ายปานกลางถึงสูง และซื้อมากกว่า 1 ครั้ง
New customers	เพิ่งซื้อไปไม่นาน และซื้อสินค้าไม่บ่อย
Promising	เพิ่งซื้อไปไม่นาน และใช้จ่ายน้อย
Long time big buy	ใช้จ่ายเยอะ นานๆ ครั้ง
At risk	ใช้จ่ายเยอะ เมื่อนานมากๆ แล้ว
Hibernating	ใช้จ่ายน้อย ซื้อไม่บ่อย และ ซื้อครั้งล่าสุดเมื่อนานมาแล้ว

ภาพที่ 2.7 แสดงตารางแบ่งกลุ่ม segment ของลูกค้า
(ที่มา: <https://predictive.co.th/blog/rfm-analysis/>)

หลังจากที่เราแบ่งลูกค้าออกเป็น 10 กลุ่มตาม Criteria ของ RFM analysis อันได้แก่

- R – Recency ซึ่งล่าสุดเมื่อไหร่
- F – Frequency ความถี่ในการซื้อ
- M – Monetary ปริมาณการซื้อ หรือ Basket Size

แล้วเราสามารถสร้างโปรไฟล์ของแต่ละกลุ่มลูกค้า เพื่อให้สะดวกต่อการเข้าใจพฤติกรรม และวิเคราะห์สัดส่วนของลูกค้าแต่ละกลุ่มได้ (Nalyn, 2565)

2.1.6 เทคนิคการจับกลุ่มข้อมูลด้วย K-Means

เทคนิค K-Means Clustering เป็นหนึ่งในวิธีการทำเหมืองข้อมูล (Data Mining) อยู่ในกลุ่มของ Unsupervised Learning ใช้สำหรับแบ่งข้อมูลออกเป็นกลุ่มย่อย (Clusters) โดยอาศัยหลักการหาจุดศูนย์กลาง (Centroid) ของแต่ละกลุ่ม และจัดกลุ่มข้อมูลให้มีความคล้ายคลึงกันมากที่สุดในแต่ละรอบของการทำงาน

2.1.6.1 หลักการทำงานของ K-Means มีขั้นตอนดังนี้

- 1) กำหนดจำนวนกลุ่ม (K) ที่ต้องการแบ่ง และเลือกจุดศูนย์กลางเริ่มต้นจำนวน K จุด
- 2) นำข้อมูลทุกตัวมาเปรียบเทียบกับจุดศูนย์กลาง และจัดข้อมูลให้อยู่ในกลุ่มที่มีศูนย์กลางใกล้ที่สุด

3) คำนวณหาค่าเฉลี่ยใหม่ของข้อมูลในแต่ละกลุ่ม และใช้เป็นศูนย์กลางใหม่

4) ทำซ้ำขั้นตอนการจัดกลุ่มและการคำนวณศูนย์กลาง จนกว่าตำแหน่งของจุดศูนย์กลางจะไม่เปลี่ยนแปลงอีก

2.1.6.2 การประเมินจำนวนกลุ่มที่เหมาะสม ในการใช้งานจริง มักใช้ Elbow Method เพื่อหาจำนวน K ที่เหมาะสม โดยสังเกตจากกราฟค่าความคลาดเคลื่อน (Within-Cluster Sum of Squares: WCSS) ว่าจุดใดเป็น “ข้อศอก” ของกราฟ ซึ่งบ่งบอกถึงจำนวนกลุ่มที่เหมาะสมที่สุด (St, 2567)

2.1.6.3 ประโยชน์ของ K-Means Clustering

- ใช้สำหรับวิเคราะห์และแบ่งกลุ่มลูกค้า (Customer Segmentation)
- ช่วยค้นหารูปแบบพฤติกรรมที่คล้ายกัน ข้อมูล
- สามารถนำไปใช้ร่วมกับเทคนิคอื่น เช่น RFM Analysis เพื่อสร้างกลยุทธ์การตลาดที่ตรงกลุ่มเป้าหมาย

2.1.7 เทคนิคเกี่ยวกับกฎความสัมพันธ์ ด้วย FP-Growth

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

2.1.7.1 หลักการทำงานของ FP-Growth

FP-Growth (Frequent Pattern Growth) เป็นอัลกอริทึมที่ใช้ในการค้นหา ลวดลายที่เกิดขึ้นบ่อยในชุดข้อมูลขนาดใหญ่ โดยอัลกอริทึมนี้ถูกพัฒนาเพื่อให้สามารถทำงานได้ อย่างมีประสิทธิภาพมากกว่าอัลกอริทึมที่ใช้วิธีการสร้างชุดข้อมูลร่วม (Apriori) ซึ่งมีข้อเสียเรื่องการ ต้องการการคำนวณและการจัดเก็บข้อมูลจำนวนมาก หลักการทำงานของ FP-Growth สามารถ แบ่งออกเป็นขั้นตอนหลักดังนี้: การสร้าง FP-Tree ขั้นตอนแรกคือการสร้างต้นไม้ FP (Frequent Pattern Tree) ซึ่งเป็นโครงสร้างข้อมูลที่ใช้ในการเก็บข้อมูลเกี่ยวกับ ลวดลายที่เกิดขึ้นบ่อย ใน ขั้นตอนนี้ อัลกอริทึมจะทำการสแกนชุดข้อมูลเพื่อคำนวณความถี่ของรายการแต่ละรายการ 22 จากนั้นจะกรองเอาเฉพาะรายการที่เกิดขึ้นบ่อยออกมาและ

สร้างต้นไม้ FP โดยการจัดระเบียบ รายการเหล่านี้ตามลำดับความถี่การค้นหาลวดลายที่เกิดขึ้นบ่อย: หลังจากที่สร้าง FP-Tree เรียบร้อยแล้ว อัลกอริทึมจะใช้ FP-Tree ในการค้นหาลวดลายที่เกิดขึ้นบ่อยโดยการใช้วิธีการ เรียกใช้ FP-Growth Algorithm ซึ่งจะค้นหาและรวบรวมลวดลายที่เกิดขึ้นบ่อยจากต้นไม้ FP โดยการแบ่งต้นไม้ออกเป็นหลาย ๆ ส่วนและดำเนินการค้นหาในแต่ละส่วนการสร้างลวดลายสุดท้าย: เมื่อได้ลวดลายที่เกิดขึ้นบ่อยจาก FP-Tree แล้ว ขั้นตอนสุดท้ายคือการสร้างลวดลายสุดท้ายที่สามารถใช้ในการวิเคราะห์ข้อมูลเพิ่มเติมหรือในการทำการค้าการณ FP-Growth เป็นอัลกอริทึมที่มีข้อดีอย่างมากในการจัดการกับชุดข้อมูลขนาดใหญ่ เนื่องจากมันสามารถลดปริมาณการคำนวณ และการจัดเก็บข้อมูลได้อย่างมีประสิทธิภาพ โดยไม่ต้องทำการสร้างชุดข้อมูลรวมทั้งหมด เหมือนกับอัลกอริทึมแบบเดิม

2.1.7.2 การประยุกต์ใช้งาน FP-Growth

1) การวิเคราะห์ตลาดและการซื้อสินค้า: FP-Growth ใช้ในการวิเคราะห์ พฤติกรรมการซื้อของลูกค้า โดยสามารถค้นหาความสัมพันธ์ระหว่างสินค้าที่ลูกค้ามักจะซื้อพร้อมกัน เช่น ถ้าลูกค้าซื้อกาแฟ อาจจะมีแนวโน้มที่จะซื้อครีมบำรุงผิวด้วย การใช้ FP-Growth ช่วยให้ธุรกิจ สามารถจัดโปรโมชั่นหรือวางแผนการจัดเรียงสินค้าในร้านค้าได้อย่างมีประสิทธิภาพ

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

3) การวิเคราะห์ข้อมูลเว็บ: FP-Growth สามารถนำมาใช้ในการวิเคราะห์ ข้อมูลการเข้าชมเว็บไซต์ เช่น การค้นหาความสัมพันธ์ระหว่างหน้าที่ผู้ใช้เข้าชมและพฤติกรรมการ ใช้เว็บไซต์ ซึ่ง ช่วยในการออกแบบเว็บไซต์ให้มีประสิทธิภาพมากขึ้นและตอบสนองความต้องการ ของผู้ใช้ได้ดียิ่งขึ้น (schooltha, 2567)

2.1.8 แนวคิดเกี่ยวกับ Decision Tree (Decision Tree Concept)

Decision Tree (ต้นไม้ตัดสินใจ) เป็นอัลกอริทึมประเภทหนึ่งของ Machine Learning ที่อยู่ในกลุ่ม Supervised Learning (การเรียนรู้แบบมีผู้สอน) โดยเป็นแบบจำลองที่ใช้ได้ทั้งสำหรับการ จำแนกประเภท (Classification) และการ พยากรณ์ค่า (Regression) ของข้อมูล

Decision Tree ทำงานโดยสร้างแบบจำลองที่มีลักษณะเป็นโครงสร้างต้นไม้ที่ประกอบด้วยเงื่อนไขการตัดสินใจ ซึ่งแนวคิดหลักคือการสร้างเงื่อนไข If-else จากตัวแปรต่าง ๆ ในชุดข้อมูล เพื่อแบ่งข้อมูลออกเป็นส่วนย่อย ๆ อย่างต่อเนื่อง คล้ายกับการแตกปัญหาใหญ่ให้เป็นปัญหาย่อย ๆ จนกว่าจะได้ข้อสรุปสุดท้ายที่ชัดเจน

2.1.8.1 โครงสร้างและองค์ประกอบหลัก ประกอบด้วยองค์ประกอบหลักที่เชื่อมโยงกันในลักษณะลำดับขั้น ได้แก่:

1) โหนดราก (Root Node): เป็นโหนดเริ่มต้นของการวิเคราะห์ ซึ่งเป็นตัวแทนของชุดข้อมูลทั้งหมดที่ใช้ในการเรียนรู้

2) โหนดภายใน (Internal Node / Test Node): เป็นโหนดที่ใช้ตั้งคำถามหรือเงื่อนไขในการตัดสินใจ โดยจะมีการทดสอบคุณสมบัติของข้อมูลและแยกข้อมูลไปยังโหนดลูกตามผลลัพธ์ของเงื่อนไขนั้น

3) กิ่ง (Branch): คือเส้นเชื่อมต่อที่แสดงผลลัพธ์ที่เป็นไปได้ของการตัดสินใจในแต่ละโหนด

4) โหนดใบ (Leaf Node / Terminal Node): เป็นโหนดสุดท้ายที่อยู่ปลายทางสุดของโครงสร้าง ซึ่งแสดงถึงผลลัพธ์ของการจำแนกประเภท หรือการพยากรณ์

2.1.8.2 หลักการทำงานของ การแบ่งข้อมูล

การสร้าง Decision Tree จะใช้วิธีการแบ่งข้อมูลแบบวนซ้ำ (Recursive Partitioning) โดยมีหลักการที่สำคัญคือ:

- การเลือกคุณสมบัติที่ดีที่สุด: ในแต่ละโหนด อัลกอริทึมจะเลือกคุณสมบัติที่เหมาะสมที่สุดในการแบ่งข้อมูล เพื่อให้ข้อมูลในโหนดลูกที่แยกออกไปนั้นมีความเป็นเนื้อเดียวกันมากที่สุด (Homogeneity)

- เกณฑ์การแบ่งข้อมูล (Splitting Criteria): การตัดสินใจเลือกคุณสมบัติในการแบ่งจะใช้ค่าทางสถิติที่เรียกว่า เกณฑ์ความบริสุทธิ์ (Impurity Measure) ซึ่งจะมีการคำนวณและเลือกการแบ่งที่ทำให้ค่าความไม่บริสุทธิ์ของข้อมูลในโหนดลูกลดลงมากที่สุด เกณฑ์ที่นิยมใช้ เช่น Gini Impurity หรือ Entropy

- Gini Impurity เป็นการวัดโอกาสที่องค์ประกอบในชุดข้อมูลที่สุ่มมาอย่างสุ่มจะถูกจัดประเภทผิดพลาด หากองค์ประกอบนั้นถูกจัดประเภทตามการแจกแจงของคลาสในชุดข้อมูลนั้น หาก Gini Impurity มีค่าเข้าใกล้ศูนย์ หมายความว่าโหนดนั้นมีความบริสุทธิ์สูง (ข้อมูลเป็นคลาสเดียวกันเกือบทั้งหมด)

- การหยุดสร้างต้นไม้ (Stopping Condition): กระบวนการแบ่งข้อมูลจะหยุดเมื่อโหนดใบมีความบริสุทธิ์ถึงระดับที่กำหนด (เช่น ข้อมูลทั้งหมดในโหนดเป็นคลาสเดียวกัน) หรือเมื่อถึงเงื่อนไขอื่น ๆ ที่ตั้งไว้ล่วงหน้า เช่น จำนวนข้อมูลในโหนดน้อยเกินไป

2.1.8.3 ความสำคัญและการประยุกต์ใช้งาน

Decision Tree ได้รับความนิยมเนื่องจากเป็นแบบจำลองที่ง่ายต่อการทำความเข้าใจ และง่ายต่อการตีความ เนื่องจากโครงสร้างของมันคล้ายกับผังงาน จึงช่วยให้ผู้ใช้งานสามารถติดตามและทำความเข้าใจเหตุผลที่นำไปสู่ผลลัพธ์สุดท้ายได้ Decision Tree สามารถนำไปประยุกต์ใช้ได้อย่างกว้างขวางในหลายด้าน เช่น:

- การวิเคราะห์และสร้างกฎเกณฑ์เพื่อสนับสนุนการตัดสินใจทางธุรกิจ
- การสร้างแบบจำลองเพื่อคาดการณ์ความเสี่ยงที่ลูกค้าจะยกเลิกบริการ (Churn Prediction)

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับเหมืองข้อมูล (Data mining)

การทำเหมืองข้อมูล (Data Mining) คือการเก็บรวบรวมข้อมูลมหาศาลเพื่อวิเคราะห์ออกมาเป็นชุดข้อมูลที่เราสามารถใช้ประโยชน์กับองค์กรได้ โดยผ่านกระบวนการตัดแยก แบ่งกลุ่ม และกระบวนการอื่น ๆ อีกมากมายเพื่อจะได้ชุดข้อมูลที่เหมาะสมกับการนำไปใช้งาน ผลลัพธ์ของการทำ Data Mining อาจจะถูกออกมาในรูปแบบของความสัมพันธ์ของข้อมูลและแนวโน้ม

2.2.1.1 กระบวนการทำ Data Mining

1) วิเคราะห์วัตถุประสงค์ ก่อนที่จะเริ่มทำ Data Mining นักวิเคราะห์ต้องรู้ก่อนว่าจะนำผลลัพธ์ของข้อมูลนี้ไปใช้เพื่ออะไร ถ้านักวิเคราะห์ไม่เข้าใจปัญหาและเป้าหมายของการทำงาน ก็อาจทำให้ดำเนินการผิดวิธีและเสียเวลาไปโดยไม่ตอบโจทย์

2) เก็บข้อมูล เมื่อทราบวัตถุประสงค์แล้วว่าองค์กรมีเป้าหมายอะไรในการทำ Data Mining นักวิเคราะห์ข้อมูลต้องรู้ว่าข้อมูลที่ต้องการคืออะไรบ้าง หาได้จากที่ไหน จะเก็บและรักษาข้อมูลอย่างไร ข้อมูลที่มีคุณภาพนั้นต้องเป็นข้อมูลที่แม่นยำและน่าเชื่อถือ โดยที่ไม่ขัดต่อหลักความปลอดภัย ความเป็นส่วนตัว และข้อจำกัดอื่น ๆ บางครั้ง นักวิเคราะห์ต้องเก็บข้อมูลเป็นล้าน ๆ ชุดเพื่อที่จะได้เห็นภาพใหญ่ของสิ่งที่ต้องการ ซึ่งกลุ่มก้อนข้อมูลเหล่านี้จะถูกเรียกว่า *Big Data* เป็นเหมือนเหมืองขนาดใหญ่ที่มีแร่หลายชนิด

3) เตรียมข้อมูล กระบวนการนี้คือการนำข้อมูลที่มีอยู่มาคัดเฉพาะส่วนที่จำเป็นต้องใช้ในประเด็นย่อยต่าง ๆ เมื่อสกัดข้อมูลมาได้แล้วก็ต้องนำมาล้างทำความสะอาดให้พร้อมใช้มากขึ้น จากนั้นก็นำมาแบ่งหมวดหมู่และจัดเรียงภายใต้มาตรฐานเดียวกัน ตรวจสอบข้อผิดพลาด จนถึงประเมินขนาดของข้อมูลก่อนนำไปสู่ขั้นตอนถัดไป

4) เจียระไนข้อมูล หลังจากเตรียมข้อมูลเสร็จ นักวิเคราะห์จะนำข้อมูลเหล่านั้นมาทำให้พร้อมนำไปใช้งานได้มากขึ้นด้วยการเข้าโปรแกรมประมวลผลผ่าน Algorithm ที่ซับซ้อนเพื่อทำให้เห็นว่าข้อมูลเหล่านี้มีความสัมพันธ์หรือรูปแบบอย่างไรบ้าง ผลลัพธ์ที่ได้ออกมาเรียกว่า *Insight* ซึ่งอาจบอกได้ตั้งแต่พฤติกรรมผู้บริโภค เทรนด์ที่กำลังมา ไปจนถึงความเสี่ยงต่อธุรกิจ

5) นำเสนอข้อมูล ผลลัพธ์ที่ได้จะถูกนำมาสร้างเป็นโมเดลที่ทำให้คนทั่วไปเห็นภาพและเข้าใจข้อมูลมากขึ้น เปรียบเสมือนการนำข้อมูลที่สวดยามมาร้อยเรียงให้มีคุณค่ามากขึ้น

2.2.1.2 Data Mining ต้องอาศัยทักษะอะไร

1) ทักษะทางสถิติ (Statistics) ในการทำ Data Mining นั้นต้องอาศัยความสามารถในการจัดการกับตัวเลขหลายล้านชุดเพื่อประมวลผลความสัมพันธ์ของข้อมูล และต้องสามารถใช้โปรแกรมทางสถิติในการสร้างและคำนวณผลลัพธ์ด้วย นั้นยังไม่รวมถึงวิธีการคิดวิเคราะห์ที่เป็นระบบ ความสามารถในการเข้าใจตรรกะและความเชื่อมโยงที่ถูกต้อง การตั้งคำถามและเก็บรายละเอียดได้เป็นอย่างดี

2) ปัญญาประดิษฐ์ (Artificial Intelligence หรือ AI) นักวิเคราะห์ต้องใช้ความสามารถของ AI ในการจัดการกับข้อมูลมหาศาลได้อย่างรวดเร็วมาลดเวลาในการประมวลผล และเพิ่มประสิทธิภาพในการทำงานเช่นกัน

3) ทักษะด้านภาษาโปรแกรม (Programming Languages) หรือการเขียนโค้ด นักวิเคราะห์ต้องใช้ภาษาโปรแกรมสื่อสารกับคอมพิวเตอร์เพื่อสร้างคำสั่งจำเพาะบางอย่าง เช่น ภาษา Python และภาษา R การมีทักษะนี้จะทำให้นักวิเคราะห์ สามารถสร้างคำสั่งแบบเฉพาะเจาะจงขึ้นมาเองได้ ทำให้การทำงานมีความสร้างสรรค์และให้ผลลัพธ์ที่ตอบโจทย์มากขึ้น

4) Machine Learning ในการประมวลผลข้อมูลอาจต้องใช้ Algorithm ที่สร้างมาเพื่อใช้ในการเจียระไนข้อมูลชุดนั้น ๆ โดยเฉพาะ หากนักวิเคราะห์สามารถทำความเข้าใจตรรกะของ Machine Learning ได้ก็จะช่วยอำนวยความสะดวกในการจัดการข้อมูลมากขึ้น

5) ทักษะด้านฐานข้อมูล (Database) เป็นทักษะที่จำเป็นเพราะนักวิเคราะห์ต้องดึงข้อมูลจากฐานข้อมูลหลายแหล่งมาวิเคราะห์และจัดการ เช่น Structured Query Language (SQL)

6) ทักษะด้าน Data Visualization ช่วยทำให้การนำเสนอข้อมูลเชิงลึกที่ค้นพบนั้นเข้าใจง่ายขึ้น เป็นประโยชน์ทั้งต่อนักวิเคราะห์ที่ต้องใช้อธิบายผลลัพธ์และต่อองค์กรที่ต้องการนำผลลัพธ์ไปใช้ประโยชน์ รูปแบบของ Data Visualization นั้นมีหลากหลาย เช่น การใช้กราฟ ชาร์ต และแผนที่แบบต่าง ๆ (TNIC, 2567)

2.2.2 ทฤษฎีเกี่ยวกับการสร้างเว็บไซต์

2.2.2.1 การออกแบบเว็บไซต์ หรือ Web Design คือ กระบวนการในการสร้างหน้าตา และองค์ประกอบภายในเว็บไซต์ ตั้งแต่การจัดวางเนื้อหา การเลือกใช้สี ฟอนต์ การเลือกใช้ภาพ หรือกราฟิก และการออกแบบระบบนำทางผู้ใช้งานหน้าเว็บ ให้สวยงาม สอดคล้อง และสามารถใช้งานได้อย่างราบรื่น โดยใช้หลักการของ Responsive Web Design เข้ามาช่วยในการออกแบบ เพื่อให้สามารถแสดงผลได้ทุก ๆ อุปกรณ์ ซึ่งจะใช้ URL ร่วมกัน แต่การแสดงผลในแต่ละอุปกรณ์ จะแตกต่างกันไป การตอบสนองเป็นการปฏิบัติของการสร้างเว็บไซต์ที่เหมาะสมเพื่อทำงานบน ทุกอุปกรณ์ และทุกขนาดหน้าจอทั้งขนาดใหญ่ หรือขนาดเล็ก เช่น มือถือ หรือเดสก์ท็อป ฯลฯ การตอบสนองจะเน้นการใช้งานง่าย จึงต้องมีการออกแบบอย่างเป็นระบบ เว็บไซต์ที่ประสบความสำเร็จ จะต้องสอดคล้องกับแบรนด์ และเป้าหมายทางธุรกิจ สามารถสื่อสารคุณค่าของสินค้ากับบริการได้อย่างโดดเด่น ขณะเดียวกันก็ต้องเข้าใจความต้องการ ของกลุ่มเป้าหมาย เพื่อมอบประสบการณ์ใช้งานที่เหมาะสมกับผู้ใช้จริง

2.2.2.2 องค์ประกอบในการออกแบบเว็บไซต์ การออกแบบเว็บไซต์ที่มีประสิทธิภาพ นั้นต้องคำนึงถึงองค์ประกอบสำคัญ ดังต่อไปนี้

1) การออกแบบที่เรียบง่าย (Simplicity) ในยุคที่ผู้คนมีความสนใจในช่วงเวลา สั้นลง (Short attention span) และต้องการข้อมูลที่กระชับ และตรงประเด็น การออกแบบเว็บไซต์ แบบมินิมอล (Minimalist) ซึ่งเน้นความเรียบง่ายจึงเป็นแนวคิดที่ควรนำมาใช้ เพื่อให้ผู้ใช้งาน สามารถมุ่งเน้นประเด็นสำคัญได้อย่างชัดเจน

ทั้งนี้ความเรียบง่ายไม่ควรถึงขั้นที่ทำให้ขาดจุดเด่นที่น่าสนใจ เพราะ การออกแบบหน้าเว็บไซต์ที่ซับซ้อนเกินไปอาจทำให้ผู้ใช้งานเกิดความสับสน และไม่สามารถค้นหาข้อมูลที่ต้องการได้อย่างสะดวก ดังนั้น ควรใช้กราฟิกและข้อความในปริมาณที่พอเหมาะ เน้นที่ความสะอาดตา และการใช้งานที่ราบรื่น โดยยังสามารถสื่อสารข้อมูลสำคัญของหน้าเว็บไซต์ ได้อย่างมีประสิทธิภาพ

2) การเลือกใช้สีที่เข้ากัน (Color Scheme) การเลือกใช้สีเป็นองค์ประกอบ สำคัญในการออกแบบเว็บไซต์ (Web Design) เนื่องจากสีสามารถสื่อถึงอารมณ์และความรู้สึก รวมทั้งส่งผลต่อการรับรู้ของผู้ใช้งาน หากเลือกใช้สีที่เหมาะสม และสอดคล้องกับลักษณะของ เว็บไซต์ ย่อมสามารถสร้างความโดดเด่น และเอกลักษณ์ได้ดี การใช้สีไม่ควรเกิน 1 - 2 สีหลัก และสีรอง 2 - 3 สี เพื่อให้ดูเป็นระเบียบน่าสนใจ

นอกจากนี้ การเลือกสีควรคำนึงถึงความคมชัดในการแสดงผล ไม่ควรใช้ สีที่ฉูดฉาดเกินไปเพราะอาจทำให้การอ่านเนื้อหาลำบาก ควรเลือกใช้คู่สีที่เป็น Complementary colors หรือใช้ Color palette ที่มีโทนสีใกล้เคียงกัน เพื่อให้เว็บไซต์ดูสวยงามและน่าสบายตา

3) การเลือกฟอนต์ที่อ่านง่าย (Readable Fonts) ฟอนต์ และการจัดวาง ข้อความเป็นองค์ประกอบสำคัญในการออกแบบหน้าเว็บไซต์ (Web Page Design) เนื่องจากมีผล ต่อความสะดวกในการอ่าน ควรเลือกใช้ฟอนต์ที่อ่านง่าย และไม่ซับซ้อน เช่น ฟอนต์ Sans-serif อย่าง Arial Verdana Helvetica สำหรับเนื้อหาทั่วไป หรือฟอนต์ Serif เช่น Times New Roman Georgia ที่เหมาะกับเนื้อหาบทความยาว

ขนาดของฟอนต์ (Font size) และการเว้นระยะห่างระหว่างบรรทัด (Line-Height) ที่เหมาะสมจะช่วยเพิ่มความสะดวกในการอ่าน และนำเสนอข้อมูลได้อย่างมีประสิทธิภาพ

4) การใช้พื้นที่ว่างอย่างเหมาะสม (White Space) หรือพื้นที่ว่าง เป็นส่วนที่ปราศจากข้อความหรือกราฟิก ซึ่งมักใช้เป็นพื้นหลังสีขาว หรือสีอ่อน เพื่อแบ่งแยก เนื้อหาให้เป็นสัดส่วน สร้างความเป็นระเบียบเรียบร้อย และช่วยให้ผู้ใช้เข้าใจเนื้อหาได้ง่ายขึ้น สีพื้น หลังไม่จำเป็นต้องเป็นสีขาว อาจใช้สีอ่อนหรือเส้นที่ไม่ทำให้ดูรกตา สิ่งสำคัญคือ ต้องใช้ White Space อย่างพอเหมาะ ไม่มาก หรือน้อยเกินไป

5) การออกแบบเว็บไซต์ให้รองรับการแสดงผลบนมือถือ ด้วยการเพิ่มขึ้นของ ผู้ใช้งานอินเทอร์เน็ตผ่านสมาร์ทโฟน และแท็บเล็ต การออกแบบเว็บไซต์ที่รองรับการแสดงผล บนอุปกรณ์พกพา หรือ Responsive Design จึงกลายเป็นปัจจัยสำคัญ Responsive Design คือการออกแบบเว็บไซต์ให้ปรับขนาด และการจัดเรียงเนื้อหาได้โดยอัตโนมัติ ไม่ว่าจะเป็นคอมพิวเตอร์ สมาร์ทโฟน หรือแท็บเล็ต ทำให้ผู้ใช้สามารถใช้งานได้อย่างราบรื่น

6) การใช้ Visual Content ในการตกแต่งเว็บไซต์ Visual Content เช่น รูปภาพ ไอคอน อินโฟกราฟิก และวิดีโอ เป็นองค์ประกอบที่ช่วยเสริมสร้างความ

น่าสนใจให้กับเว็บไซต์ นอกจากการตกแต่งให้สวยงาม ยังช่วยดึงดูดความสนใจ และอธิบายเนื้อหาที่ยากให้เข้าใจง่ายขึ้น การใช้ Visual Content ควรอยู่ในปริมาณที่เหมาะสม และต้องคำนึงถึงขนาดไฟล์เพื่อไม่ให้กระทบ ต่อความเร็วในการโหลดเว็บไซต์

7) การวางตำแหน่งเนื้อหาหลักให้เด่นชัด (Highlight Key Content) การวางตำแหน่งเนื้อหาที่ต้องการเน้นให้เห็นได้ชัดเจนเป็นสิ่งสำคัญ เพราะผู้ชมมักให้ความสนใจกับ สิ่งที่เห็นชัดเจนที่สุดก่อน เช่น ข้อความขนาดใหญ่ หรือรูปภาพเด่นชัด ควรใช้หลักการจัดลำดับ ความสำคัญขององค์ประกอบ (Visual Hierarchy) เพื่อทำให้เนื้อหาสำคัญโดดเด่น และสะดุดตา มากขึ้น

8) การออกแบบระบบนำทางเว็บไซต์ (Navigation) การออกแบบระบบนำทาง หรือเมนูที่ใช้งานง่ายเป็นหัวใจสำคัญ เพราะช่วยให้ผู้ใช้งานสามารถค้นหา และเข้าถึงข้อมูลได้อย่าง รวดเร็ว ควรวางโครงสร้างเมนูให้ชัดเจน และเป็นหมวดหมู่ การใช้ข้อความในเมนูควรกระชับ และเข้าใจง่าย

9) การใส่ Call-to-Action (CTA) เป็นองค์ประกอบที่กระตุ้นให้ผู้ใช้ดำเนินการ ตามที่เว็บไซต์ต้องการ เช่น การกดสั่งซื้อ หรือการสมัครสมาชิก ปุ่ม CTA ควรมีขนาดใหญ่พอ ที่จะสังเกตเห็นได้ชัดเจน และข้อความควรกระชับ ชัดเจน เช่น "สั่งซื้อทันที" หรือ "ดาวน์โหลดฟรี"

10) การทดสอบ และปรับปรุงการออกแบบเว็บไซต์อยู่เสมอ การออกแบบ เว็บไซต์เป็นกระบวนการที่ต้องมีการทดสอบ และปรับปรุงอย่างต่อเนื่อง เพื่อให้สอดคล้อง กับเทคโนโลยีที่เปลี่ยนแปลง และพฤติกรรมของผู้ใช้ การใช้เครื่องมืออย่าง Google Analytics ช่วยให้เห็นพฤติกรรมการใช้งาน และสามารถปรับปรุงเว็บไซต์ให้ตอบโจทย์ความต้องการ ได้อย่างมีประสิทธิภาพ

2.2.2.3 รูปแบบโครงสร้างของเว็บไซต์ การออกแบบโครงสร้างของเว็บไซต์ สามารถ ทำได้หลากหลายแบบ ซึ่งก็ขึ้นอยู่กับความชอบ และความถนัดของแต่ละบุคคล นอกจากนี้ยังขึ้นอยู่กับกลุ่มเป้าหมายที่ต้องการนำเสนอ เพราะจะต้องออกแบบให้เหมาะสม กับ การใช้งาน ของกลุ่มเป้าหมายมากที่สุด โดยโครงสร้างของเว็บไซต์ส่วนใหญ่ก็จะประกอบไปด้วย 4 รูปแบบดังนี้

1) โครงสร้างเว็บไซต์แบบเรียงลำดับ จะเป็นโครงสร้างแบบธรรมดาที่นิยม ใช้งานกันมากที่สุด เนื่องจากมีความง่ายต่อการจัดระบบข้อมูล และสามารถนำเสนอเรื่องราว ตามลำดับได้เป็นอย่างดี เหมาะกับเว็บไซต์ที่มีขนาดเล็ก มีเนื้อหาที่ไม่ซับซ้อน ส่วนใหญ่ก็จะเป็น พวกเว็บไซต์ที่ให้ความรู้ หรือเว็บไซต์องค์กรขนาดย่อม โดยลักษณะการลิ้งค์

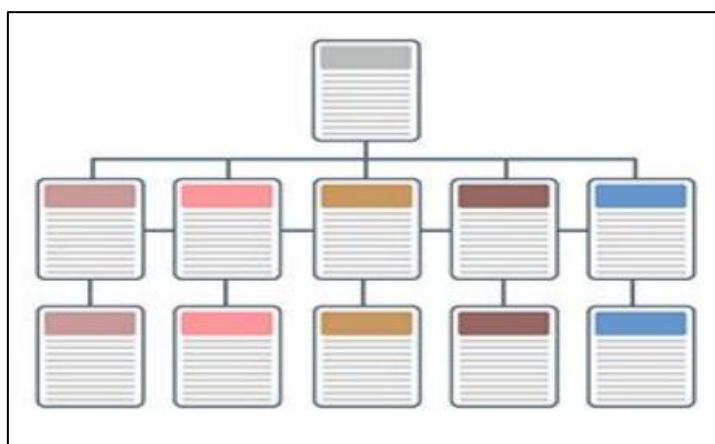
เนื้อหา ก็จะลึ้งค์ไป ทีละหน้า มีทิศทางกรเข้าสู่เนื้อหาต่าง ๆ ในแบบเส้นตรง ใช้ปุ่มเดินหน้า - ถอยหลังในการกำหนด ทิศทาง จึงทำให้การใช้งานเป็นไปอย่างง่าย แต่โครงสร้างเว็บไซต์แบบ เรียงลำดับมีข้อเสีย คือจะทำให้ผู้ใช้งานต้องเสียเวลาในการเข้าสู่เนื้อหาเพราะไม่สามารถ กำหนดทิศทางกรเข้า สู่เนื้อหาด้วยตัวเองได้



ภาพที่ 2.8 แสดงโครงสร้างแบบเรียงลำดับ

(ที่มา: <https://asria.org/website-structure/>)

2) โครงสร้างแบบลำดับชั้น นิยมใช้กับเว็บที่มีความซับซ้อนของข้อมูล เพื่อให้ สามารถเข้าถึงข้อมูลต่าง ๆ ได้ง่ายขึ้น โดยจะมีการแบ่งเนื้อหาออกเป็น ส่วน ๆ และมีการนำเสนอ รายละเอียดย่อย ๆ ที่ลดหลั่นกันมา ทำให้สามารถทำความเข้าใจกับโครงสร้าง เนื้อหาได้ง่ายขึ้น โดยจะมีโฮมเพจเป็นจุดเริ่มต้น และจุดรวมจุดเดียวที่จะนำไปสู่การเชื่อมโยง เนื้อหาเป็นลำดับ จากบนลงล่าง

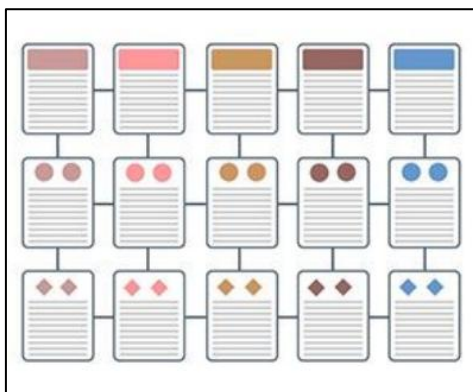


ภาพที่ 2.9 แสดงโครงสร้างแบบลำดับชั้น

(ที่มา: <https://asria.org/website-structure/>)

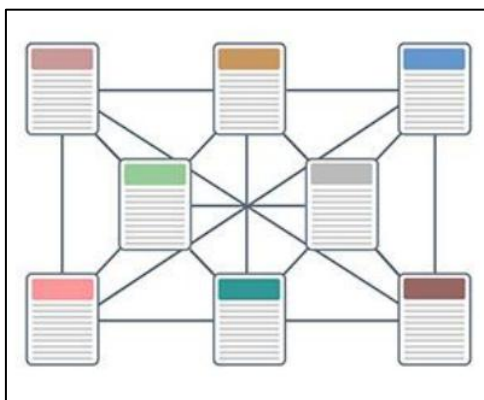
3) โครงสร้างแบบตาราง เป็นโครงสร้างการออกแบบเว็บไซต์ที่มีความซับซ้อน แต่ก็มีความยืดหยุ่นในระดับหนึ่ง เพื่อให้ผู้ใช้งานสามารถเข้าสู่เนื้อหาต่าง ๆ ได้ง่ายขึ้น การ

ออกแบบ ในลักษณะนี้จะมีการเชื่อมโยงเนื้อหาในแต่ละส่วนซึ่งกัน และกัน ทำให้ผู้ใช้งานสามารถเปลี่ยน ทิศทาง หรือกำหนดทิศทางในการเข้าสู่เนื้อหาด้วยตัวเองได้ จึงไม่ทำให้เสียเวลา และทำให้เว็บไซต์ มีความทันสมัยขึ้น



ภาพที่ 2.10 แสดงโครงสร้างแบบตาราง
(ที่มา: <https://asria.org/website-structure/>)

4) โครงสร้างแบบใยแมงมุม เป็นโครงสร้างที่ได้รับความนิยมเป็นอย่างมาก เพราะมีความยืดหยุ่นมากที่สุด โดยทุกหน้าเว็บจะมีการเชื่อมโยงถึงกันหมด ทำให้สามารถเข้า ถึงหน้าเว็บเพจต่าง ๆ ที่ต้องการได้อย่างง่าย และมีความอิสระมากขึ้น นอกจากนี้ก็สามารถเชื่อมโยงไปสู่เว็บไซต์ภายนอกได้ดี



ภาพที่ 2.11 แสดงโครงสร้างแบบใยแมงมุม
(ที่มา: <https://asria.org/website-structure/>)

2.2.2.4 ส่วนประกอบหลักในหน้าเว็บไซต์ มีองค์ประกอบหลัก ๆ เช่น ชื่อเว็บไซต์ แถบเมนูนำทาง แถบเมนูย่อย ข้อความหัวเรื่อง ภาพประกอบตกแต่ง หรือภาพ

เนื้อหา ป้ายโฆษณา และอื่น ๆ ตามที่ต้องการนำเสนอ บางเว็บไซต์อาจจะแสดงภาพกราฟิกของสินค้า หรือบริการ แทนการแสดงข้อความ ทั้งนี้ก็ขึ้นอยู่กับการออกแบบหน้าเว็บไซต์ให้เหมาะสมกับเนื้อหาที่ต้องการ นำเสนอ

ส่วนประกอบบนหน้าเว็บเพจ มีทั้งหมด 6 องค์ประกอบ ดังนี้

1) ส่วนหัว (Header) สิ่งสำคัญที่ช่วยสร้างการจดจำให้กับธุรกิจ คือ โลโก้ (Logo) ซึ่งทำหน้าที่เป็นสัญลักษณ์ที่สะท้อนถึงตัวตน และลักษณะเฉพาะของธุรกิจอย่างชัดเจน การออกแบบโลโก้ที่เหมาะสมจะช่วยให้ผู้เยี่ยมชมเว็บไซต์สามารถเข้าใจได้ทันทีว่าเว็บไซต์นั้นมีเนื้อหาหรือบริการเกี่ยวข้องกับอะไร โลโก้จึงถือเป็นองค์ประกอบที่สำคัญในการสื่อสารภาพลักษณ์ ของธุรกิจ ควรจัดวางโลโก้ให้อยู่บริเวณด้านบนสุดของหน้าแรกของเว็บไซต์ เพื่อให้ผู้เข้าชม สามารถมองเห็นได้ทันทีที่เข้าสู่หน้าเว็บไซต์ และสร้างการจดจำธุรกิจของคุณอย่างมีประสิทธิภาพ

2) เมนู (Navigator) แถบเมนูเป็นองค์ประกอบสำคัญที่ช่วยอำนวยความสะดวก ในการใช้งานเว็บไซต์ให้กับผู้เยี่ยมชม ควรออกแบบแถบเมนูให้สอดคล้องกับโครงสร้าง และการใช้งานของเว็บไซต์ โดยตำแหน่งของเมนูสามารถอยู่ในตำแหน่งเดียวกับโลโก้ด้านบนของหน้าเว็บไซต์ หรือวางไว้ที่ด้านซ้ายของหน้าจอก็ได้ ทั้งนี้ เมนูควรถูกจัดเรียงตามลำดับความสำคัญ เพื่อให้ผู้เยี่ยมชม สามารถเข้าถึงข้อมูลและบริการต่าง ๆ ของเว็บไซต์ได้อย่างรวดเร็ว และสะดวก อย่างไรก็ตาม ควรหลีกเลี่ยงการใส่เมนูมากเกินไป เพราะอาจก่อให้เกิดความสับสนในการใช้งาน และส่งผลให้การนำทางภายในเว็บไซต์ดูซับซ้อน และไม่เป็นระเบียบ

3) เนื้อหา (Contents) ในหน้าแรกของเว็บไซต์ธุรกิจ ควรนำเสนอเนื้อหาที่ชัดเจน และครอบคลุมรายละเอียดเกี่ยวกับธุรกิจ หรือบริการของคุณ เพื่อให้ผู้เยี่ยมชมได้รู้จักธุรกิจของคุณมากยิ่งขึ้น ข้อมูลที่ควรนำเสนอ ได้แก่ การแนะนำว่าสินค้าหรือบริการของคุณคืออะไร คุณสมบัติเด่นของสินค้า หรือบริการ ประโยชน์ที่ผู้ใช้จะได้รับ ข่าวสารหรือกิจกรรมสำคัญ รวมถึง โปรโมชั่นที่เกี่ยวข้อง ช่องทางการติดต่อที่สะดวกและเข้าถึงง่าย คำรับรองจากลูกค้าที่เคยใช้บริการ รวมถึงรางวัล หรือการรับรองคุณภาพบริการที่ธุรกิจของคุณได้รับ นอกจากนี้ หากต้องการเพิ่มโอกาสให้ธุรกิจของคุณถูกค้นพบได้ง่ายขึ้น ควรใส่ คำสำคัญ (Keywords) ที่เกี่ยวข้องกับสินค้า หรือบริการในเนื้อหาของเว็บไซต์ ซึ่งจะช่วยให้ กลุ่มเป้าหมายสามารถเข้าถึงเว็บไซต์ของคุณได้สะดวก และตรงกับความต้องการมากยิ่งขึ้น

4) ฟังก์ชันการค้นหา (Search Function) การจัดทำฟังก์ชันการค้นหาในเว็บไซต์ เป็นสิ่งสำคัญที่ช่วยให้ผู้เยี่ยมชมสามารถค้นหาสินค้าหรือบริการได้อย่างรวดเร็วและสะดวก

ยิ่งขึ้น โดยเฉพาะในกรณีที่เว็บไซต์ของคุณมีสินค้าหรือบริการจำนวนมาก ฟังก์ชันนี้จะช่วยให้ผู้ใช้งานสามารถ ค้นหาสิ่งที่ต้องการได้อย่างง่ายดาย การออกแบบ และพัฒนาฟังก์ชันการค้นหาที่มีประสิทธิภาพ จึงเป็นสิ่งที่ควรให้ความสำคัญ เพื่อให้ผู้เข้าชมสามารถเข้าถึงข้อมูลที่ต้องการได้โดยไม่ยุ่งยาก และเพิ่มประสบการณ์การใช้งานที่ดีขึ้นบนเว็บไซต์

5) ส่วนท้าย (Footer) ส่วนท้ายของเว็บไซต์ควรถูกออกแบบให้ช่วยอำนวยความสะดวกแก่ผู้เข้าชมในการนำทางไปยังหน้าต่าง ๆ ภายในเว็บไซต์ พร้อมทั้งเชื่อมโยงไปยังแพลตฟอร์ม โซเชียลมีเดีย (Social Media) ข้อมูลที่ควรแสดงในส่วนนี้ประกอบด้วย ชื่อของธุรกิจ ช่องทางการติดต่อ เช่น หมายเลขโทรศัพท์และอีเมล ที่ตั้งของบริษัท รวมถึงลิงก์ไปยังโซเชียลมีเดีย ต่าง ๆ และลิงก์ที่เชื่อมโยงไปยังหน้าหลักอื่น ๆ ของเว็บไซต์ ทั้งนี้ การจัดวางข้อมูลในส่วนท้ายของเว็บไซต์ควรคำนึงถึงการใช้งานที่ง่ายและการเข้าถึงข้อมูลได้อย่างรวดเร็ว

2.2.2.5 ทฤษฎีสีในการออกแบบเว็บไซต์

การสร้างสีบนหน้าเว็บเป็นสิ่งที่สื่อความหมายของเว็บไซต์ได้อย่าง ชัดเจนการเลือกใช้สีให้เหมาะสม กลมกลืน ไม่เพียงแต่จะสร้างความพึงพอใจให้กับผู้ใช้ แต่ยังสามารถทำให้เห็นถึงความแตกต่างระหว่างเว็บไซต์ได้ สีเป็นองค์ประกอบหลักสำหรับการ ตกแต่งเว็บ จึงจำเป็นอย่างยิ่งที่จะต้องทำความเข้าใจเกี่ยวกับการใช้สีระบบสีที่แสดงบน จอคอมพิวเตอร์ มีระบบการแสดงผลผ่านหลอดลำแสงที่เรียกว่า CRT (Cathode ray tube) โดย มีลักษณะระบบสีแบบบวก อาศัยการผสมของแสงสีแดง สีเขียว และสีน้ำเงิน หรือระบบสี RGB สามารถกำหนดค่าสีจาก 0 ถึง 255 ได้ จากการรวมสีของแม่สีหลักจะทำให้เกิดแสงสีขาว มีลักษณะเป็นจุดเล็ก ๆ บนหน้าจอไม่สามารถมองเห็นด้วยตาเปล่าได้ จะมองเห็นเป็นสีที่ถูก ผสมเป็นเนื้อสีเดียวกันแล้ว จุดแต่ละจุด หรือ พิกเซล (Pixel) เป็นส่วนประกอบของภาพบนหน้าจอคอมพิวเตอร์ โดยจำนวนบิตที่ใช้ในการกำหนดความสามารถของการแสดงสีต่าง ๆ เพื่อสร้างภาพบนจอ นั้นเรียกว่า บิตเดป (Bit depth) ในภาษา HTML มีการกำหนดสีด้วยระบบเลขฐานสิบหก ซึ่งมีเครื่องหมาย (#) อยู่ด้านหน้าและตามด้วยเลขฐานสิบหกจำนวนอักษรอีก 6 หลัก โดยแต่ละไบต์ (byte) จะมีตัวอักษรสองตัว แบ่งออกเป็น 3 กลุ่ม เช่น #FF12AC การใช้ตัวอักษรแต่ละไบต์นี้เพื่อกำหนดระดับความเข้มของแม่สีแต่ละสีของชุดสี RGB โดย 2 หลักแรกแสดงถึงความเข้มของสีแดง 2 หลักต่อมา แสดงถึงความเข้มของสีเขียว 2 หลักสุดท้ายแสดงถึงความเข้มของสีน้ำเงิน

สีมีอิทธิพลในเรื่องของอารมณ์การสื่อความหมายที่เด่นชัด กระตุ้นการ รับรู้ทางด้านจิตใจมนุษย์ สีแต่ละสีให้ความรู้สึกถึงอารมณ์ที่ไม่เหมือนกัน สีบางสีให้ความรู้สึก สงบ บางสีให้ความรู้สึกตื่นเต้นรุนแรง สีจึงเป็นปัจจัยสำคัญอย่างยิ่งต่อการออกแบบเว็บไซต์ ดังนั้น

การเลือกใช้โทนสีภายในเว็บไซต์เป็นการแสดงถึงความแตกต่างของสีที่แสดงออกทางอารมณ์ มีชีวิตชีวาหรือเศร้าโศก รูปแบบของสีที่สายตาของมนุษย์มองเห็น สามารถแบ่ง ออกเป็น 3 กลุ่ม คือ

1) สีโทนร้อน (Warm Colors) เป็นกลุ่มสีที่แสดงถึงความสุข ความ ปลอดภัย ความอบอุ่น และดึงดูดใจ สีกลุ่มนี้เป็นกลุ่มสีที่ช่วยให้หายจากความเฉื่อยชา มีชีวิตชีวามากยิ่งขึ้น

2) สีโทนเย็น (Cool Colors) แสดงถึงความที่ดูสุภาพ อ่อนโยน เรียบร้อย เป็นกลุ่มสีที่มีคนชอบมากที่สุด สามารถโน้มน้าวในระยะไกลได้

3) สีโทนกลาง (Neutral Colors) สีที่เป็นกลาง ประกอบด้วย สีดำ สีขาว สีเทา และสีน้ำตาล กลุ่มสีเหล่านี้คือ สีกลางที่สามารถนำไปผสมกับสีอื่นๆ เพื่อให้เกิดสีกลางขึ้นมา

2.2.3 ทฤษฎีเกี่ยวกับ Machine Learning

Machine Learning (การเรียนรู้ของเครื่อง) คือ สาขาหนึ่งของ ปัญญาประดิษฐ์ (Artificial Intelligence: AI) ที่ทำให้คอมพิวเตอร์สามารถเรียนรู้และปรับปรุง ประสิทธิภาพการทำงานได้เองจากข้อมูลที่มีอยู่ โดยไม่จำเป็นต้องมีการเขียนโปรแกรมทุก ขั้นตอน เหมือนกับที่มนุษย์สามารถเรียนรู้ได้จากสิ่งที่เกิดขึ้นแล้วหรือประสบการณ์

2.2.3.1 Machine Learning คือ กระบวนการที่ระบบคอมพิวเตอร์

สามารถเรียนรู้จากข้อมูลและประสบการณ์ เพื่อปรับปรุงความสามารถในการทำงานให้ดีขึ้นในอนาคต โดยที่ระบบจะสร้างตัวแบบ หรือโมเดล (Model) ขึ้นจากชุดข้อมูลตัวอย่าง (Training Data) เพื่อคาดการณ์หรือตัดสินใจ โดยไม่จำเป็นต้องมีการเขียนชุดคำสั่งหรือโปรแกรมเพื่อควบคุมการทำงานในทุกขั้นตอน

หรืออาจกล่าวได้ว่า คอมพิวเตอร์สามารถเรียนรู้จากข้อมูลที่มีอยู่ แล้วใช้สิ่งที่ได้เรียนรู้ในการทำนายหรือตัดสินใจผลลัพธ์ที่ (อาจจะ) เกิดขึ้นเมื่อมีข้อมูลใหม่เข้ามานั่นเอง

ตัวอย่างของการใช้งาน Machine Learning ได้แก่:

- การคาดการณ์ราคาหุ้น
- ระบบแนะนำ (Recommendation Systems) เช่นใน Netflix หรือ YouTube
- การจดจำเสียงและภาพ เช่น Google Assistant และ Face ID
- การวิเคราะห์ทางการแพทย์ เช่น การวินิจฉัยโรคจากภาพเอกซเรย์

2.2.3.2 ประเภทของ Machine Learning การเรียนรู้ของเครื่องสามารถ

แบ่งได้เป็น 3 ประเภทหลัก ได้แก่

1) การเรียนรู้แบบมีผู้สอน (Supervised Learning): โมเดลเรียนรู้จากข้อมูลที่มีป้ายกำกับ (Labeled Data) เช่น การทำนายราคาบ้านโดยใช้ข้อมูลราคาบ้านที่เคยถูกขายไปก่อนหน้านี้

2) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning): โมเดลวิเคราะห์ข้อมูลที่ไม่มีป้ายกำกับ (Unlabeled Data) เช่น การจัดกลุ่มลูกค้า (Customer Segmentation) โดยวิเคราะห์ถึงพฤติกรรมต่าง ๆ เช่น ประเภทสินค้าที่ซื้อเป็นประจำ หรือจำนวนเงินที่ใช้จ่าย

3) การเรียนรู้แบบเสริมแรง (Reinforcement Learning): โมเดลเรียนรู้ผ่านการทดลองและการให้รางวัล เช่น การพัฒนาหุ่นยนต์ที่สามารถเดินได้อย่างมีประสิทธิภาพ หรือการเรียนรู้โดยการลองผิดลองถูกนั่นเอง (Blog, 2568)

2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่ง HTML

HTML (HyperText Markup Language) เป็นภาษาที่ใช้สำหรับสร้างเว็บเพจ มีโครงสร้างภาษาโดยใช้ตัวกำกับ (Markup Tag) เพื่อควบคุมการแสดงผลข้อมูล รูปภาพ และวัตถุอื่น ๆ ผ่านทาง Web Browser เช่น Google Chrome , Firefox , Safari , Microsoft Edge

2.2.4.1 การสร้างไฟล์ HTML จะต้องอาศัย Text Editor เพื่อใช้สำหรับเขียนคำสั่งต่าง ๆ ที่ต้องการแสดงผลทางจอภาพ / เว็บเบราว์เซอร์ และเก็บเป็นไฟล์โดยมีนามสกุล .html

2.2.4.2 HTML 5.0 มาตรฐานของภาษา HTML มีการจัดโครงสร้างและการแสดงผลของเนื้อหาสำหรับ www มาตรฐานใหม่มีคุณลักษณะเด่นที่สำคัญ ตัวอย่าง เช่น

- เล่นวิดีโอ
- แสดงตำแหน่งทางภูมิศาสตร์
- เก็บไฟล์ในลักษณะออฟไลน์
- แสดงกราฟิก
- การป้อนข้อมูลแบบใหม่ เช่น search, number, range, color, tel, url, email, date, month, week, time, datetime, datetime-local

2.2.4.3 กำหนดรูปแบบ Character encoding ในหน้าเว็บ คือ การกำหนดรูปแบบการเข้ารหัสอักขระ(Character encoding) โดยใช้แท็ก<meta> กำหนด Attribute charset ลงไป

```
<meta charset="utf-8">
<meta http-equiv="Content-Type" content="text/html; charset=utf-8">
```

ภาพที่ 2.12 กำหนดรูปแบบ Character encoding

(ที่มา: <https://kongruksiam.medium.com/พื้นฐาน-html5-สำหรับการพัฒนาเว็บแอปพลิเคชัน-44d815e68390/>)

2.2.4.4 โครงสร้างของ HTML

1) ส่วน head เป็นส่วนที่อยู่ภายใน <head> ... </head> ใช้สำหรับอธิบายข้อมูลเกี่ยวกับเว็บ เช่น ชื่อเรื่องของเว็บเพจ (Title) ชื่อผู้จัดทำเว็บ (Author) คีย์เวิร์ด (Keywords) เพื่อใช้สำหรับให้ผู้ใช้ค้นหาข้อมูลเกี่ยวกับเว็บได้

2) ส่วน body เป็นส่วนที่อยู่ระหว่าง <body> ... </body> ใช้อธิบายเนื้อหาหลักของเว็บ เช่น ใส่ข้อความต่างๆ รูปภาพ แบบฟอร์ม วิดีโอและยังสามารถกำหนดคุณสมบัติพื้นฐานของเว็บได้ เช่น รูปแบบของพื้นหลัง สีของตัวอักษร (Studio, 2567)

2.2.5 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS

CSS (Cascading Style Sheets) เป็นภาษาที่ใช้ในการออกแบบและจัดรูปแบบเว็บไซค์ โดยใช้ควบคู่กับ HTML (Hypertext Markup Language) ที่เป็นภาษาหลักในการสร้างโครงสร้างของเว็บไซค์ CSS มีบทบาทสำคัญในการแยกความรับผิดชอบระหว่างโครงสร้างเนื้อหาและการนำเสนอ ทำให้เว็บไซค์มีความสวยงามและใช้งานง่ายขึ้น โดย CSS ย่อมาจากคำว่า “Cascading Style Sheets” หรือ “สไตลชีตแบบซ้อนทับ” คือ ภาษาที่ใช้กำหนดรูปแบบการแสดงผลของ HTML โดยใช้หลักการของการกำหนดกฎเกณฑ์หรือ “กฎของ CSS” (CSS Rules) เพื่อควบคุมลักษณะการแสดงผลขององค์ประกอบต่างๆ ในหน้าเว็บ เช่น ขนาดตัวอักษร สีพื้นหลัง ขอบของกล่อง และการจัดวางองค์ประกอบ

2.2.5.1 ประโยชน์ของ CSS

1) การแยกเนื้อหาและการออกแบบ หนึ่งในประโยชน์หลักของ CSS คือการแยกโครงสร้างเนื้อหา (HTML) ออกจากการออกแบบ (CSS) ซึ่งทำให้การปรับแต่งรูปแบบของเว็บไซค์ทำได้ง่ายขึ้น และสามารถนำสไตลชีตไปใช้กับหลายๆ หน้าได้โดยไม่ต้องแก้ไขโค้ด HTML หลายครั้ง

2) ความสะดวกในการบำรุงรักษา CSS ช่วยให้การบำรุงรักษาเว็บไซค์ทำได้ง่ายขึ้น หากต้องการปรับเปลี่ยนรูปแบบของเว็บไซค์ เพียงแค่แก้ไขไฟล์ CSS ไฟล์เดียว การเปลี่ยนแปลงจะส่งผลกับทุกหน้าเว็บที่ใช้ไฟล์นั้นทันที

3) เพิ่มประสิทธิภาพในการโหลดหน้าเว็บ CSS ช่วยลดขนาดของโค้ด HTML โดยการแยกการจัดรูปแบบออกมาเป็นไฟล์แยก ทำให้หน้าเว็บมีขนาดเล็กลงและโหลดได้เร็วขึ้น ซึ่งเป็นปัจจัยสำคัญที่ช่วยเพิ่มประสิทธิภาพในการทำ SEO

4) การออกแบบที่ยืดหยุ่นและหลากหลาย CSS ทำให้สามารถออกแบบเว็บไซต์ให้มีความยืดหยุ่นและหลากหลายมากขึ้น สามารถปรับเปลี่ยนรูปแบบการแสดงผลของเว็บไซต์ได้ตามต้องการ เช่น การแสดงผลที่แตกต่างกันระหว่างหน้าจอคอมพิวเตอร์และอุปกรณ์มือถือ

2.2.5.2 โครงสร้างของ CSS

CSS ประกอบด้วยสามส่วนหลักคือ ตัวเลือก (Selector) คุณสมบัติ (Property) และค่าของคุณสมบัติ (Value)

- Selector เป็นส่วนที่ระบุว่าจะประกอบใดใน HTML ที่จะนำกฎ CSS ไปใช้
- Property เป็นคุณสมบัติที่ต้องการกำหนดให้กับองค์ประกอบ
- Value เป็นค่าที่กำหนดให้กับคุณสมบัติที่ระบุไว้

2.2.5.3 ประเภทของ CSS

CSS สามารถนำไปใช้ในเว็บไซต์ได้หลายวิธี โดยแต่ละวิธีมีข้อดีและข้อเสียที่แตกต่างกันไป

1) Inline CSS เป็นการเขียน CSS ในบรรทัดเดียวกับโค้ด HTML ใช้เมื่อจำเป็นต้องปรับเปลี่ยนรูปแบบขององค์ประกอบเฉพาะเจาะจง แต่ไม่เหมาะสำหรับการใช้ในเว็บไซต์ขนาดใหญ่เพราะยากต่อการแก้ไข

2) Internal CSS เป็นการเขียน CSS ภายในไฟล์ HTML โดยใช้แท็ก <style> ภายใน <head> ใช้ในกรณีที่ต้องการควบคุมรูปแบบในหน้าเดียว แต่หากมีหลายหน้าก็ยังคงต้องเขียน CSS ในแต่ละหน้า

3) External CSS เป็นวิธีที่แนะนำและใช้บ่อยที่สุด โดยเขียน CSS ในไฟล์แยกออกมา (.css) และเชื่อมโยงไฟล์นี้กับหน้า HTML โดยใช้แท็ก <link> ภายใน <head> ของ HTML เหมาะสำหรับการใช้ในเว็บไซต์ขนาดใหญ่ที่มีหลายหน้า

2.2.5.4 เทคนิคการใช้ CSS ในการออกแบบเว็บไซต์

1) Responsive Design CSS ช่วยในการทำให้เว็บไซต์มีการแสดงผลที่ดีบนทุกขนาดหน้าจอ ไม่ว่าจะเป็นคอมพิวเตอร์ แท็บเล็ต หรือมือถือ โดยใช้ media queries ที่ช่วยกำหนดกฎ CSS ที่แตกต่างกันตามขนาดหน้าจอ

2) การใช้ Flexbox และ Grid เป็นเครื่องมือที่ช่วยในการจัดวางองค์ประกอบในหน้าเว็บได้อย่างยืดหยุ่นและง่ายตายไม่ว่าจะเป็นการจัดเรียงกล่อง การแบ่งส่วนต่างๆ ของหน้าเว็บให้มีรูปแบบที่สวยงาม

3) การใช้ CSS Variables ช่วยในการจัดการค่าที่ใช้ซ้ำใน CSS โดยกำหนดค่าไว้ในตัวแปร ทำให้การแก้ไขค่าเป็นไปได้อย่างขึ้น (visperhost, 2567)

2.3 วรรณกรรมที่เกี่ยวข้อง

สุภาพวรรณ ดงมณีพรรณ (2562) ศึกษาการวิเคราะห์ข้อมูลการซื้อผลิตภัณฑ์เสริมอาหารแบรนด์ MYHEALTH ของลูกค้าในร้านยา โดยใช้กรอบกระบวนการ CRISP-DM และเทคนิคเหมืองข้อมูลหลายรูปแบบร่วมกัน เพื่อค้นหาความสัมพันธ์ของการซื้อสินค้าและแบ่งกลุ่มลูกค้าตามพฤติกรรมการซื้อ โดยงานวิจัยใช้ข้อมูลธุรกรรมจำนวน 93,278 แถวสำหรับการหาความสัมพันธ์ และข้อมูลลูกค้า 63,860 แถวสำหรับการจัดกลุ่มลูกค้า จากนั้นผู้วิจัยนำข้อมูลเข้าสู่กระบวนการทำเหมืองข้อมูล ได้แก่ การทำความเข้าใจธุรกิจ การทำความเข้าใจข้อมูล การเตรียมข้อมูล การสร้างแบบจำลอง การประเมินผล และการนำแบบจำลองไปใช้

เทคนิคที่ใช้ประกอบด้วย FP-Growth สำหรับการค้นหารูปแบบการซื้อสินค้าที่เกิดขึ้นร่วมกันบ่อยและสร้างกฎความสัมพันธ์ (Association Rules) โดยกำหนดค่า minimum confidence ไม่น้อยกว่า 70% และใช้ RFM Analysis เพื่อวิเคราะห์พฤติกรรมด้าน Recency, Frequency และ Monetary ของลูกค้า ก่อนแบ่งกลุ่มลูกค้าด้วยเทคนิค K-Means Clustering ซึ่งเลือกจำนวนกลุ่มที่เหมาะสมจาก Elbow Method พบว่าเหมาะสมที่สุดที่ 6 กลุ่ม ได้แก่ Blue Diamond, Diamond, Platinum, Gold, Silver และ Welcome

ผลการวิจัยพบกฎความสัมพันธ์รวม 7 กฎ เช่น เมื่อลูกค้าซื้อ L-Carnitine และ Royal Jelly จะมีโอกาสซื้อ Carbo-X ร่วมด้วยถึง 71.4% และกฎหลายชุดมีค่า Confidence สูงถึง 100% นอกจากนี้การแบ่งกลุ่ม RFM ทำให้ทราบลักษณะลูกค้าแต่ละกลุ่มอย่างชัดเจน เช่น กลุ่ม Blue Diamond มีระยะกลับมาซื้อเฉลี่ย 91 วัน และมีค่าใช้จ่ายสูงที่สุด โดยผลลัพธ์เหล่านี้สามารถนำไปใช้วางแผนโปรโมชัน การจัดตำแหน่งสินค้า การเพิ่มยอดขาย และการวางกลยุทธ์ทางการตลาดอย่างมีประสิทธิภาพ เช่น การจัดเซตสินค้า การเสนอขายพ่วง รวมถึงการวางแผน CRM เพื่อรักษาลูกค้าคุณภาพสูงของธุรกิจ

ปัทมากรีย์ สุนทรวราภาส และคณะ (2563) มุ่งวิเคราะห์พฤติกรรมผู้ใช้บริการห้องสมุดของมหาวิทยาลัยสงขลานครินทร์ โดยใช้เทคนิค RFM (Recency, Frequency, Monetary) เพื่อ

จัดกลุ่มผู้ใช้บริการตามรูปแบบการเยี่ยมชมพยากรสารสนเทศ โดยวิเคราะห์ข้อมูลจากระบบห้องสมุดอัตโนมัติ (ALIST) จำนวน 43,077 รายการ จากผู้ใช้ 6,155 คน ตลอดปี 2562 ด้วยโปรแกรม RapidMiner

จากการศึกษานี้ ผู้วิจัยได้ทำการคำนวณค่า RFM ให้แก่ผู้ใช้แต่ละราย จากนั้นกำหนดระดับคะแนน R, F, M แยกเป็น 3 ระดับเท่ากันตามเปอร์เซ็นต์ และรวมรหัสเป็นกลุ่ม เช่น RFM133 หรือ RFM322 แล้วทำการวิเคราะห์จัดกลุ่มพฤติกรรม พบว่าแบ่งได้ทั้งหมด 18 กลุ่ม โดยกลุ่ม RFM133 มีผู้ใช้บริการมากที่สุด (17.55%) คือกลุ่มที่มาใช้บริการล่าสุดไม่นาน มีความถี่และปริมาณการเยี่ยมชมสูง และกลุ่ม RFM322 และ RFM222 ตามลำดับ

ผลการวิเคราะห์เชิงพฤติกรรมพบว่า บริการยอดนิยมอันดับหนึ่งในเกือบทุกกลุ่มคือการยืมหนังสือ (มากกว่า 40%) รองลงมาคือหูฟัง (มากกว่า 10%) โดยเฉพาะกลุ่ม RFM133 ที่ยืมหนังสือถึง 14,441 รายการ และหูฟัง 4,341 รายการ สะท้อนพฤติกรรมการใช้ห้องสมุดที่ยังคงให้ความสำคัญกับทรัพยากรทางกายภาพ

จากการแบ่งกลุ่มนี้ ผู้วิจัยเสนอว่าห้องสมุดสามารถใช้ข้อมูล RFM ไปวางแผนการจัดซื้อทรัพยากร เช่น จัดซื้อหนังสือในกลุ่มยอดนิยมอย่าง RFM133, 322, 222, 311, 122 และ 233 เพื่อใช้งบประมาณอย่างคุ้มค่า และจัดกิจกรรมส่งเสริมการใช้บริการโดยใช้ลักษณะการใช้งานร่วม เช่น วางโปสเตอร์วิธีใหม่ใกล้จุดคืนหูฟัง เพื่อกระตุ้นให้ผู้ใช้กลับมาใช้งานห้องสมุด

ทูซาร์ กานซาล และคณะ (2563) ได้ทำการศึกษาวิจัยเกี่ยวกับการแบ่งกลุ่มลูกค้าด้วยเทคนิคการจัดกลุ่มแบบไม่มีผู้สอน (Unsupervised Learning) โดยใช้ 3 อัลกอริทึมหลัก ได้แก่ K-Means, Agglomerative และ Mean Shift เพื่อทำการวิเคราะห์ความคล้ายคลึงกันของพฤติกรรมลูกค้า จากนั้นเปรียบเทียบประสิทธิภาพของแต่ละอัลกอริทึมโดยใช้ Internal Clustering Validation เป็นตัวชี้วัดหลักเนื่องจากชุดข้อมูลที่ใช้เป็นข้อมูลที่ไม่มีป้ายกำกับ (Unlabeled Data) ในส่วนของกระบวนการวิจัยได้เริ่มจากการเตรียมชุดข้อมูลลูกค้าจำนวนหนึ่งแล้วใช้แต่ละอัลกอริทึมทำการจัดกลุ่มเพื่อเปรียบเทียบจำนวนคลัสเตอร์และคุณภาพของแต่ละคลัสเตอร์ที่ได้

ผลการวิจัยพบว่าอัลกอริทึม K-Means สามารถจัดกลุ่มลูกค้าได้ 5 กลุ่ม ได้แก่ Careless, Careful, Standard, Target และ Sensible ส่วนอัลกอริทึม Mean Shift สามารถค้นพบกลุ่มเพิ่มเติมอีก 2 กลุ่ม คือ High Buyers and Frequent Visitors และ High Buyers and Occasional Visitors ซึ่งเป็นกลุ่มลูกค้าที่มีพฤติกรรมซื้อสินค้าที่โดดเด่นกว่ากลุ่มอื่นๆ ทั้งนี้งานวิจัยยังระบุว่าการใช้ค่าชี้วัดแบบ Internal Validation เช่น Silhouette หรือ Davies-Bouldin สามารถช่วยให้

นักวิเคราะห์เลือกอัลกอริทึมที่เหมาะสมได้อย่างแม่นยำและเพิ่มความถูกต้องของการจัดกลุ่มข้อมูล พร้อมทั้งเสนอแนะว่างานวิจัยในอนาคตสามารถเพิ่มกระบวนการ Feature Extraction เพื่อช่วยลดข้อมูลที่ซ้ำซ้อนและเพิ่มประสิทธิภาพของโมเดลได้มากยิ่งขึ้น

พิมพ์พนิต สิทธิทวี และธันยานี โพธิสาร (2563) ศึกษาการใช้เทคนิคเหมืองข้อมูลเพื่อวิเคราะห์พฤติกรรมของผู้บริโภคที่ซื้อผลิตภัณฑ์เสริมอาหารคู่แข่งในร้านขายยา Pharmacy Chain AA โดยใช้กระบวนการ CRISP-DM ประกอบด้วยขั้นตอนความเข้าใจธุรกิจ ความเข้าใจข้อมูล การเตรียมข้อมูล การสร้างแบบจำลอง การประเมินผล และการวางแผนนำไปใช้ โดยอาศัยข้อมูลการขายจริงจาก 3 สาขา

ประยุกต์เทคนิค FP-Growth ในการค้นหาความสัมพันธ์ของการซื้อสินค้า รวมถึงใช้เทคนิค RFM Analysis เพื่อจัดกลุ่มลูกค้าตามพฤติกรรมการซื้อ

ผลการวิเคราะห์พบว่ากฎความสัมพันธ์ที่เกิดขึ้นในหลายสาขามีรูปแบบที่คล้ายคลึงกัน เช่น ลูกค้าที่ซื้อ Fish Oil มักซื้อร่วมกับ Magnesium, Vitamin B, Bilberry หรือ Calcium ซึ่งช่วยบ่งชี้การซื้อแบบเป็นชุดของผลิตภัณฑ์เสริมอาหาร นอกจากนี้ผลการจัดกลุ่มลูกค้าด้วย RFM สามารถแบ่งได้เป็น 3 กลุ่ม ได้แก่ Gold, Silver และ Bronze โดยกลุ่ม Gold มีค่า Recency ต่ำ ความถี่สูง และยอดใช้จ่ายสูงที่สุด แสดงถึงความเป็นลูกค้าคุณภาพที่ควรได้รับการดูแลเป็นพิเศษ งานวิจัยสรุปว่าแบบจำลองดังกล่าวสามารถช่วยสนับสนุนการวางแผนโปรโมชัน การจัดกลุ่มลูกค้า และการผลักดันยอดขายผลิตภัณฑ์ให้มีประสิทธิภาพมากขึ้น รวมถึงเพิ่มความสามารถในการแข่งขันด้านการตลาดของร้านยาได้เป็นอย่างดี

ชิตนภัส มาศชาย และคณะ (2564) มุ่งวิเคราะห์หาความสัมพันธ์ของการประเมินสมรรถภาพทางกายของนักเรียนโรงเรียนราชประชานุเคราะห์ 1 โดยใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) เพื่อหารูปแบบความสัมพันธ์ของรายการทดสอบสมรรถภาพในแต่ละด้าน โดยวิเคราะห์ข้อมูลจากกลุ่มตัวอย่างนักเรียนจำนวน 812 คน ในปีการศึกษา 2563 ด้วยโปรแกรม Weka 3.8.5

จากการศึกษาครั้งนี้ ผู้วิจัยได้ใช้เทคนิคหลัก 2 รูปแบบ คือ การจำแนกประเภทข้อมูล (Classification) ด้วยวิธีต้นไม้ตัดสินใจ (Decision Tree) แบบ J48 และป่าสุ่ม (Random Forest) และการหาความสัมพันธ์ (Association Rule) ด้วยขั้นตอนวิธีแบบ Apriori และ FP-Growth โดยกำหนดรายการทดสอบ 5 ด้าน ได้แก่ ดัชนีมวลกาย (BMI), การยืนยกเข่าขึ้นลง, การลุก-นั่ง,

การค้นพบบ่อยที่สุด และการนั่งอตัวไปข้างหน้า ซึ่งแบ่งระดับผลการทดสอบออกเป็น 5 ระดับ ตั้งแต่ระดับต่ำมากไปจนถึงระดับดีมาก

ผลการวิเคราะห์เชิงพฤติกรรมพบว่าเทคนิคการจำแนกประเภทข้อมูล: วิธีต้นไม้ตัดสินใจ (Decision Tree) แบบ J48 ร่วมกับการประเมินแบบ Percentage Split 70% มีประสิทธิภาพสูงที่สุด โดยให้ค่าความแม่นยำ (Accuracy), ค่าความไว (Recall) และค่าวัดประสิทธิภาพโดยรวม (F-measure) สูงถึงร้อยละ 100 เทคนิคกฎความสัมพันธ์: จากการกำหนดค่าความเชื่อมั่น (Confidence) ที่ 0.9 และค่าสนับสนุนขั้นต่ำ (Support) ที่ 0.1 พบกฎความสัมพันธ์จำนวน 5 ข้อ จาก Apriori และ 4 ข้อจาก FP-Growth โดยกฎที่มีค่าความเชื่อมั่นสูงสุด (93%) ระบุว่า หากนักเรียนมีดัชนีมวลกาย "สมส่วน" และมีผลสรุปสมรรถภาพทางกาย "ดี" จะมีโอกาสที่ผลการยืนยกเข้าชั้นลง 3 นาทีอยู่ในระดับ "ดีมาก"

จากการวิเคราะห์นี้ ผู้วิจัยเสนอว่าผลลัพธ์ที่ได้จากการจำแนกประเภทข้อมูลและกฎความสัมพันธ์ให้ผลลัพธ์ที่คล้ายคลึงกันและไปในทิศทางเดียวกัน ซึ่งทางโรงเรียนสามารถนำกฎความสัมพันธ์ที่ค้นพบไปประยุกต์ใช้ในการทดสอบสมรรถภาพ และเป็นแนวทางในการพัฒนารูปแบบการออกกำลังกายที่เหมาะสมเพื่อส่งเสริมสุขภาพของนักเรียนในแต่ละกลุ่มได้อย่างมีประสิทธิภาพต่อไป

ชัยโรจน์ หล่อไพบูลย์ (2565) ทำการศึกษาการวิเคราะห์พฤติกรรมลูกค้าในธุรกิจก่อสร้างของประเทศไทยโดยใช้เทคนิคเหมืองข้อมูลหลายรูปแบบร่วมกัน เพื่อสนับสนุนการวางกลยุทธ์ทางการตลาดและการบริหารความสัมพันธ์ลูกค้า งานวิจัยนี้เริ่มจากการนำข้อมูลธุรกรรมการซื้อสินค้ามาทำการคำนวณตัวแปร RFM (Recency, Frequency, Monetary) เพื่อใช้เป็นตัวบ่งชี้พฤติกรรมการซื้อของลูกค้าในเชิงเวลา ความถี่ และมูลค่า ก่อนนำข้อมูลที่ผ่านการเตรียมความพร้อมแล้วไปเข้าสู่ขั้นตอน การจัดกลุ่ม (Clustering) โดยใช้วิธี K-Means Clustering เพื่อแบ่งลูกค้าออกเป็นหลายกลุ่มตามลักษณะพฤติกรรม ผลการวิเคราะห์พบว่าการใช้จำนวนกลุ่มที่เหมาะสมช่วยให้เห็นรูปแบบของลูกค้าแต่ละประเภทได้ชัดเจนขึ้น เช่น กลุ่มลูกค้าที่ซื้อบ่อยแต่มูลค่าต่ำ และกลุ่มที่ซื้อมูลค่าสูงแต่ความถี่ไม่มาก เป็นต้น

หลังจากจัดกลุ่มแล้ว ผู้วิจัยได้นำผลลัพธ์จากการจัดกลุ่มลูกค้าไปวิเคราะห์ต่อยอดด้วยเทคนิค Hybrid-Dimension Association Rules ซึ่งเป็นการขยายขีดความสามารถของกฎความสัมพันธ์ (Association Rules) แบบปกติ ให้สามารถหาความสัมพันธ์เชิงลึกระหว่าง “พฤติกรรมซื้อ” และ “ประเภทสินค้า” ได้พร้อมกัน ทำให้สามารถตรวจพบรูปแบบการซื้อที่เกิดขึ้นบ่อยในแต่ละ

กลุ่มลูกค้าได้ละเอียดกว่าเดิม เช่น การซื้อสินค้าประเภทหนึ่งมักจะร่วมกับสินค้าประเภทอื่นในกลุ่มลูกค้าบางกลุ่มโดยเฉพาะ

ผลการศึกษาพบว่า การผสมเทคนิคทั้ง 3 ได้แก่ RFM + K-Means + Hybrid Association Rules ช่วยให้สามารถระบุลักษณะเฉพาะของลูกค้าแต่ละกลุ่มได้อย่างแม่นยำ และช่วยในการคาดการณ์โอกาสการซื้อสินค้าชนิดต่าง ๆ ของลูกค้า ทำให้ธุรกิจสามารถนำผลลัพธ์ไปประยุกต์ใช้เพื่อออกแบบโปรโมชั่น วางแผนการตลาด และจัดการสต็อกสินค้าได้อย่างมีประสิทธิภาพมากยิ่งขึ้น ข้อค้นพบสำคัญคือความสัมพันธ์ของรูปแบบการซื้อที่มีความแตกต่างกันชัดเจนในแต่ละกลุ่มลูกค้า ซึ่งเป็นประโยชน์อย่างยิ่งสำหรับการวางแผนกลยุทธ์ CRM และการเพิ่มมูลค่าทางธุรกิจในระยะยาว

ปริญทร์ กิตติโสภณธรรม และกวีศิลป์ คักดีวรพงศ์ (2565) ศึกษาการสร้างโอกาสในการเสนอขายสินค้าด้วยเทคนิคเหมืองข้อมูล โดยมุ่งวิเคราะห์พฤติกรรมกรรมการซื้อสินค้าร่วมกันของลูกค้าร้านกาแฟ อเมซอน สาขามหาวิทยาลัยสยาม จากข้อมูลผู้ใช้บริการจำนวน 535 คน และผ่านการคัดเลือกข้อมูลจนเหลือ 315 รายการเพื่อนำมาวิเคราะห์เชิงลึก กระบวนการศึกษาดำเนินตามขั้นตอน CRISP-DM ครอบคลุมการทำความเข้าใจธุรกิจ การทำความเข้าใจข้อมูล การเตรียมข้อมูลด้วยการทำความสะอาด การคัดเลือกตัวแปร และการแปลงข้อมูลด้วย Numerical to Binominal เพื่อให้เหมาะสมกับอัลกอริทึม

จากนั้นจึงดำเนินการสร้างแบบจำลองด้วยเทคนิค Association Rules โดยใช้อัลกอริทึม FP-Growth ผ่านโปรแกรม RapidMiner กำหนดค่า Minimum Support = 0.50 และ Confidence = 0.69-1.00 เพื่อค้นหากฎความสัมพันธ์ที่เกิดขึ้นบ่อย

ผลการวิจัยพบกฎความสัมพันธ์จำนวน 3 กฎ ได้แก่ (1) เอสเปรสโซ่ร้อน → เอสเปรสโซ่เย็น (Confidence 81%) (2) แบล็คคอฟฟี่เย็น → เอสเปรสโซ่เย็น (Confidence 73%) และ (3) ชานมเย็น → ชาเขียวเย็น (Confidence 71%) ซึ่งสะท้อนถึงพฤติกรรมกรรมการซื้อคู่กันของลูกค้า และสามารถนำไปใช้วางแผนโปรโมชั่นแบบแพ็คเกจเพื่อเพิ่มยอดขายในร้านได้อย่างมีประสิทธิภาพ นอกจากนี้ การประเมินผลด้วยการแบ่งข้อมูล Train/Test พบว่ากฎที่ได้มีเสถียรภาพและสามารถนำไปใช้จริงในการพยากรณ์พฤติกรรมกรรมการซื้อของลูกค้าได้อย่างเหมาะสม ส่งผลให้ผู้ประกอบการสามารถออกแบบกลยุทธ์การตลาดตรงตามความต้องการของลูกค้ากลุ่มเป้าหมายมากขึ้น

ณรรฐคุณ วิรุฬห์ศรี และคณะ (2565) ศึกษาการวิเคราะห์แบ่งกลุ่มลูกค้าของบริษัทผู้ผลิตอาหารสัตว์เลี้ยง โดยใช้ข้อมูลการซื้อขายตั้งแต่ปี ค.ศ. 2018–2020 รวมจำนวนลูกค้า 588 ราย เพื่อจำแนกลักษณะลูกค้าตามพฤติกรรมการซื้อที่แท้จริง โดยผู้วิจัยนำตัวแปรเชิงปริมาณทั้งหมด 15 ตัว เช่น จำนวนครั้งที่ซื้อ-คืนสินค้า อัตราการเติบโตยอดซื้อ ความสม่ำเสมอรายเดือน ระยะทางจากศูนย์กระจายสินค้า และมูลค่าการซื้อ มาวิเคราะห์ด้วยเทคนิค K-Means Clustering ซึ่งเป็นวิธีการแบ่งกลุ่มแบบไม่เป็นขั้นตอน พร้อมตรวจสอบความเหมาะสมของจำนวนกลุ่มด้วย Elbow Method และใช้ดัชนี AD และ FOM เพื่อตรวจสอบเสถียรภาพของผลการแบ่งกลุ่ม

ผลการวิเคราะห์พบว่า จำนวนกลุ่มที่เหมาะสมคือ 8 กลุ่ม โดยลักษณะเด่นของแต่ละกลุ่มแตกต่างกัน เช่น กลุ่มลูกค้าซื้อมาก-ซื้อถี่ กลุ่มลูกค้าคืนสินค้าสูง กลุ่มลูกค้าทั่วไปคืนสินค้าน้อย กลุ่มลูกค้ากระเป่าหนัก และกลุ่มลูกค้าห่างไกล เป็นต้น ซึ่งช่วยให้บริษัทสามารถเข้าใจพฤติกรรมการซื้อของแต่ละกลุ่มอย่างชัดเจน และสามารถนำไปใช้ในการกำหนดกลยุทธ์การขาย การส่งเสริมการตลาด การดูแลลูกค้ารายสำคัญ รวมถึงการวางแผนขยายศูนย์กระจายสินค้า เพื่อเพิ่มประสิทธิภาพในการให้บริการในอนาคต

สิทธิพันธ์ สำเร็จประสงค์ (2566) ทำการศึกษาการแบ่งกลุ่มลูกค้าในธุรกิจค้าปลีกอิเล็กทรอนิกส์สินค้าสำหรับสตรีมีครรภ์ โดยใช้ข้อมูลธุรกรรมของร้าน MamyAnn จำนวน 12,641 รายการ ซึ่งเป็นข้อมูลจากช่องทางออนไลน์หลายแพลตฟอร์ม งานวิจัยดำเนินการตามกระบวนการ CRISP-DM ครบทั้ง 6 ขั้นตอน ได้แก่ การทำความเข้าใจธุรกิจ การทำความเข้าใจข้อมูล การเตรียมข้อมูล การสร้างแบบจำลอง การประเมินประสิทธิภาพ และการนำแบบจำลองไปใช้งาน

โดยใช้ตัวแปร RFM (Recency, Frequency, Monetary) เพื่อวัดพฤติกรรมการซื้อของลูกค้า และนำไปจัดกลุ่มด้วยอัลกอริทึม K-Means Clustering พร้อมประเมินความเหมาะสมของจำนวนคลัสเตอร์ด้วยเทคนิค Elbow Method, Sum Square Error (SSE) และ Silhouette Score

ผลการวิจัยสามารถแบ่งลูกค้าออกเป็น 7 กลุ่มที่มีลักษณะแตกต่างกันอย่างชัดเจน โดยกลุ่มที่มีคุณค่าต่อธุรกิจที่สุดจะมีค่า Recency ต่ำ ความถี่ในการซื้อสูง และมูลค่าการซื้อสูง แสดงถึงลูกค้าชั้นดีของร้าน ค่า SSE จากโมเดลลดลงเหลือ 12.098 และค่า Silhouette Score เท่ากับ 0.405 ซึ่งอยู่ในระดับที่แสดงว่าการจัดกลุ่มมีความเหมาะสม ผลลัพธ์การแบ่งกลุ่มนี้สามารถนำไปใช้สนับสนุนการวางแผนกลยุทธ์การตลาดเฉพาะกลุ่ม การจัดโปรโมชั่น และการ

ส่งเสริมการขาย เพื่อรักษาลูกค้าคุณภาพสูงและเพิ่มศักยภาพการสร้างรายได้ให้กับร้านค้าได้อย่างมีประสิทธิภาพ

กมลทิพย์ มนตรีสา และเรืองศักดิ์ ตระกูลพุทธิรักษ์ (2567) ได้ศึกษาวิธีการแบ่งกลุ่มลูกค้าตามพฤติกรรมการซื้อสินค้าของร้านค้าปลีก โดยใช้ข้อมูลสาธารณะจากเว็บไซต์ Kaggle ตั้งแต่ปี ค.ศ. 2019 ถึงต้นปี ค.ศ. 2023 รวมทั้งสิ้น 29,103 รายการ เพื่อวิเคราะห์ลักษณะการซื้อของลูกค้าและจัดแบ่งกลุ่มที่มีความคล้ายคลึงกัน โดยนำกระบวนการวิเคราะห์ข้อมูลตามขั้นตอน CRISP-DM เริ่มจากการทำความเข้าใจปัญหา การทำความเข้าใจข้อมูล และการเตรียมข้อมูล ซึ่งมีการทำความสะอาดข้อมูล การจัดการ Missing Value การลบข้อมูลผิดปกติ และการสร้างตัวแปรใหม่ที่สำคัญ เช่น Recency, Frequency, Monetary ร่วมกับตัวแปร Basket Size และ Day Type เพื่อสร้าง Profile ลูกค้า ก่อนนำมาสร้างแบบจำลองการจัดกลุ่มด้วยเทคนิค K-Means Clustering ทั้ง 2 รูปแบบ ได้แก่ แบบจำลองที่ 1 ใช้ข้อมูล RFM และแบบจำลองที่ 2 ใช้ข้อมูล RFMBD โดยใช้วิธี Elbow Method, Silhouette Score และ Davies-Bouldin Index ในการหาจำนวนกลุ่มที่เหมาะสม

ผลการทดลองแบบจำลอง RFM พบว่าการแบ่งกลุ่มที่เหมาะสมที่สุดคือ 2 กลุ่ม โดยได้กลุ่มลูกค้าที่มีลักษณะ “ซื้อเยอะ จ่ายหนัก” และ “ซื้อน้อย คิดนาน” ขณะที่แบบจำลอง RFMBD พบจำนวนกลุ่มที่เหมาะสมคือ 4 กลุ่ม ได้แก่ “กลุ่มลูกค้ามาน้อย จ่ายน้อย ซื้อวันธรรมดา”, “กลุ่มลูกค้ามาบ่อย จ่ายหนัก ซื้อวันธรรมดา”, “กลุ่มลูกค้ามาบ่อย จ่ายหนัก ซื้อวันหยุด” และ “กลุ่มลูกค้ามาน้อย จ่ายน้อย ซื้อวันหยุด” ซึ่งสะท้อนให้เห็นถึงความแตกต่างทางพฤติกรรมการซื้อที่ชัดเจนมากขึ้นเมื่อเพิ่มตัวแปร B และ D เข้ามา ส่งผลให้การจำแนกประเภทลูกค้าแม่นยำขึ้นและสามารถนำไปใช้วางกลยุทธ์ทางการตลาด เช่น การสร้างโปรโมชั่นเฉพาะกลุ่ม การรักษาลูกค้ามูลค่าสูง และการปรับปรุงประสบการณ์ลูกค้าในแต่ละกลุ่มได้อย่างมีประสิทธิภาพ

ชยพัทธ์ แดงทองดี (2567) ได้ทำการศึกษาการวิเคราะห์การแบ่งกลุ่มลูกค้าในอุตสาหกรรมเครื่องประดับ โดยมุ่งทำความเข้าใจพฤติกรรมการซื้อของลูกค้าผ่านข้อมูลเชิงธุรกรรมและข้อมูลคุณลักษณะของลูกค้า เพื่อนำผลลัพธ์ไปใช้กำหนดกลยุทธ์ทางการตลาดอย่างเหมาะสม โดยใช้กระบวนการวิเคราะห์ข้อมูลตามกรอบ CRISP-DM ครอบคลุมการทำ ความเข้าใจธุรกิจ การสำรวจและเตรียมข้อมูลด้วยการทำความสะอาด จัดรูปแบบข้อมูล และสร้างตัวแปรสำคัญ ได้แก่ ตัวแปร RFM (Recency, Frequency, Monetary) ซึ่งสะท้อน

พฤติกรรมการณ์ซื้อในช่วงเวลาล่าสุด ความถี่ในการซื้อ และมูลค่าที่ใช้จ่าย จากนั้นทำการปรับมาตรฐานข้อมูล (Normalization) และนำข้อมูลที่ผ่านมาเตรียมเข้าสู่กระบวนการจัดกลุ่มด้วยเทคนิค K-Means Clustering พร้อมประเมินจำนวนคลัสเตอร์ที่เหมาะสมโดยใช้ Elbow Curve, Silhouette Score และ Davies-Bouldin Index เพื่อให้ได้ผลการแบ่งกลุ่มที่มีความแตกต่างกันชัดเจนที่สุด

ผลการวิเคราะห์พบว่าลูกค้าสามารถแบ่งได้เป็นหลายกลุ่มตามลักษณะการซื้อ เช่น กลุ่มลูกค้ามูลค่าสูงที่ซื้อซ้ำบ่อย กลุ่มที่ซื้อน้อยแต่มีศักยภาพเติบโตสูง และกลุ่มลูกค้าที่มีแนวโน้มลดการซื้อ ซึ่งช่วยให้ธุรกิจสามารถออกแบบกลยุทธ์ทางการตลาดที่ตอบสนองแต่ละกลุ่มได้อย่างตรงจุด เช่น การจัดโปรโมชั่นเฉพาะกลุ่ม การบริหารความสัมพันธ์กับลูกค้ากลุ่มสำคัญ การรักษาลูกค้าที่มีคุณค่าสูง และการวางแผนเพิ่มยอดขายจากลูกค้าที่มีพฤติกรรมซื้อไม่สม่ำเสมอ ส่งผลให้การวางแผนเชิงกลยุทธ์ของธุรกิจระดับมีประสิทธิภาพมากยิ่งขึ้น และสามารถแข่งขันได้ดีขึ้นในระยะยาว

ฐิติพร ฐิติพรธรรม และศุภร คนธภักดี (2567) ศึกษาการจัดกลุ่มลูกค้าบัตรเครดิตของบริษัทแห่งหนึ่ง โดยมีวัตถุประสงค์เพื่อทำความเข้าใจพฤติกรรมลูกค้าและนำไปใช้ในการออกแบบแคมเปญการตลาดเฉพาะกลุ่ม งานวิจัยแบ่งการวิเคราะห์ออกเป็น 2 ระยะ ได้แก่ ระยะแรกใช้เทคนิค K-Means Clustering เพื่อจัดกลุ่มลูกค้าจากข้อมูลประชากรศาสตร์และพฤติกรรมการใช้บัตรเครดิต เช่น อายุ เพศ รายได้ วงเงินบัตรเครดิต ยอดใช้จ่าย และจำนวนครั้งในการทำธุรกรรม โดยมีการตรวจสอบความเหมาะสมของจำนวนกลุ่มด้วย Elbow Method, Silhouette Analysis, Davies-Bouldin Index และ Calinski-Harabasz Index ผลการวิเคราะห์พบว่าจำนวนกลุ่มที่เหมาะสมคือ 3 กลุ่ม ซึ่งมีลักษณะพฤติกรรมการใช้จ่ายแตกต่างกันอย่างชัดเจน จากนั้นในระยะที่สองได้นำผลการจัดกลุ่มมาใช้สร้างแบบจำลอง Decision Tree เพื่อทำนายกลุ่มลูกค้าและอธิบายลักษณะของแต่ละกลุ่ม โดยผลการประเมินประสิทธิภาพของแบบจำลองมีความแม่นยำสูง (Accuracy = 0.98) และสามารถระบุปัจจัยสำคัญที่ส่งผลต่อการจำแนกกลุ่มลูกค้าได้ เช่น วงเงินบัตรเครดิต เพศ รายได้สูงสุด และยอดใช้จ่ายรวม ผลการวิจัยช่วยให้ธุรกิจสามารถเข้าใจพฤติกรรมลูกค้าเชิงลึก และนำไปใช้วางแผนกลยุทธ์การตลาด โปรโมชั่น และการบริหารความสัมพันธ์กับลูกค้าได้อย่างมีประสิทธิภาพ

จากการทบทวนวรรณกรรมที่เกี่ยวข้องกับการวิเคราะห์ข้อมูลลูกค้า สามารถสรุปได้ว่าเทคนิคที่สำคัญและถูกนำมาใช้ในการศึกษาพฤติกรรมลูกค้า ได้แก่ Association Rule ซึ่งใช้ค้นหาความสัมพันธ์ของประเภทลูกค้าที่มาใช้บริการโดยใช้ Algorithm หลักอย่าง และ FP-

Growth เพื่อให้ได้กฎที่มีความแข็งแกร่งและน่าเชื่อถือตามตัวชี้วัด Support, Confidence, และ Lift; รวมไปถึง Decision Tree ที่ใช้ในการวิเคราะห์กลุ่มลูกค้า RFM ผ่านการใช้ต้นไม้ตัดสินใจ และ RFM Analysis เป็นวิธีการที่จำเป็นในการจำแนกลูกค้าตามมูลค่าและความภักดีโดยพิจารณาจาก Recency, Frequency, และ Monetary Value ในขณะที่ Clustering (K-Means) เป็นเทคนิคที่ใช้ในการจัดกลุ่มข้อมูลจากการวิเคราะห์ RFM ซึ่งผลลัพธ์จากเทคนิคเหล่านี้มีประโยชน์อย่างยิ่งในการกำหนดกลยุทธ์การจัดวางสินค้า, การทำนายพฤติกรรม, และการออกแบบแคมเปญโปรโมชันที่ตรงกลุ่มเป้าหมาย

2.4 บทสรุป

จากการศึกษาพบว่า แนวคิด ทฤษฎี และเครื่องมือที่เกี่ยวข้องกับการวิเคราะห์ข้อมูล เช่น Data Analytics, RFM Analysis, Decision Tree การจัดกลุ่มด้วย K-Means และ Machine Learning ล้วนมีความสำคัญต่อการทำความเข้าใจพฤติกรรมลูกค้าอย่างเป็นระบบ เครื่องมือเหล่านี้ช่วยให้สามารถจำแนกกลุ่มลูกค้า วิเคราะห์แนวโน้มการซื้อได้อย่างแม่นยำ นอกจากนี้ยังสนับสนุนการตัดสินใจทางธุรกิจและการกำหนดกลยุทธ์ทางการตลาดให้สอดคล้องกับความต้องการของลูกค้าได้ดียิ่งขึ้น