

บทที่ 2

แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงการเรื่องการวิเคราะห์ข้อมูลเพื่อสร้างเครื่องมือช่วยตัดสินใจสำหรับผู้สนใจเข้าศึกษาในคณะบริหารธุรกิจและศิลปศาสตร์มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ โดยใช้เทคนิคการวิเคราะห์ข้อมูล ในบทนี้เป็นกรนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้อง ของการวิเคราะห์ข้อมูลเพื่อสร้างเครื่องมือช่วยตัดสินใจสำหรับผู้สนใจเข้าศึกษาในคณะบริหารธุรกิจและศิลปศาสตร์มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง เพื่อใช้เป็นแนวทางการศึกษา ประกอบด้วย รายละเอียดตามลำดับ ดังนี้

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytics)

2.1.2 แนวคิดการใช้ข้อมูลเพื่อการตัดสินใจ (Data-Driven Decision Making)

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

2.1.4 แนวคิดการตรวจสอบคุณภาพเครื่องมือ (Validity และ Reliability)

2.2 ทฤษฎี

2.2.1 Big Five

2.2.2 การทำเหมืองข้อมูล (Data Mining)

2.2.3 หลักในการทำเหมืองข้อมูล

2.2.4 ทฤษฎีการกำหนดขนาดกลุ่มตัวอย่าง (Taro Yamane)

2.2.5 ทฤษฎีการวิเคราะห์ความแปรปรวน (ANOVA)

2.2.6 ทฤษฎีเกี่ยวข้องกับการสร้างเว็บไซต์

2.3 วรรณกรรมที่เกี่ยวข้อง

2.4 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.4.1 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

2.4.2 เทคนิคต้นไม้ตัดสินใจ (Decision Tree)

2.5 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytics) ในการดำเนินงานเรื่องการวิเคราะห์เพื่อสร้างเครื่องมือช่วยตัดสินใจสำหรับผู้สนใจเข้าศึกษาในคณะบริหารธุรกิจและศิลปศาสตร์มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ โดยใช้เทคนิคการทำเหมืองข้อมูล ผู้ศึกษาได้ศึกษาหลักการ และทฤษฎีต่าง ๆ ซึ่งองค์ประกอบที่สำคัญ คือ การวิเคราะห์ข้อมูล ซึ่งมีรายละเอียด ดังนี้

2.1.1.1 การวิเคราะห์ข้อมูล(Data Analytics) คือ การนำข้อมูลตั้งแต่อดีตจนถึงปัจจุบัน มาวิเคราะห์เพื่อคาดการณ์อนาคต และนำผลลัพธ์ไปใช้ในหลายด้าน ซึ่งครอบคลุมทุกอุตสาหกรรม ทั้งภาครัฐ และเอกชน โดยอาศัยเทคโนโลยี หรือซอฟต์แวร์เฉพาะด้าน เพื่อให้การประมวลผล และวิเคราะห์มีความแม่นยำรวมถึงสามารถนำไปใช้งานได้อย่างมีประสิทธิภาพ แบ่งออกเป็น 4 ประเภท ดังนี้

1) Diagnostic Analytics คือ การวิเคราะห์เชิงวินิจฉัย ซึ่งเป็นการวิเคราะห์ เชิงลึกที่ต่อยอดจาก Descriptive Analytics เพื่อตอบคำถามว่าเหตุการณ์นั้น ๆ จึงเกิดขึ้นโดยใช้ เทคนิคต่าง ๆ เช่น Data Discovery หรือ Data Mining เพื่อหาปัจจัย และตัวแปรที่ส่งผล ตัวอย่าง เช่น การวิเคราะห์สาเหตุที่ยอดขายเพิ่มขึ้น การเพิ่มขึ้นของผู้เข้าชมเว็บไซต์ การลดลงของจำนวน ลูกค้าที่เข้าร้าน โปรโมชั่นที่ไม่ค่อยได้รับความนิยม หรือเนื้อหาโฆษณาที่มีอัตราการคลิกสูงกว่า ปกติ นอกจากนี้ยังสามารถนำมาใช้วิเคราะห์การทำงานของคอมพิวเตอร์ เพื่อตรวจจับความ ผิดปกติ ข้อมูลที่ได้จะช่วยให้องค์กรเข้าใจความต้องการของตลาด พฤติกรรมของลูกค้า และสาเหตุของปัญหาทางเทคโนโลยี รวมถึงสามารถปรับปรุงวัฒนธรรมองค์กร เพื่อเพิ่มประสิทธิภาพในการทำงาน

2) การวิเคราะห์ข้อมูลในปัจจุบันเป็นสิ่งสำคัญอย่างยิ่งสำหรับองค์กร หรือธุรกิจในการดำเนินงานเชิงกลยุทธ์และตัดสินใจอย่างมีข้อมูล การวิเคราะห์เชิงพยากรณ์ (Predictive Analytics) และการวิเคราะห์เชิงคาดการณ์(Forecast Analytics) เป็นวิธีการวิเคราะห์ ข้อมูลที่มักถูกใช้งานเพื่อให้ข้อมูลเชิงลึกที่มีคุณค่าแก่ผู้บริหาร ซึ่งทั้งสองวิธีการ มีความแตกต่างกัน ในด้านของวัตถุประสงค์ และวิธีการใช้งาน ซึ่งสามารถแยกได้อย่างชัดเจนตามบริบทของการ ใช้งาน(P3 Adaptive, n.d.) Predictive Analytics คือ การวิเคราะห์ข้อมูลในอดีต และปัจจุบัน เพื่อตรวจสอบแนวโน้ม ทำนาย หรือพยากรณ์สิ่งที่จะเกิดขึ้นตามวัตถุประสงค์ที่กำหนด โดยใช้ แบบจำลองทางสถิติร่วมกับเทคโนโลยีปัญญาประดิษฐ์(AI) ซึ่งสามารถนำไปใช้ประโยชน์ 11 ได้หลายด้าน เช่น การคาดการณ์ความเสี่ยง และโอกาส ยอดขาย ภัยไซเบอร์ สภาพอากาศ การลงทุน หุ้น หรือผลการเลือกตั้ง ทั้งนี้ความแม่นยำของ Predictive Analytics

ขึ้นอยู่กับคุณภาพ ของข้อมูลที่องค์กรควรให้ความสำคัญเป็นอันดับแรก เพื่อให้ผลการวิเคราะห์มีประสิทธิภาพ ลดความผิดพลาด และสร้างประโยชน์ได้อย่างแท้จริง Forecast Analytics คือ การวิเคราะห์เชิงคาดการณ์ เป็นการวิเคราะห์เชิงพยากรณ์มุ่งเน้นไปที่การทำนายเหตุการณ์ ที่เฉพาะเจาะจง การวิเคราะห์เชิงคาดการณ์จะมุ่งเน้นไปที่การคาดการณ์แนวโน้มของเหตุการณ์ ที่เกิดขึ้นในระยะยาว โดยใช้ข้อมูลในอดีตเป็นฐานในการคาดการณ์อนาคต ซึ่งมักใช้กับตัวแปร ที่เป็นค่าเชิงเวลา(time-series data) เช่น การคาดการณ์ยอดขายรายปี หรือ การประเมินแนวโน้ม การใช้สินค้าคงคลัง(P3 Adaptive, n.d.)

3) Predictive Analytics และ Forecast Analytics จะเป็นการใช้ข้อมูลในอดีต ในการคาดการณ์อนาคต แต่ทั้งสองวิธีการมีวัตถุประสงค์และการใช้งานที่แตกต่างกันอย่างชัดเจน Predictive Analytics มักใช้ในการคาดการณ์พฤติกรรมของบุคคล หรือเหตุการณ์เฉพาะเจาะจง โดยเน้นการวิเคราะห์ในระดับจุลภาค(micro-level) ขณะที่ Forecast Analytics มุ่งเน้นไปที่แนวโน้ม ในอนาคตในระดับมหภาค(macro-level) และให้ภาพรวมที่กว้างขึ้นสำหรับการตัดสินใจ ในระยะยาว(One Model, n.d.)

4) Prescriptive Analytics คือ การวิเคราะห์เพื่อให้คำแนะนำ ซึ่งเป็นการวิเคราะห์ที่มีความซับซ้อนที่สุด โดยต่อเนื่องจากการทำ Predictive Analytics เมื่อได้ข้อมูลแนวโน้มที่จะเกิดขึ้น Prescriptive Analytics จะช่วยแนะนำแนวทางปฏิบัติที่เหมาะสมที่สุด รวมถึง วิเคราะห์ผลลัพธ์ที่อาจเกิดขึ้นจากการเลือกแนวทาง หรือแนะนำวิธีการแก้ไขปัญหา เครื่องมือนี้ จึงมีความสำคัญอย่างยิ่งในการตัดสินใจที่ขับเคลื่อนด้วยข้อมูล Prescriptive Analytics ใช้ Big Data อัลกอริทึม Machine Learning และเทคโนโลยี AI ในการวิเคราะห์ข้อมูลจำนวนมาก และซับซ้อน ซึ่งเกินขีดความสามารถของมนุษย์(Blendedata, 2024)

2.1.2 แนวคิดการใช้ข้อมูลเพื่อการตัดสินใจ(Data-Driven Decision Making)เป็นแนวคิดที่มุ่งเน้นให้การตัดสินใจอยู่บนพื้นฐานของข้อมูลเชิงประจักษ์ แทนการอาศัยเพียงประสบการณ์ สัญชาตญาณ หรือความคิดเห็นส่วนบุคคล แนวคิดนี้เกิดขึ้นจากการพัฒนาเทคโนโลยีสารสนเทศและฐานข้อมูลขนาดใหญ่(Big Data) ที่ทำให้องค์กรสามารถรวบรวม จัดเก็บ และประมวลผลข้อมูลในปริมาณมากได้อย่างมีประสิทธิภาพ ความสำคัญของ DDDM ดังนี้

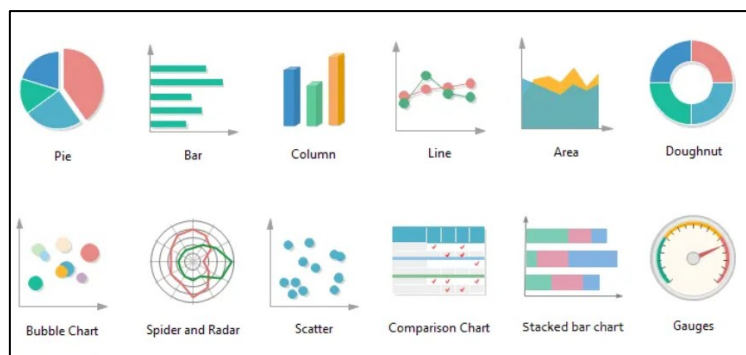
- 1) ความน่าเชื่อถือ(Reliability) – การตัดสินใจที่อ้างอิงจากข้อมูลจริงย่อมมีความถูกต้องและตรวจสอบย้อนกลับได้ ต่างจากการคาดเดาที่ไม่มีหลักฐานรองรับ
- 2) การลดความเสี่ยง(Risk Reduction) – การใช้ข้อมูลเชิงสถิติและการวิเคราะห์ ช่วยคาดการณ์แนวโน้มและผลลัพธ์ที่อาจเกิดขึ้น ทำให้สามารถป้องกันปัญหาล่วงหน้า

3) การเพิ่มประสิทธิภาพ(Efficiency Improvement) – การวิเคราะห์ข้อมูลช่วยให้ค้นหาทางเลือกที่เหมาะสมที่สุดได้เร็วขึ้น ลดเวลาและทรัพยากรที่สูญเปล่า

4) การสร้างคุณค่าเชิงกลยุทธ์(Strategic Value) – การตัดสินใจเชิงข้อมูลช่วยให้องค์กรสร้างความได้เปรียบในการแข่งขัน และสามารถปรับกลยุทธ์ตามพฤติกรรมหรือความต้องการของผู้ใช้ได้อย่างทันเวลา (จำกัด, 2568)

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization) Data Visualization เป็นการนำข้อมูลมาแสดงในรูปแบบของ ภาพกราฟิก แผนภูมิ กราฟ ไดอะแกรม หรืออื่น ๆ (Charts Graphs Diagrams หรือ Maps) เพื่อให้เราเห็น และเข้าใจลักษณะ แนวโน้ม และความสัมพันธ์ของข้อมูล (Patterns trends และ correlations) ได้ง่ายขึ้น ชัดเจนขึ้น เนื่องจากโดยทั่วไปแล้วข้อมูลที่เราสวมใจมักจะอยู่ในรูปแบบข้อความ อักขระ หรือตัวเลข ซึ่งก็มีอยู่ เป็นจำนวนมาก การศึกษาหรือการวิเคราะห์โดยพิจารณา หรือดูจากข้อมูลดังกล่าวตรง ๆ หรือดูจากตารางที่มีข้อมูลจำนวนมาก ๆ จะทำได้ค่อนข้างยาก แต่การใช้ Data Visualization จะช่วยให้การศึกษาและวิเคราะห์ทำได้ง่ายขึ้น เนื่องจากนอกจากจะเห็นภาพรวมของข้อมูลทั้งหมด สามารถสั่งให้แสดงเน้นข้อมูลบางส่วนที่สนใจ แยกสีแยกกลุ่ม ทำให้เข้าใจถึงลักษณะ แนวโน้ม การเปรียบเทียบข้อแตกต่าง และความสัมพันธ์ของข้อมูลได้ง่ายขึ้น ดังคำกล่าวที่ว่า "ภาพหนึ่งภาพแทนคำได้นับพันคำ" การทำ Data Visualization จะต้องพิจารณาถึงวัตถุประสงค์ ว่าต้องการค้นหา หรือต้องการดูข้อมูลสรุปอะไร เพื่อที่จะเลือกชนิดของกราฟที่เหมาะสม เช่น ต้องการดูจำนวน ความถี่ ใช้กราฟแท่ง Histogram ต้องการดูการกระจาย ต้องการดูความสัมพันธ์ หรือแนวโน้ม ของข้อมูล ใช้ Scatter และยังต้องคำนึงถึงด้วยว่า ข้อมูลที่นำมาพล็อตเป็นข้อมูลชนิดใด ชนิดหนึ่ง ต่อเนื่อง หรือแบบ Category ฯลฯ (สระอุบล, 2563)

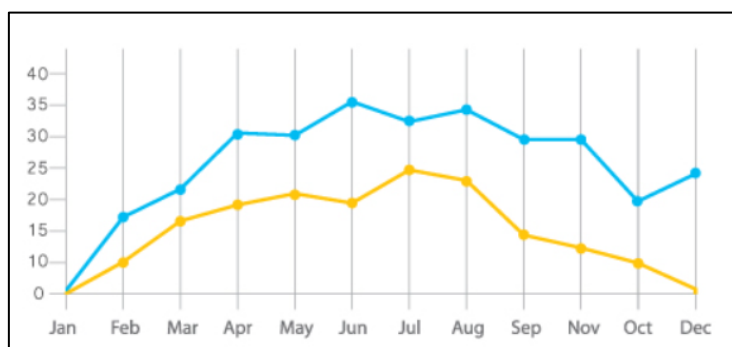
1) แผนภูมิ (Charts) เป็นรูปแบบที่นิยมมากที่สุด และเป็นรูปแบบที่มีหลากหลายชนิดที่เหมาะสมกับการนำเสนอข้อมูลที่แตกต่างกันไปตามวัตถุประสงค์ เช่น Pie Chart จะช่วยให้เราเห็นปริมาณความแตกต่างได้ชัดเจน Comparison Chart เหมาะสำหรับการเปรียบเทียบคุณสมบัติหลาย ๆ ข้อมูลมาตรวัด (Gauges) จะช่วยให้เห็นความเข้มข้น ความรุนแรง หรือน้ำหนัก



ภาพที่ 2.1 ตัวอย่างแผนภูมิรูปแบบต่าง ๆ

(ที่มา: medium.com/@Lynia_Li/)

2) กราฟ (Graphs) คือ Subset หรือประเภทหนึ่งของแผนภูมิ โดยกราฟจะทำหน้าที่แสดงความสัมพันธ์ระหว่างข้อมูล 2 ตัวแปร ผ่านแกนแนวนอน (แกน X) และแกนแนวตั้ง (แกน Y) ช่วยให้แนวโน้มสถานการณ์ประกอบกับบริบทได้เป็นอย่างดี



ภาพที่ 2.2 ตัวอย่างกราฟเส้นตัวอย่าง

(ที่มา: <https://www.def2design.com/Mkt-Corner/81>)

3) ตาราง (Tables) ก็เป็นอีกรูปแบบที่ใช้กันมากเพื่อนำเสนอข้อมูลให้ออกมา ดูง่าย ตารางประกอบไปด้วย 2 ส่วน ได้แก่ คอลัมน์ และแถว ซึ่งช่วยจัดการข้อมูลให้เรียบร้อย ช่วยให้มองเห็นบริบท และความสัมพันธ์ของข้อมูลหลาย ๆ ชุดได้อย่างง่าย

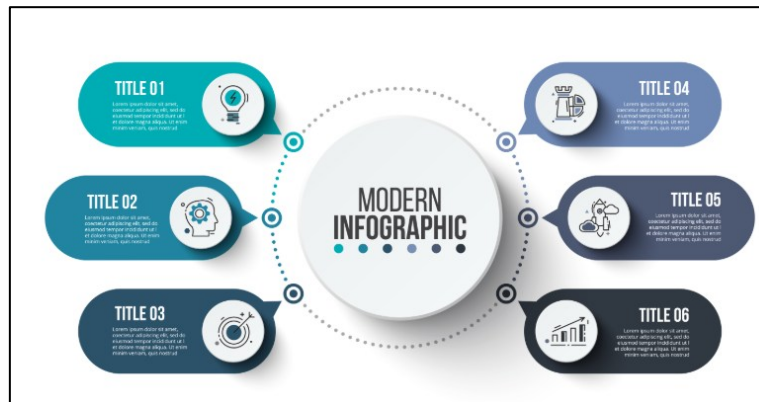
ม.ก. - ม.ย. 55				ม.ก. - ม.ย. 56		การเติบโต (%)
รายการ						
ตามบัญชี	อุตสาหกรรม	กลุ่ม	รวม	รวม		
ตัวแทน	4,717,862	408,911	147,896	5,274,669	4,779,573	10.36
ผู้บริหาร	3,551,398	41,807	24,578	8,862,582	2,842,459	27.28
รวม(ทุกรายการ)	8,248,163	450,718	163,700	8,862,581	7,582,236	16.89

หมายเหตุ : กำหนด 1 หน่วย : 1,000 บาท

ภาพที่ 2.3 ตัวอย่างตารางแสดงข้อมูล

(ที่มา: <https://ms1--q2--57.blogspot.com/p/week7.html>)

4) อินโฟกราฟิก (Infographic) คือ การนำเสนอสารสนเทศ (Info: information) ด้วยภาพกราฟิก (Graphic) เป็นรูปแบบการนำเสนอข้อมูลที่ใช้ภาพสื่อแทน ทำให้ผู้อ่านข้อมูลเข้าใจข้อมูลได้ง่ายหรือสามารถทำความเข้าใจผ่านภาพแทนที่คุ้นเคย นอกจากนี้ อินโฟกราฟิกยังเป็น 16 รูปแบบการนำเสนอข้อมูลที่น่าสนใจ มีการนำเทคนิคการเล่าเรื่อง (Storytelling) มาใช้ทำให้ข้อมูล น่าสนใจ น่าดึงดูด จึงมักจะใช้เพื่อนำเสนอเนื้อหา ความรู้ หรือเป็นสื่อการเรียนการสอน



ภาพที่ 2.4 ตัวอย่างอินโฟกราฟิก

(ที่มา: <https://www.datawow.co.th/blogs/what-is-data-visualization>)

5) แดชบอร์ด (Dashboards) คือ การนำข้อมูลต่าง ๆ มาเรียบเรียง และสรุป เป็นภาพ โดยใช้แผนภูมิและกราฟต่าง ๆ มาใช้นำเสนอ ปัจจุบันแดชบอร์ดเป็น Data Visualization ที่นิยมใช้กับการนำเสนอข้อมูลแบบ Real-time ผ่านซอฟต์แวร์ หรือเครื่องมือจัดการ และวิเคราะห์ ข้อมูลต่าง ๆ เช่น เครื่องมือการตลาด เครื่องมือบริหารจัดการข้อมูล เครื่องมือติดตาม และดูแล เว็บไซต์ ฯลฯ



ภาพที่ 2.5 ตัวอย่างแดชบอร์ด

(ที่มา: <https://www.datawow.co.th/blogs/what-is-data-visualization>)

2.1.4 แนวคิดการตรวจสอบคุณภาพเครื่องมือ (Validity และ Reliability) ในการศึกษาวิจัยเชิงปริมาณ (Quantitative Research) โดยเฉพาะการวิจัยที่เกี่ยวข้องกับพฤติกรรมศาสตร์และสังคมศาสตร์ที่ใช้แบบสอบถาม (Questionnaire) เป็นเครื่องมือหลักในการเก็บรวบรวมข้อมูลนั้น การตรวจสอบคุณภาพของเครื่องมือวิจัยถือเป็นขั้นตอนที่มีความสำคัญอย่างยิ่ง เพื่อเป็นหลักประกันว่าข้อมูลที่รวบรวมได้จะมีความถูกต้อง แม่นยำ และสามารถนำไปตอบวัตถุประสงค์ของการวิจัยได้อย่างแท้จริง การตรวจสอบคุณภาพเครื่องมือวิจัยประเภทแบบสอบถามที่ได้รับการยอมรับตามหลักสากล จะพิจารณาคุณภาพใน 2 มิติที่สำคัญ ได้แก่ ความตรง (Validity) และความเชื่อมั่น (Reliability) ซึ่งมีรายละเอียดดังต่อไปนี้

1) แนวคิดเกี่ยวกับความตรงของเครื่องมือ (Validity) ความตรง หรือ ความเที่ยงตรง หมายถึง ความสามารถและระดับความแม่นยำของเครื่องมือในการวัดสิ่งที่ผู้วิจัยต้องการจะวัดได้อย่างถูกต้องและครอบคลุมตามสภาพความเป็นจริง เครื่องมือที่ดีจะต้องสามารถเก็บรวบรวมข้อมูลได้ตรงตามวัตถุประสงค์ของการวิจัย ไม่ใช่การวัดสิ่งอื่นที่คลาดเคลื่อนไปจากกรอบแนวคิด โดยทั่วไปในการสร้างแบบสอบถาม มักจะพิจารณาความตรงใน 2 ลักษณะหลัก ได้แก่ ความตรงตามเนื้อหา (Content Validity) เป็นการพิจารณาว่าข้อคำถามที่สร้างขึ้นมีความครอบคลุมและสอดคล้องกับขอบเขตเนื้อหาหรือวัตถุประสงค์ที่ต้องการศึกษาหรือไม่ ซึ่งมักจะดำเนินการโดยการให้ผู้เชี่ยวชาญ (Experts) ในสาขาวิชาที่เกี่ยวข้อง จำนวนไม่น้อยกว่า 3 คน เป็นผู้ตรวจสอบและพิจารณาให้คะแนนความสอดคล้องของข้อคำถามแต่ละข้อ จากนั้นจึงนำมาคำนวณหาค่าดัชนีความสอดคล้อง (Index of Item-Objective Congruence: IOC) โดยข้อคำถามที่มีคุณภาพและสามารถนำไปจัดเก็บข้อมูลจริงได้นั้น ควรมีค่า IOC อยู่ในระดับตั้งแต่ 0.50 ไปจนถึง 1.00 ความตรงตามโครงสร้าง (Construct Validity) เป็นการประเมินว่าเครื่องมือ

นั้นสามารถวัดตัวแปรได้ตรงตามทฤษฎีหรือแนวคิดที่เป็นรากฐานของการศึกษานั้นหรือไม่ โดยเฉพาะอย่างยิ่งในการวัดตัวแปรที่เป็นคุณลักษณะแฝง (Latent Variables) หรือตัวแปรที่เป็นนามธรรม เช่น ทักษะคิด ความเชื่อ หรือบุคลิกภาพ ซึ่งไม่สามารถสังเกตเห็นได้โดยตรงทางกายภาพ การยืนยันความตรงตามโครงสร้างจะช่วยรับประกันว่าข้อคำถามเหล่านั้นสามารถสะท้อนพฤติกรรมตามทฤษฎีได้อย่างแท้จริง

2) แนวคิดเกี่ยวกับความเชื่อมั่นของเครื่องมือ (Reliability) ความเชื่อมั่น หรือ ความเที่ยง หมายถึง ความคงเส้นคงวา (Consistency) หรือความมีเสถียรภาพของเครื่องมือวัด ซึ่งแสดงให้เห็นว่า ไม่ว่าจะนำเครื่องมือนั้นไปใช้วัดซ้ำกับกลุ่มตัวอย่างเดิมอีกกี่ครั้ง ผลลัพธ์ที่ได้จากการวัดจะต้องมีความสอดคล้อง หรือออกมาในทิศทางเดียวกันหรือใกล้เคียงกับผลการวัดในครั้งแรก ในการวัดตัวแปรทางสังคมศาสตร์ที่ใช้มาตรวัดแบบประเมินค่า (Rating Scale) เช่น มาตรวัดแบบลิเคิร์ต (Likert Scale) การประเมินความเชื่อมั่นจะนิยมใช้วิธีการตรวจสอบความสอดคล้องภายใน (Internal Consistency) ผ่านการคำนวณหาค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Alpha Coefficient) โดยมีหลักเกณฑ์ที่ได้รับการยอมรับในระดับสากลว่า เครื่องมือแบบสอบถามที่มีคุณภาพดีและมีความน่าเชื่อถือสูง ควรมีค่าสัมประสิทธิ์แอลฟาตั้งแต่ 0.70 หรือ 0.80 ขึ้นไป (ยิ่งค่าเข้าใกล้ 1.00 ยิ่งมีความเชื่อมั่นสูง) หากเครื่องมือมีค่าความเชื่อมั่นต่ำ อาจมีสาเหตุมาจากความไม่ชัดเจนของข้อคำถาม จำนวนข้อคำถามที่น้อยเกินไป หรือความคลาดเคลื่อนแบบสุ่ม (Random Error) ในขณะที่ผู้ตอบแบบสอบถามทำการประเมิน

2.2 ทฤษฎี

2.2.1 Big Five

Big Five คือ ทฤษฎีบุคลิกภาพห้าปัจจัย เป็นกรอบแนวคิดทางจิตวิทยาที่อธิบายลักษณะบุคลิกภาพของมนุษย์โดยแบ่งออกเป็น 5 มิติหลัก ได้แก่ การเปิดรับประสบการณ์ใหม่ (Openness) ความมีระเบียบวินัย (Conscientiousness) การแสดงออกภายนอก (Extraversion) ความเป็นมิตร (Agreeableness) และความไม่มั่นคงทางอารมณ์ (Neuroticism) ความสำคัญของทฤษฎี Big Five มักถูกนำไปใช้ในการแนะแนวการศึกษาและอาชีพ เนื่องจากบุคลิกภาพมีอิทธิพลโดยตรงต่อความชอบ ความถนัด และการปรับตัวในการเรียน เช่น ผู้ที่มี Openness สูง เหมาะกับสาขาที่ใช้ความคิดสร้างสรรค์ เช่น ศิลปะ การออกแบบ หรือการตลาด ผู้ที่มี Conscientiousness สูง เหมาะกับสาขาที่ต้องใช้ความละเอียดรอบคอบ เช่น บัญชี หรือการเงิน ผู้ที่มี Extraversion สูง มักประสบความสำเร็จในสาขาที่ต้องการการสื่อสาร เช่น การจัดการธุรกิจ การขาย หรือการโรงแรม ผู้ที่มี Agreeableness สูง เหมาะกับงานบริการหรือสาขาที่เกี่ยวข้องกับการ

ช่วยเหลือผู้อื่น เช่น การบริหารทรัพยากรมนุษย์ หรือการสังเคราะห์ผู้ที่มี Neuroticism ต่ำ มักจะรับมือกับความกดดันได้ดี เหมาะกับสาขาที่มีการแข่งขันสูง เช่น โลจิสติกส์ หรือการจัดการโครงการ

2.2.2 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) คือ การค้นหาข้อมูลที่มีประโยชน์จากแหล่งข้อมูล ที่มีเป็นจำนวนมาก เพื่อดึงข้อมูลที่มีประโยชน์มาทำการวิเคราะห์ค้นหา รูปแบบ หรือความสัมพันธ์ที่เกิดในฐานข้อมูล และจัดทำเป็นสารสนเทศเพื่อใช้ในการวางแผนบริหารจัดการ ธุรกิจ โดยการแยกข้อมูลที่มีประโยชน์ออกมาใช้งานเปรียบเทียบกับการทำเหมืองแร่ ที่จะต้องทำการแยกเศษหินดินทรายที่ไม่มีค่า และมีปริมาณมากออกจากแร่ที่มีมูลค่ามาก และมักจะมีปริมาณน้อย

2.2.3 หลักในการทำเหมืองข้อมูล

เหมืองข้อมูลเป็นการวิเคราะห์ความสัมพันธ์ของข้อมูลจากฐานข้อมูลขนาดใหญ่ในอดีต เพื่อการค้นหากฎความสัมพันธ์ของข้อมูล การจัดกลุ่มข้อมูล และจำแนกประเภทข้อมูล เพื่อพยากรณ์สิ่งที่จะเกิดขึ้นในอนาคตโดยมีรายละเอียด ดังต่อไปนี้

1) การค้นหากฎความสัมพันธ์ของข้อมูล (Association Rule) เป็นเทคนิคที่ใช้ในการค้นหารูปแบบกฎความสัมพันธ์ที่สนใจ หรือความสัมพันธ์ที่มีความถี่เกิดขึ้นร่วมกันบ่อยในฐานข้อมูลขนาดใหญ่ โดยพิจารณาจากค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) ที่ผ่านเกณฑ์ขั้นต่ำ (Minimum Threshold) ในบริบทของการศึกษา อาจนำเทคนิคนี้มาใช้ค้นหาความสัมพันธ์ระหว่างตัวแปรต่างๆ เช่น ความสัมพันธ์ระหว่างความถนัดของนักศึกษากับผลการเรียนในรายวิชาต่างๆ เพื่อนำผลลัพธ์มาใช้ในการวางแผนหลักสูตรหรือการแนะแนว

2) การจัดกลุ่มข้อมูล (Clustering) เป็นเทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) ที่ใช้ในการลดขนาดของข้อมูล ด้วยวิธีการจัดกลุ่มข้อมูลที่มีลักษณะหรือคุณสมบัติคล้ายคลึงกันให้อยู่ในกลุ่มเดียวกัน (จำนวน k กลุ่ม) โดยที่ยังไม่มีการกำหนดเป้าหมายหรือคำตอบ (Target Label) ไว้ล่วงหน้า มีวัตถุประสงค์สำคัญเพื่อจัดกลุ่มหน่วยวิเคราะห์ เช่น การจัดกลุ่มนักศึกษาที่มีโปรไฟล์บุคลิกภาพ (Personality Profile) หรือลักษณะนิสัยที่คล้ายคลึงกันให้อยู่ในกลุ่มเดียวกัน เพื่อให้ง่ายต่อการวิเคราะห์พฤติกรรมในภาพรวม

3) การจำแนกประเภทข้อมูล (Classification) เป็นเทคนิคที่ใช้ในการจำแนกประเภทข้อมูลด้วยคุณลักษณะต่าง ๆ ที่ใช้ในการสร้างแบบจำลอง (Model) เพื่อจำแนกประเภทข้อมูลตามคุณลักษณะต่างๆ โดยโมเดลจะเรียนรู้จากชุดข้อมูลในอดีตที่มีการกำหนดคำตอบหรือ

กลุ่มเป้าหมาย (Target Label) ไว้แล้วล่วงหน้า เพื่อนำไปทำนายคลาสของข้อมูลใหม่ที่กำลังจะเกิดขึ้น สำหรับโครงการนี้ คณะผู้จัดทำได้เลือกใช้เทคนิคการจำแนกประเภทข้อมูลด้วยอัลกอริทึมต้นไม้ตัดสินใจ (Decision Tree) มาทำการเรียนรู้ชุดข้อมูลแบบสอบถาม เพื่อวิเคราะห์หาความสัมพันธ์ระหว่าง "คะแนนบุคลิกภาพ Big Five ทั้ง 5 ด้าน" (ข้อมูลนำเข้า/คุณลักษณะ) กับ "สาขาวิชาทั้ง 9 สาขา" (กลุ่มเป้าหมาย/คำตอบ) ทำให้แบบจำลองสามารถทำนายและจำแนกความเหมาะสมได้ว่า ผู้ใช้งานที่มีลักษณะบุคลิกภาพแบบใด จะเหมาะสมกับการศึกษาต่อในสาขาวิชาใดมากที่สุด

2.2.4 ทฤษฎีการกำหนดขนาดกลุ่มตัวอย่าง (Taro Yamane)

ในการศึกษาวิจัยเชิงปริมาณ (Quantitative Research) ที่ประชากรมีขนาดใหญ่และไม่สามารถเก็บรวบรวมข้อมูลจากประชากรทั้งหมดได้ การกำหนดขนาดกลุ่มตัวอย่าง (Sample Size Estimation) ถือเป็นขั้นตอนที่มีความสำคัญอย่างยิ่งยวด เพื่อให้ได้กลุ่มตัวอย่างที่มีคุณสมบัติเป็นตัวแทนที่ดีของประชากร (Representative) และมีขนาดปริมาณที่เพียงพอต่อการวิเคราะห์ทางสถิติ ทาโร ยามาเน่ ได้เสนอวิธีการและสูตรการคำนวณเพื่อหาขนาดกลุ่มตัวอย่างที่เหมาะสมสำหรับการวิจัยเชิงสำรวจ (Survey Research) โดยวิธีการนี้ถูกออกแบบมาเพื่อใช้ประเมินสัดส่วนของประชากรในกรณี ทฤษฎีของยามาเน่มีข้อตกลงเบื้องต้น (Assumptions) ในการคำนวณ คือ กำหนดให้สัดส่วนของลักษณะที่สนใจในประชากรมีค่าเท่ากับ 0.5 (ซึ่งเป็นค่าที่ทำให้เกิดความแปรปรวนสูงสุด) และตั้งอยู่บนระดับความเชื่อมั่น (Confidence Level) ที่ร้อยละ 95 ในการคำนวณหาขนาดกลุ่มตัวอย่างตามหลักการของยามาเน่ ผู้วิจัยจำเป็นต้องทราบขนาดของประชากรทั้งหมด (N) และต้องกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ (Margin of Error: \$e\$) ยกตัวอย่างเช่น หากประชากรเป้าหมายมีจำนวน 2,000 คน และผู้วิจัยกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ไว้ที่ร้อยละ 5 เมื่อผ่านกระบวนการคำนวณแล้ว จะได้ขนาดของกลุ่มตัวอย่างขั้นต่ำที่ต้องการคือ 333 คน เป็นต้น

2.2.5 ทฤษฎีการวิเคราะห์ความแปรปรวน (ANOVA)

การวิเคราะห์ความแปรปรวน เป็นวิธีการทางสถิติที่นิยมใช้ในการเปรียบเทียบค่าเฉลี่ยของข้อมูลที่มีการแบ่งกลุ่มตั้งแต่ 3 กลุ่มขึ้นไป โดยอาศัยหลักการแยกความแปรผันทั้งหมด (Total Variation) ออกเป็นความแปรผันระหว่างกลุ่ม (Between Groups) และความแปรผันภายในกลุ่ม (Within Groups) การนำสถิติการวิเคราะห์ความแปรปรวนมาใช้นั้น ผู้วิจัยจำเป็นต้องคำนึงถึงข้อตกลงเบื้องต้น (Assumptions) ที่สำคัญ ได้แก่ ข้อมูลตัวแปรตามต้องมีการวัดตั้งแต่มาตราอันตรภาค (Interval Scale) ขึ้นไป กลุ่มตัวอย่างแต่ละกลุ่มต้องมาจากประชากรที่มีการแจกแจงแบบปกติ (Normal Distribution) และประชากรแต่ละกลุ่มต้องมีความ

แปรปรวนที่เท่ากัน ทั้งนี้ ในการทดสอบทางสถิติ หากพบว่าอัตราส่วนความแปรปรวน (F-Ratio) มีนัยสำคัญทางสถิติ จะสามารถสรุปได้ว่ามีประชากรอย่างน้อย 1 คู่ที่มีค่าเฉลี่ยแตกต่างกัน

คณะผู้จัดทำได้นำทฤษฎีการวิเคราะห์ความแปรปรวน มาประยุกต์ใช้ในขั้นตอนการตรวจสอบคุณภาพของชุดข้อมูลรายข้อคำถาม (Item Analysis) เพื่อพิสูจน์ความแตกต่างของระดับความคิดเห็นในแบบสอบถามวัดบุคลิกภาพ 5 องค์ประกอบ (Big Five Personality Traits) ก่อนนำข้อมูลเข้าสู่กระบวนการประมวลผล ผลจากการวิเคราะห์ด้วยสถิติ ANOVA ช่วยยืนยันเชิงประจักษ์ว่า ผู้ตอบแบบสอบถามได้พิจารณาและให้คะแนนตามบริบทเฉพาะของแต่ละข้อคำถามอย่างแท้จริง ปราศจากการให้คะแนนในระดับเดียวกันทั้งหมด (Response Set Bias) ส่งผลให้คะแนนในแต่ละมิติของบุคลิกภาพมีความแตกต่างกันอย่างมีนัยสำคัญ ข้อมูลชุดนี้จึงมีคุณสมบัติที่ดี ปราศจากความคลาดเคลื่อน และมีประสิทธิภาพสูงเพียงพอสำหรับการนำไปใช้เป็นข้อมูลนำเข้า (Input Data) เพื่อให้แบบจำลองทำหน้าที่จำแนกและทำนายความเหมาะสมของทั้ง 9 สาขาวิชาในขั้นตอนการทำเหมืองข้อมูล (Data Mining) ต่อไปได้อย่างแม่นยำ

2.2.6 ทฤษฎีเกี่ยวข้องกับการสร้างเว็บไซต์

2.2.6.1 การออกแบบเว็บไซต์ หรือ Web Design คือ กระบวนการในการสร้างหน้าตา และองค์ประกอบภายในเว็บไซต์ ตั้งแต่การจัดวางเนื้อหา การเลือกใช้สี ฟอนต์ การเลือกใช้ภาพ หรือกราฟิก และการออกแบบระบบนำทางผู้ใช้งานหน้าเว็บ ให้สวยงามสอดคล้อง 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 ขณะเดียวกันก็ต้องเข้าใจความต้องการของกลุ่มเป้าหมาย เพื่อมอบประสบการณ์ใช้งานที่เหมาะสมกับผู้ใช้จริง

2.2.6.2 องค์ประกอบในการออกแบบเว็บไซต์ การออกแบบเว็บไซต์ที่มีประสิทธิภาพ นั้นต้องคำนึงถึงองค์ประกอบสำคัญ ดังต่อไปนี้

1) การออกแบบที่เรียบง่าย (Simplicity) ในยุคที่ผู้คนมีความสนใจในช่วงเวลา สั้นลง (Short attention span) และต้องการข้อมูลที่กระชับ และตรงประเด็น การออกแบบเว็บไซต์

แบบมินิมอล (Minimalist) ซึ่งเน้นความเรียบง่ายจึงเป็นแนวคิดที่ควรนำมาใช้ เพื่อให้ผู้ใช้งานสามารถมุ่งเน้นประเด็นสำคัญได้อย่างชัดเจน

ทั้งนี้ความเรียบง่ายไม่ควรถึงขั้นที่ทำให้ขาดจุดเด่นที่น่าสนใจ เพราะ การออกแบบหน้าเว็บไซต์ที่ซับซ้อนเกินไปอาจทำให้ผู้ใช้งานเกิดความสับสน และไม่สามารถค้นหาข้อมูลที่ต้องการได้อย่างสะดวก ดังนั้น ควรใช้กราฟิกและข้อความในปริมาณที่พอเหมาะ เน้นที่ความสะอาดตา และการใช้งานที่ราบรื่น โดยยังสามารถสื่อสารข้อมูลสำคัญของหน้าเว็บไซต์ ได้อย่างมีประสิทธิภาพ

2) การเลือกใช้สีที่เข้ากัน (Color Scheme) การเลือกใช้สีเป็นองค์ประกอบ สำคัญในการออกแบบเว็บไซต์ (Web Design) เนื่องจากสีสามารถสื่อถึงอารมณ์ และความรู้สึก รวมทั้งส่งผลต่อการรับรู้ของผู้ใช้งาน หากเลือกใช้สีที่เหมาะสม และสอดคล้องกับลักษณะของเว็บไซต์ 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 ข้อความ เป็นองค์ประกอบสำคัญในการออกแบบหน้าเว็บไซต์ (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 สิ่งสำคัญ คือ ต้องใช้ White Space อย่างพอเหมาะ ไม่มาก หรือน้อยเกินไป

5) การออกแบบเว็บไซต์ให้รองรับการแสดงผลบนมือถือ ด้วยการเพิ่มขึ้นของ ผู้ใช้งานอินเทอร์เน็ตผ่านสมาร์ทโฟน และแท็บเล็ต การออกแบบเว็บไซต์ที่รองรับการแสดงผล บนอุปกรณ์พกพา หรือ Responsive Design จึงกลายเป็นปัจจัยสำคัญ Responsive Design คือ

การออกแบบเว็บไซต์ให้ปรับขนาด และการจัดเรียงเนื้อหาได้โดยอัตโนมัติ ไม่ว่าจะเป็นคอมพิวเตอร์ สมาร์ทโฟน หรือแท็บเล็ต ทำให้ผู้ใช้สามารถใช้งานได้อย่างราบรื่น

6) การใช้ Visual Content ในการตกแต่งเว็บไซต์ Visual Content เช่น รูปภาพ ไอคอน อินโฟกราฟิก และวิดีโอ เป็นองค์ประกอบที่ช่วยเสริมสร้างความน่าสนใจให้กับเว็บไซต์ นอกจากการตกแต่งให้สวยงาม ยังช่วยดึงดูดความสนใจ และอธิบายเนื้อหาที่ยากให้เข้าใจง่ายขึ้น การใช้ Visual Content ควรอยู่ในปริมาณที่เหมาะสม และต้องคำนึงถึงขนาดไฟล์เพื่อไม่ให้เกิดกระทบ ต่อความเร็วในการโหลดเว็บไซต์

7) การวางตำแหน่งเนื้อหาหลักให้เด่นชัด (Highlight Key Content) การวางตำแหน่งเนื้อหาที่ต้องการเน้นให้เห็นได้ชัดเจนเป็นสิ่งสำคัญ เพราะผู้ชมมักให้ความสนใจกับ สิ่ง ที่เห็นชัดเจนที่สุดก่อน เช่น ข้อความขนาดใหญ่ หรือรูปภาพเด่นชัด ควรใช้หลักการจัดลำดับ ความสำคัญขององค์ประกอบ (Visual Hierarchy) เพื่อให้เนื้อหาสำคัญโดดเด่น และสะดุดตา มากขึ้น

8) การออกแบบระบบนำทางเว็บไซต์ (Navigation) การออกแบบระบบนำทาง หรือเมนูที่ใช้งานง่ายเป็นหัวใจสำคัญ เพราะช่วยให้ผู้ใช้งานสามารถค้นหา และเข้าถึงข้อมูลได้ อย่าง รวดเร็ว ควรวางโครงสร้างเมนูให้ชัดเจน และเป็นหมวดหมู่ การใช้ข้อความในเมนูควร กระชับ 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 กับ เทคโนโลยีที่เปลี่ยนแปลง และพฤติกรรมของผู้ใช้ การใช้เครื่องมืออย่าง Google Analytics ช่วย ให้เห็นพฤติกรรมการใช้งาน และสามารถปรับปรุงเว็บไซต์ให้ตอบโจทยความต้องการ ได้อย่าง มีประสิทธิภาพ

2.2.6.3 รูปแบบโครงสร้างของเว็บไซต์ การออกแบบโครงสร้างของเว็บไซต์ สามารถ ทำ ได้หลากหลายแบบ ซึ่งก็ขึ้นอยู่กับความชอบ และความถนัดของแต่ละบุคคล นอกจากนี้ยัง ขึ้นอยู่กับกลุ่มเป้าหมายที่ต้องการนำเสนอ เพราะจะต้องออกแบบให้เหมาะสมกับ การใช้งาน ของ กลุ่มเป้าหมายมากที่สุด โดยโครงสร้างของเว็บไซต์ส่วนใหญ่ก็จะประกอบ ไปด้วย 4 รูปแบบ ดังนี้

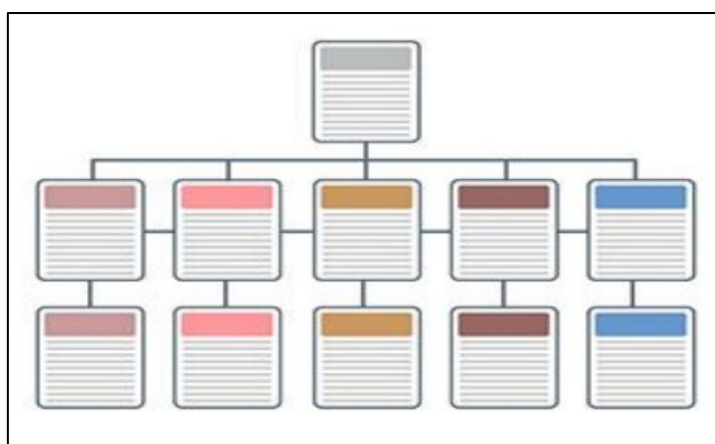
1) โครงสร้างเว็บไซต์แบบเรียงลำดับ จะเป็นโครงสร้างแบบธรรมดาที่นิยม ใช้งานกันมากที่สุด เนื่องจากมีความง่ายต่อการจัดระบบข้อมูล และสามารถนำเสนอเรื่องราวตามลำดับได้เป็นอย่างดี เหมาะกับเว็บไซต์ที่มีขนาดเล็ก มีเนื้อหาที่ไม่ซับซ้อน ส่วนใหญ่ก็จะเป็นพวกเว็บไซต์ที่ให้ความรู้ หรือเว็บไซต์องค์กรขนาดย่อม โดยลักษณะการลิงค์เนื้อหา ก็จะลิงค์ไปที่หน้า มีทิศทาง การเข้าสู่เนื้อหาต่าง ๆ ในแบบเส้นตรง ใช้ปุ่มเดินหน้า - ถอยหลังในการกำหนด ทิศทาง จึงทำให้การใช้งานเป็นไปอย่างง่าย แต่โครงสร้างเว็บไซต์แบบเรียงลำดับมีข้อเสีย คือจะทำให้ผู้ใช้งานต้องเสียเวลาในการเข้าสู่เนื้อหาเพราะไม่สามารถกำหนดทิศทาง การเข้า สู่เนื้อหาด้วยตัวเองได้



ภาพที่ 2.6 แสดงโครงสร้างแบบเรียงลำดับ

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

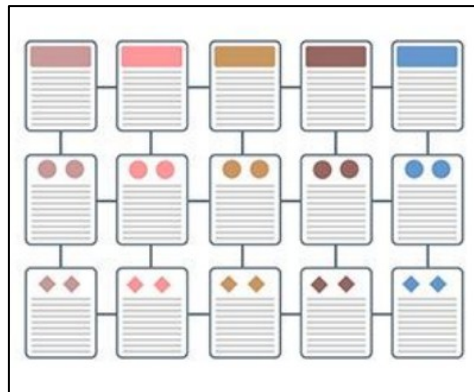
2) โครงสร้างแบบลำดับขั้น นิยมใช้กับเว็บที่มีความซับซ้อนของข้อมูล เพื่อให้สามารถเข้าถึงข้อมูลต่าง ๆ ได้ง่ายขึ้น โดยจะมีการแบ่งเนื้อหาออกเป็น ส่วน ๆ และมีการนำเสนอ รายละเอียดย่อย ๆ ที่ลดหลั่นกันมา ทำให้สามารถทำความเข้าใจกับโครงสร้างเนื้อหาได้ง่ายขึ้น โดยจะมีโฮมเพจเป็นจุดเริ่มต้น และจุดรวมจุดเดียวที่จะนำไปสู่การเชื่อมโยงเนื้อหาเป็นลำดับ จากบนลงล่าง



ภาพที่ 2.7 แสดงโครงสร้างแบบลำดับขั้น

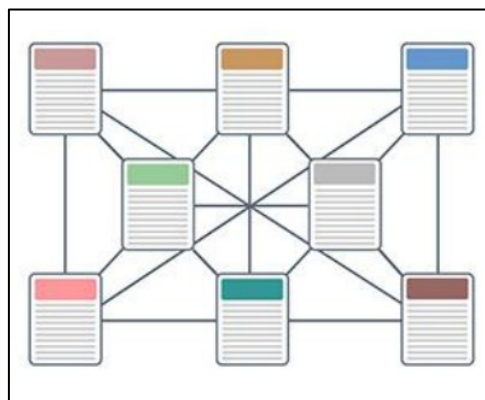
(ที่มา: <https://asria.org/website-structure/>)

3) โครงสร้างแบบตาราง เป็นโครงสร้างการออกแบบเว็บไซต์ที่มีความซับซ้อน แต่ก็มีคามยืดหยุ่นในระดับหนึ่ง เพื่อให้ผู้ใช้งานสามารถเข้าสู่เนื้อหาต่าง ๆ ได้ง่ายขึ้น การออกแบบ ในลักษณะนี้จะมีการเชื่อมโยงเนื้อหาในแต่ละส่วนซึ่งกัน และกัน ทำให้ผู้ใช้งานสามารถเปลี่ยน ทิศทาง หรือกำหนดทิศทางในการเข้าสู่เนื้อหาด้วยตัวเองได้ จึงไม่ทำให้เสียเวลา และทำให้เว็บไซต์ มีความทันสมัยขึ้น



ภาพที่ 2.8 แสดงโครงสร้างแบบตาราง
(ที่มา: <https://asria.org/website-structure/>)

4) โครงสร้างแบบใยแมงมุม เป็นโครงสร้างที่ได้รับความนิยมเป็นอย่างมาก เพราะมีความยืดหยุ่นมากที่สุด โดยทุกหน้าเว็บจะมีการเชื่อมโยงถึงกันหมด ทำให้สามารถเข้า ถึงหน้าเว็บเพจต่าง ๆ ที่ต้องการได้อย่างง่าย และมีความอิสระมากขึ้น นอกจากนี้ก็สามารถ เชื่อมโยงไปสู่เว็บไซต์ภายนอกได้ดี



ภาพที่ 2.9 แสดงโครงสร้างแบบใยแมงมุม
(ที่มา: <https://asria.org/website-structure/>)

2.2.6.4 ส่วนประกอบหลักในหน้าเว็บไซต์ มีองค์ประกอบหลัก ๆ เช่น ชื่อเว็บไซต์ แถบเมนู นำทาง แถบเมนูย่อย ข้อความหัวเรื่อง ภาพประกอบตกแต่ง หรือภาพเนื้อหา ป้ายโฆษณา และอื่น ๆ ตามที่ต้องการนำเสนอ บางเว็บไซต์อาจจะแสดงภาพกราฟิกของสินค้า หรือบริการ แทนการแสดงข้อความ ทั้งนี้ขึ้นอยู่กับการออกแบบหน้าเว็บไซต์ให้เหมาะสมกับเนื้อหาที่ต้องการนำเสนอ

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 เพื่อให้ผู้เยี่ยมชมได้รู้จักธุรกิจ ของ คุณมากยิ่งขึ้น ข้อมูลที่ควรนำเสนอ ได้แก่ การแนะนำว่าสินค้าหรือบริการของคุณคืออะไร คุณสมบัตินเด่นของสินค้า หรือบริการ ประโยชน์ที่ผู้ใช้จะได้รับ ข่าวสารหรือกิจกรรมสำคัญ รวมถึง โปรโมชันที่เกี่ยวข้อง ช่องทางการติดต่อที่สะดวกและเข้าถึงง่าย คำรับรองจากลูกค้าที่เคยใช้บริการ รวมถึงรางวัล หรือการรับรองคุณภาพบริการที่ธุรกิจของคุณได้รับ นอกจากนี้ หากต้องการเพิ่มโอกาสให้ธุรกิจของคุณถูกค้นพบได้ง่ายขึ้น ควรใส่ คำสำคัญ (Keywords) ที่เกี่ยวข้องกับสินค้า หรือบริการในเนื้อหาของเว็บไซต์ ซึ่งจะช่วยให้ กลุ่มเป้าหมายสามารถเข้าถึงเว็บไซต์ของคุณได้สะดวก และตรงกับความต้องการมากยิ่งขึ้น

4) ฟังก์ชันการค้นหา (Search Function) การจัดทำฟังก์ชันการค้นหาในเว็บไซต์เป็นสิ่งสำคัญที่ช่วยให้ผู้เยี่ยมชมสามารถค้นหาสินค้าหรือบริการได้อย่างรวดเร็วและสะดวกยิ่งขึ้น โดยเฉพาะในกรณีที่เว็บไซต์ของคุณมีสินค้าหรือบริการจำนวนมาก ฟังก์ชันนี้จะช่วยให้ผู้ใช้สามารถ ค้นหาสิ่งที่ต้องการได้อย่างง่ายดาย การออกแบบ และพัฒนาฟังก์ชันการค้นหาที่มีประสิทธิภาพ จึงเป็นสิ่งที่คุณควรให้ความสำคัญ 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

2.2.6.5 ทฤษฎีสีในการออกแบบเว็บไซต์

การสร้างสีบนหน้าเว็บเป็นสิ่งสื่อความหมายของเว็บไซต์ได้อย่าง ชัดเจนการเลือกใช้สีให้เหมาะสม กลมกลืน ไม่เพียงแต่จะสร้างความพึงพอใจให้กับผู้ใช้ แต่ยังสามารถทำให้เห็นถึงความแตกต่างระหว่างเว็บไซต์ได้ สีเป็นองค์ประกอบหลักสำหรับการ ตกแต่งเว็บ จึงจำเป็นอย่างยิ่งที่จะต้องทำความเข้าใจเกี่ยวกับการใช้สีระบบสีที่แสดงบน จอคอมพิวเตอร์ มีระบบการแสดงผลผ่านหลอดลำแสงที่เรียกว่า CRT (Cathode ray tube) โดย มีลักษณะระบบสีแบบบวก อาศัยการผสมของของแสงสีแดง สีเขียว และสีน้ำเงิน หรือระบบสี RGB สามารถกำหนดค่าสีจาก 0 ถึง 255 ได้ จากการรวมสีของแม่สีหลักจะทำให้เกิดแสงสีขาว มีลักษณะเป็นจุดเล็ก ๆ บนหน้าจอไม่สามารถมองเห็นด้วยตาเปล่าได้ จะมองเห็นเป็นสีที่ถูก ผสมเป็นเนื้อสีเดียวกันแล้ว จุดแต่ละจุด หรือ พิกเซล (Pixel) เป็นส่วนประกอบของภาพบนหน้าจอคอมพิวเตอร์ โดยจำนวนบิตที่ใช้ในการกำหนดความสามารถของการแสดงสีต่าง ๆ เพื่อสร้างภาพบนจอ นั้นเรียกว่า บิตเด็ป (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 กระตุ้นการรับรู้ทางด้านจิตใจมนุษย์ สีแต่ละสีให้ความรู้สึกถึงอารมณ์ที่ไม่เหมือนกัน สีบางสีให้ความรู้สึกสงบ บางสีให้ความรู้สึกตื่นเต้นรุนแรง สีจึงเป็นปัจจัยสำคัญอย่างยิ่งต่อการออกแบบเว็บไซต์ ดังนั้นการเลือกใช้โทนสีภายในเว็บไซต์เป็นการแสดงถึงความแตกต่างของสีที่แสดงออกทางอารมณ์ มีชีวิตชีวาหรือเศร้าโศก รูปแบบของสีที่สายตาของมนุษย์มองเห็น สามารถแบ่ง ออกเป็น 3 กลุ่ม คือ

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

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

3) สีโทนกลาง (Neutral Colors) สีที่เป็นกลาง ประกอบด้วย สีดำ สีขาว สีเทา และสีน้ำตาล กลุ่มสีเหล่านี้คือ สีกลางที่สามารถนำไปผสมกับสีอื่นๆ เพื่อให้เกิดสีกลางขึ้นมา (visperhost, 2567)

2.3 วรรณกรรมที่เกี่ยวข้อง

สุจารัตน์ ชาวนาฟาง (2023) ได้ประยุกต์ใช้เทคนิค ต้นไม้ตัดสินใจ(Decision Tree) เพื่อสร้างโมเดลซ์เทคนิคเหมืองข้อมูลเพื่อวิเคราะห์และพยากรณ์พฤติกรรมการเลือกสาขาวิชาของผู้สมัครเข้าศึกษาระดับปริญญาตรี โดยใช้ข้อมูลจากระบบทะเบียนของมหาวิทยาลัยราชภัฏนครปฐมในช่วงปีการศึกษา 2557–2561โดยอาศัยข้อมูลจากระบบทะเบียนนักศึกษาที่ประกอบด้วยคุณลักษณะส่วนบุคคล เช่น เพศ จังหวัดที่อยู่ คะแนนสอบ GAT/PAT สาขาวิชาที่เลือก และข้อมูลด้านประชากรศาสตร์อื่น ๆ ต้นไม้ตัดสินใจถูกนำมาใช้เพื่อแยกและจัดกลุ่มข้อมูลตามเงื่อนไขที่สำคัญ โดยใช้หลักการของอัลกอริทึม ID3 ที่อาศัยทฤษฎีสารสนเทศ (Information Gain) ในการเลือกคุณลักษณะที่มีความสำคัญสูงสุดเป็นโหนดหลักของการจำแนก มีข้อดีคือสามารถตีความได้ง่าย เข้าใจเงื่อนไขที่นำไปสู่การเลือกสาขาวิชา และสามารถนำเสนอผลลัพธ์ในลักษณะของแผนผังการตัดสินใจที่ชัดเจน เหมาะสมต่อการประยุกต์ใช้เป็นเครื่องมือสนับสนุนการตัดสินใจของผู้สมัครและผู้บริหารการศึกษา

อนันต์ ปินะเต (2024) ได้ศึกษาการวิเคราะห์คุณลักษณะของผู้สมัครในการเลือกสาขาวิชาในระบบ TCAS งานวิจัยนี้มุ่งเน้นไปที่กลุ่มเป้าหมายคือผู้สมัครเข้าศึกษาในระดับปริญญาตรีในระบบ TCAS ช่วงปีการศึกษา 2561-2564 โดยใช้ชุดข้อมูลจำนวนมากถึง 58,110 รายการ นักวิจัยได้เลือกปัจจัยที่เกี่ยวข้องกับการเลือกสาขาวิชามาทำการวิเคราะห์ ได้แก่ เพศ, จังหวัด, คะแนนความถนัดทั่วไป (GAT) และคะแนนความถนัดทางด้านวิชาชีพและวิชาการ (PAT) งานวิจัยใช้วิธีการทางเทคนิคเหมืองข้อมูล (Data Mining) โดยเฉพาะอย่างยิ่งการค้นหากฎความสัมพันธ์ (Association Rule) ด้วยอัลกอริทึมเอปริออริ (Apriori Algorithm) เพื่อให้ได้กฎความสัมพันธ์ที่มีความน่าเชื่อถือสูง งานวิจัยจึงกำหนดค่าความเชื่อมั่นไว้ที่ 80% จากการวิเคราะห์ผลพบว่าสามารถสร้างกฎความสัมพันธ์ได้ทั้งหมด 36 กฎ ซึ่งปัจจัยที่มีผลต่อการตัดสินใจเลือกอันดับสาขาวิชามากที่สุดคือ เพศ, จังหวัด, คะแนน TGAT และคะแนน TPAT ตามลำดับ

วรการ ใจดี และ นรินทร์ จิวิตัน (2023) ได้ศึกษาปัจจัยที่ส่งผลกระทบต่อตัดสินใจเข้าศึกษาต่อในระดับ ปริญญาตรี คณะบริหารธุรกิจและศิลปศาสตร์ โดยใช้เทคนิคการคัดเลือกคุณลักษณะและเทคนิค การ ทำเหมืองข้อมูล การวิจัยนี้ศึกษาปัจจัยที่มีผลต่อการตัดสินใจเข้าศึกษาต่อในระดับปริญญาตรี คณะ บริหารธุรกิจและศิลปศาสตร์ โดยวิเคราะห์ข้อมูลจากการรับสมัครนักศึกษามหาวิทยาลัยเทคโนโลยี ราชมนคล ล้านนาเชียงใหม่ ช่วงปีการศึกษา 2563-2565 ซึ่งมีทั้งหมด 2,509 รายการ โดยใช้เทคนิค การคัดเลือก คุณลักษณะ 3 เทคนิค และเทคนิคการทำเหมืองข้อมูล 6 เทคนิค ผลการศึกษาแสดงให้เห็นว่าปัจจัยสำคัญที่ ส่งผลต่อการตัดสินใจ ได้แก่ สาขาวิชาหรือหลักสูตรที่เลือก สายการเรียนเดิม และวุฒิการศึกษาเดิม ส่วน เทคนิคการทำเหมืองข้อมูลที่มีความแม่นยำสูงสุดคือ เทคนิคการโหวต ด้วย ความถูกต้อง 73.44% มากกว่า เทคนิคอื่น ๆ เช่น เทคนิคป่าสุ่มและเทคนิคการถดถอยโลจิสติก. การใช้เทคนิคเหมืองข้อมูลเป็นการค้นหาองค์ ความรู้ (Knowledge) จากข้อมูลที่มีอยู่มาทำการ วิเคราะห์ปัจจัยที่ส่งผลต่อการตัดสินใจเข้าศึกษาต่อในระดับ ปริญญาตรี คณะบริหารธุรกิจและศิลป ศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมนคลล้านนาเชียงใหม่ ซึ่งในการ วิจัยครั้งต่อไป คณะผู้วิจัยมีความเห็นว่า 1. ควรเพิ่มปัจจัยอื่น ๆ ที่อาจจะมีผลต่อการเข้าศึกษาต่อในระดับ ปริญญาตรี เช่น ปัจจัย ด้านผู้ปกครอง อาชีพ ผู้ปกครอง รายได้ผู้ปกครอง ปัจจัยด้านข้อมูลส่วนตัวนักศึกษา เช่น งานอดิเรก ความสามารถ อาชีพในอนาคต เป็นต้น 2. ควรมีการใช้เทคนิคการพยากรณ์อื่น ๆ เข้ามา เปรียบเทียบ เช่น เทคนิคการเรียนรู้แบบรวมกลุ่ม (Ensemble Model) ประเภทต่าง ๆ เพื่อเพิ่มประสิทธิภาพ ให้กับตัว แบบพยากรณ์ รวมถึงมีการปรับปรุงพารามิเตอร์ของแต่ละ

เทคนิคการพยากรณ์ให้ประสิทธิภาพมากขึ้น 3. ควรมีการใช้เทคนิคการปรับสมดุลข้อมูลที่หลากหลาย เพื่อเพิ่มประสิทธิภาพของตัวแบบ พยากรณ์ เช่น วิธีการ สุ่มเพิ่ม (Oversampling) วิธีการผสมผสาน (Hybrid Method) เป็นต้น 4. ควรมีการทดสอบเทคนิคการ คัดเลือกคุณลักษณะและเทคนิคการพยากรณ์หลายๆ ครั้ง เพื่อให้ได้ ประสิทธิภาพมากขึ้นและไม่ให้เกิด ปัญหา overfitting

จิฎติภักษ์ ศรีสุรธีกุล, บุญมา อิมวิเศษและสำราญ บุญเจริญ (2024) มีวัตถุประสงค์ เพื่อศึกษาและเปรียบเทียบปัจจัยการตัดสินใจเลือกศึกษาต่อในสถาบันอุดมศึกษาเอกชนในภาคตะวันออกเฉียงเหนือ กลุ่มตัวอย่าง ได้แก่ นักศึกษาชั้นปีที่ 1 ปีการศึกษา 2565 ที่กำลังศึกษาอยู่ใน สถาบันอุดมศึกษาเอกชนในภาคตะวันออกเฉียงเหนือ จำนวน 389 คน เครื่องมือการวิจัย คือ แบบสอบถาม สถิติวิเคราะห์ข้อมูล สามารถแบ่งออกเป็น 3 ตอน คือตอนที่ 1 แบบสอบถามเกี่ยวกับข้อมูลส่วนบุคคลของผู้ตอบแบบสอบถาม เป็นคำถามปลายปิด โดยมีตอบแบบเลือกตอบ (Check List) ที่ตรงกับสภาพความเป็นจริง ประกอบด้วย เพศ อาชีพบิดา-มารดา/ผู้ปกครองรายได้ครอบครัว ผลการเรียนเฉลี่ยสะสม และ คณะวิชาที่ ศึกษาอยู่จำนวน 5 ข้อ ส่วนตอนที่ 2 เป็นแบบสอบถามเกี่ยวกับปัจจัยสำคัญที่ส่งผลต่อการตัดสินใจ (Kotler & Keller, 2009) เป็นคำถามปลายปิด โดยมีตอบแบบเลือกตอบประกอบด้วย 4 ด้าน ได้แก่ ด้านสังคม ด้าน วัฒนธรรม ด้านจิตวิทยา และด้านการตลาดบริการ (7Ps) จำนวน 26 ข้อซึ่งในแต่ละคำถามจะมีคำตอบให้เลือก 5 ระดับ ตามวิธีของลิเคิร์ท(บุญชม ศรีสะอาด, 2560) และตอนที่ 3 เป็นแบบสอบถามเกี่ยวกับข้อเสนอแนะซึ่ง เป็นคำถามปลายเปิดเพื่อเปิดโอกาสให้ผู้ตอบแบบสอบถามได้แสดงความคิดเห็นและข้อเสนอแนะ

Soto, C. J., & John, O. P. (2017) ได้พัฒนาแบบสอบถามเวอร์ชันใหม่ที่เรียกว่า BFI2 (Big Five Inventory-2) โดยวัตถุประสงค์ของงานวิจัยนี้คือการพัฒนาแบบสอบถาม BFI-2 ที่มีโครงสร้าง แบบลำดับชั้น (hierarchical) โดยแต่ละปัจจัยหลักของ Big Five จะถูกแบ่งออกเป็น 3 มิติย่อย (facets) เพื่อเพิ่ม ความแม่นยำ ความละเอียด และความสามารถในการทำนายพฤติกรรม และยังคงยึดตามโครงสร้างของ Big Five ที่ประกอบด้วย 5 ปัจจัยหลัก ได้แก่ การเปิดเผย (Extraversion) การยอมรับผู้อื่น (Agreeableness) ความ มุ่งมั่นตั้งใจ (Conscientiousness) ความไม่มั่นคงทางอารมณ์ (Negative Emotionality) ความเปิดกว้างทางความคิด (Open-Mindedness หรือ Openness) โดยแบบสอบถาม BFI-2 ประกอบด้วย 60 ข้อ ใช้มาตราส่วน Likert 5 ระดับ โดยมีทั้งข้อที่เป็นบวกและข้อกลับด้าน ใช้กลุ่มตัวอย่าง

หลากหลายทั้งเชื้อชาติ อายุ และภูมิหลัง โดยใช้การวิเคราะห์โครงสร้างเชิงยืนยัน (Confirmatory Factor Analysis) และการวัดความเชื่อมั่นของ แบบสอบถาม (เช่น ค่า Cronbach's alpha) รวมถึงการทดสอบความสามารถในการพยากรณ์พฤติกรรม (predictive validity) และผลการวิจัยแสดงให้เห็นถึงความเที่ยงตรงและความตรงในระดับสูง โดยค่า Cronbach's alpha ของแต่ละมิติย่อยอยู่ในเกณฑ์ดีมาก (> 0.75) แบบสอบถามนี้สามารถทำนายพฤติกรรม และลักษณะต่างๆ ได้อย่างแม่นยำ เช่น การเรียน การทำงาน และความสัมพันธ์ทางสังคม

วัลลภ รัฐฉัตรานนท์ (2562) ได้ศึกษาวิเคราะห์เรื่อง การหาขนาดตัวอย่างที่เหมาะสมสำหรับการวิจัย: มายาคติในการใช้สูตรของ ทาโร ยามาเน่ และเครทซี-มอร์แกน งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างความเข้าใจที่ถูกต้องตามหลักสถิติศาสตร์ ในการกำหนดขนาดกลุ่มตัวอย่างสำหรับการวิจัยเชิงปริมาณ โดยได้อธิบายถึงเงื่อนไขที่สำคัญที่สุดของการใช้สูตรทาโร ยามาเน่ (Taro Yamane) ว่าเป็นสูตรที่ถูกออกแบบมาสำหรับโครงการวิจัยที่ "ทราบจำนวนประชากรเป้าหมายที่แน่นอน (Finite Population)" ผลการศึกษาชี้ให้เห็นว่า การใช้สูตรของยามาเน่ให้เกิดความน่าเชื่อถือและสามารถเป็นตัวแทนของประชากรได้อย่างแท้จริงนั้น นักวิจัยควรกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ (Margin of Error) ไว้ที่ไม่เกินร้อยละ 5 ($e = 0.05$) หรือร้อยละ 1 ($e = 0.01$) เท่านั้น การกำหนดค่าความคลาดเคลื่อนที่สูงกว่านี้จะทำให้ผลลัพธ์ที่ได้ขาดความน่าเชื่อถือทางสถิติและเกิดความลำเอียง (Bias) ในการวิเคราะห์ข้อมูล

กิตติพงษ์ พิพิธกุล (2561) ได้ศึกษาเรื่องคุณภาพเครื่องมือแบบสอบถาม Validity กับ Reliability ในการวิจัยทางรัฐประศาสนศาสตร์ โดยมุ่งเน้นการอธิบายและกำหนดมาตรฐานการตรวจสอบคุณภาพของเครื่องมือประเภทแบบสอบถามที่ใช้มาตรวัดแบบประเมินค่า (Rating Scale) งานวิจัยนี้ได้เสนอว่าเครื่องมือที่ดีจะต้องผ่านการทดสอบ 2 มิติหลัก คือ ความตรงตามเนื้อหา (Content Validity) ซึ่งพิจารณาจากค่าดัชนีความสอดคล้อง (IOC) ที่ต้องมีค่าตั้งแต่ 0.50 ขึ้นไป และความเชื่อมั่น (Reliability) ซึ่งเป็นการทดสอบความคงเส้นคงวาของการวัดผ่านค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Alpha) ที่ควรมีค่าตั้งแต่ 0.70 ขึ้นไป การนำแนวคิดนี้มาใช้มีข้อดีคือ ช่วยเป็นหลักประกันว่าชุดข้อมูลที่เก็บรวบรวมมาได้นั้นมีความถูกต้องแม่นยำ และปราศจากความคลาดเคลื่อนแบบสุ่ม เหมาะสมต่อการนำไปเป็นข้อมูลตั้งต้น (Input Data) สำหรับการวิเคราะห์สถิติขั้นสูงหรือการทำเหมืองข้อมูลต่อไป

ปริญญา สิริอรรถะกุล (2555) ได้ศึกษาและรวบรวมองค์ความรู้เรื่องการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-Way ANOVA) สำหรับการวิจัยทางสังคมศาสตร์ โดยอธิบายหลักการทางสถิติที่ใช้ในการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของข้อมูลที่มีการแบ่งกลุ่มตั้งแต่ 3 กลุ่มขึ้นไป การศึกษาได้ระบุถึงเงื่อนไขและข้อตกลงเบื้องต้น (Assumptions) ที่สำคัญก่อนการวิเคราะห์ เช่น ข้อมูลต้องมีระดับการวัดตั้งแต่มาตราอันตรภาค (Interval Scale) ข้อมูลมีการแจกแจงแบบปกติ และมีความแปรปรวนระหว่างกลุ่มเท่ากัน โดยอาศัยหลักการแยกความแปรผันทั้งหมดออกเป็นความแปรผันระหว่างกลุ่มและภายในกลุ่ม วิธีการนี้มีข้อดีคือสามารถใช้ประเมินและยืนยันความแตกต่างของระดับข้อมูลในแต่ละกลุ่ม (รวมถึงการประเมินความแตกต่างของรายข้อคำถาม) ได้อย่างมีประสิทธิภาพ ช่วยให้การแปลผลข้อมูลมีความน่าเชื่อถือและเป็นไปตามหลักการทางสถิติอย่างถูกต้อง

วัลลภ รัฐฉัตรานนท์ (2562) ได้ศึกษาวิเคราะห์เรื่องการหาขนาดตัวอย่างที่เหมาะสมสำหรับการวิจัย: มายาคติในการใช้สูตรของทาโร ยามาเน และเครจซี-มอร์แกน โดยมุ่งเน้นสร้างความเข้าใจที่ถูกต้องตามหลักสถิติศาสตร์ในการกำหนดขนาดกลุ่มตัวอย่างสำหรับการวิจัยเชิงปริมาณ งานวิจัยนี้ได้อธิบายถึงเงื่อนไขสำคัญของการใช้สูตรทาโร ยามาเน (Taro Yamane) ว่าเป็นสูตรที่เหมาะสมกับโครงการวิจัยที่ "ทราบจำนวนประชากรเป้าหมายที่แน่นอน (Finite Population)" พร้อมทั้งชี้ให้เห็นว่า เพื่อให้ได้กลุ่มตัวอย่างที่เป็นตัวแทนของประชากรอย่างแท้จริง และมีความน่าเชื่อถือ ผู้วิจัยควรกำหนดระดับความคลาดเคลื่อนที่ยอมรับได้ (Margin of Error) ไว้ที่ไม่เกินร้อยละ 5 ($e = 0.05$) หรือระดับความเชื่อมั่นร้อยละ 95 การตั้งค่าความคลาดเคลื่อนที่สูงกว่านี้อาจทำให้เกิดความลำเอียง (Bias) และทำให้ผลลัพธ์ขาดความน่าเชื่อถือทางสถิติ การนำแนวคิดนี้มาใช้มีข้อดีคือ ช่วยเป็นเกณฑ์มาตรฐานที่ได้รับการยอมรับในการคำนวณและกำหนดขนาดกลุ่มตัวอย่างเป้าหมายของโครงการ (จำนวน 348 คน) ทำให้มั่นใจได้ว่าชุดข้อมูลที่จัดเก็บมามีปริมาณที่ครอบคลุมและเพียงพอต่อการนำไปสอนแบบจำลองต้นไม้ตัดสินใจ (Decision Tree) ซึ่งจะช่วยลดปัญหาข้อมูลไม่เพียงพอหรือแบบจำลองเรียนรู้จำกัด (Overfitting) ได้อย่างมีประสิทธิภาพ

2.4 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.4.1 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

ขั้นตอนการทำเหมืองข้อมูลที่นิยมใช้ในปัจจุบันมีหลายแบบ และรูปแบบที่เป็นที่นิยม คือ CrossIndustry Standard Process for Data Mining (CRISP-DM) ที่มีการพัฒนา เป็น Workflow มาตรฐาน สำหรับการทำเหมืองข้อมูล ประกอบด้วย 6 ขั้นตอนคือ

1) Business Understanding คือ การทำความเข้าใจในงาน ระบุโอกาส และหาปัญหาที่จะเกิดขึ้นกับธุรกิจ กำหนดขอบเขตของข้อมูลที่จะนำวิเคราะห์หาความได้เปรียบทางการตลาด และแก้ไขปัญหาคงที่ ซึ่งต้องสามารถระบุผลลัพธ์ที่มีได้

2) Data Understanding ทำความเข้าใจข้อมูลโดยการรวบรวมข้อมูลที่เกี่ยวข้อง คัดเลือกให้เหลือเพียงข้อมูลที่มีความถูกต้อง และสำคัญต่องานมาทำการวิเคราะห์

3) Data Preparation ทำการแปลงข้อมูล (Raw Data) ให้กลายเป็นข้อมูลที่สามารถ นำมาช่วยในการวิเคราะห์ต่อไปได้ ขั้นตอนนี้จะใช้เวลามากที่สุดในทุกขั้นตอน เพราะคุณภาพของ งานที่ได้จะดีเพียงใดขึ้นอยู่กับคุณภาพข้อมูลที่จัดเตรียมในขั้นนี้ การเตรียมข้อมูล ประกอบด้วย การคัดเลือกข้อมูล การกลั่นกรองข้อมูล และแปลงรูปแบบของข้อมูล

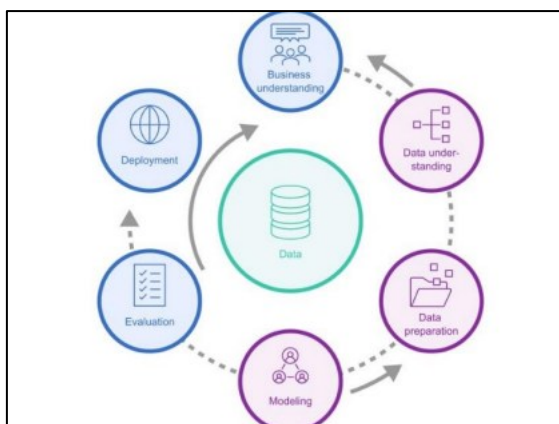
4) Modeling การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูลที่ได้จากขั้นตอนที่ 3 พร้อมทดสอบผลลัพธ์แบบจำลองเพื่อให้ได้คำตอบที่ดีที่สุด บางครั้งอาจมีการย้อนกลับไปปรับการเตรียม ข้อมูลเพื่อให้ได้แบบจำลองที่เหมาะสมที่สุด

5) Evaluation การประเมินผลลัพธ์ที่ได้ก่อนที่จะนำไปใช้จริง ว่าตรงกับวัตถุประสงค์ หรือเป้าหมายที่ได้ตั้งไว้ หรือมีความน่าเชื่อถือมากน้อยเพียงใด หากไม่ได้ผลลัพธ์ตามวัตถุประสงค์ ต้องย้อนกลับไปปรับปรุงแก้ไขการดำเนินงานในขั้นตอนก่อนหน้า

6) Deployment การนำเอาข้อมูลที่เป็นผลลัพธ์จากทั้งหมด มาลองปฏิบัติจริงกับธุรกิจ ในองค์กร โดยแปลงแนวคิดที่มีให้เกิดเป็นสารสนเทศเพื่อให้ผู้บริหาร หรือนักการตลาด เข้าใจ สามารถนำไปใช้ประโยชน์ในทางธุรกิจได้จริง และติดตามประเมินผลที่ได้เพื่อนำกลับไปปรับปรุง Data Mining ต่อเนื่องต่อไป ซึ่งการประเมินผลสามารถทำได้หลายทาง เช่น วัดจากส่วนแบ่งของ ตลาด วัดจากปริมาณลูกค้า หรือ วัดจากกำไรสุทธิ เป็นต้น

ทั้งนี้การทำเหมืองข้อมูลในงานระบบทางธุรกิจ เป็นกระบวนการทางสถิติที่เน้นการจัดการข้อมูลที่มีจำนวนมากในหลากหลายรูปแบบ คัดเลือกเฉพาะข้อมูลที่สำคัญ และจำเป็นต้องนำมาใช้งาน ทำการกำหนดรูปแบบจัดแบ่งกลุ่มลำดับความสำคัญ จากนั้นจึงจะเริ่มค้นหารูปแบบแนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นโดยแต่ละขั้นตอนจะอาศัยผลลัพธ์ จากอีกขั้นตอนหนึ่งกลายเป็นข้อมูลให้ขั้นตอนต่อไป

การทำเหมืองข้อมูลจะช่วยเปลี่ยนข้อมูลดิบให้เป็นสารสนเทศที่มีประโยชน์ การระบุแหล่งข้อมูลที่ถูกต้องจึงเป็นสิ่งที่สำคัญต่อผลลัพธ์ที่ได้ จากการวิเคราะห์ การทำเหมืองข้อมูลจำเป็นต้องอาศัยบุคลากรจากหลายฝ่าย และต้องอาศัยความรู้ จำนวนมากถึงจะได้รับประโยชน์อย่างแท้จริง เพราะสิ่งที่ได้เป็นเพียงตัวเลข และข้อมูล ที่อาจจะนำไปใช้ประโยชน์ได้ หรือใช้ประโยชน์อะไรไม่ได้เลยก็เป็นได้ ผู้ที่ศึกษาการทำเหมืองข้อมูลจึงควรมี ความรู้รอบด้าน และต้องติดต่อกับทุก ๆ ฝ่าย เพื่อให้เข้าใจถึงขอบเขตของปัญหาโดยแท้จริงก่อน เพื่อให้การทำเหมืองข้อมูลเกิดประโยชน์อย่างแท้จริง



ภาพที่ 2.10 แบบจำลอง CRISP-DM

(ที่มา: ist.fraunhofer.de)

2.4.2 เทคนิคต้นไม้ตัดสินใจ (Decision Tree)

เป็นเทคนิคหนึ่งของ Classification ซึ่งเป็นวิธีการแบ่งประเภทหรือแยกหมวดหมู่ข้อมูล โดย Classification นั้นเป็นเทคนิคหนึ่งของเหมืองข้อมูล (Data Mining)

การเรียนรู้แบบต้นไม้ตัดสินใจ (อังกฤษ: decision tree learning) เป็นหนึ่งในวิธีการเรียนรู้ซึ่งใช้ในสถิติ, การเรียนรู้ของเครื่อง และการทำเหมืองข้อมูล โดยพิจารณาการสังเกตการแบ่งแยกข้อมูลโดยพิจารณาข้อมูล

ในการเรียนรู้ของเครื่อง (machine learning) ต้นไม้ตัดสินใจ เป็นโมเดลทางคณิตศาสตร์ที่ใช้ทำนายประเภทของวัตถุโดยพิจารณาจากลักษณะของวัตถุ บัพภายใน (inner node) ของต้นไม้จะแสดงตัวแปร ส่วนกิ่งจะแสดงค่าที่เป็นไปได้ของตัวแปร ส่วนบัพใบ (leaf node) จะแสดงประเภทของวัตถุ

ต้นไม้ตัดสินใจที่บัพใบแสดงถึงข้อมูลที่เป็นข้อมูลไม่ต่อเนื่อง (discrete values) จะเรียกว่าต้นไม้ตัดสินใจแบบจำแนก (classification trees) และต้นไม้ตัดสินใจที่บัพใบเป็นข้อมูลต่อเนื่อง (continuous values) จะเรียกว่าต้นไม้ตัดสินใจแบบถดถอย (regression trees)

ต้นไม้การตัดสินใจในการบริหารธุรกิจ เป็นแผนผังต้นไม้ช่วยในการ ตัดสินใจ โดยแสดงถึงมูลค่าของทรัพยากรที่จะใช้ ความเสี่ยงในการลงทุนและผลลัพธ์ที่มี โอกาสเกิดขึ้น ต้นไม้ตัดสินใจสร้างขึ้นเพื่อช่วยการตัดสินใจเพื่อใช้ในการสร้างแผนงาน นิยมใช้ มากในการบริหารความเสี่ยง (risk management) ต้นไม้ตัดสินใจเป็นส่วนหนึ่งของทฤษฎีการ ตัดสินใจ (decision theory) และ ทฤษฎีกราฟ ต้นไม้ตัดสินใจเป็นวิธีการพื้นฐานอย่างหนึ่ง สำหรับการทำความเข้าใจข้อมูล

ขั้นตอนในการสร้าง Decision Tree เพื่อใช้จำแนกข้อมูล มีดังนี้

- 1) เลือก Attribute ที่ทำหน้าที่เป็น Root Node
- 2) จาก Root Node สร้างเส้นเชื่อมโยงไปยังโหนดลูกจำนวนเส้น เชื่อมโยงจะเท่ากับจำนวนค่าที่เป็นไปได้ทั้งหมดของ Attribute ที่เป็น root node
- 3) ถ้าโหนดลูกเป็นกลุ่มของข้อมูลที่อยู่ในคลาสเดียวกันทั้งหมด ให้หยุด สร้างต้นไม้ แต่ถ้าโหนดลูกมีข้อมูลของหลายคลาสปะปนกันอยู่ ต้อง สร้าง subtree เพื่อจำแนก ข้อมูลต่อไป โดยเลือก subtree มาทำหน้าที่ เป็น root node ของ subtree มาทำซ้ำในขั้นตอนที่ 2.) , 3.) ซึ่งการคำนวณมีดังนี้

เอนโทรปี (Entropy) การสร้างต้นไม้การตัดสินใจจากบนลงล่างด้วยการ ถามว่าลักษณะใดควรจะเป็นรากของต้นไม้การตัดสินใจต้นนี้ และถามซ้ำๆ ไปเรื่อยๆ เพื่อหา ต้นไม้ทั้งต้นด้วยการเขียนโปรแกรมด้วยความสัมพันธ์แบบเวียนเกิด (อังกฤษ: recursion) โดยใน การเลือกว่าลักษณะใดดีที่สุดที่สุคนั้นดูจากค่าของลักษณะเรียกว่าเกนความรู้ (Information gain) ก่อนที่จะรู้จักเกนความรู้จะต้องนิยามค่าหนึ่งที่ใช้บอกความไม่บริสุทธิ์ของข้อมูลก่อน เรียกว่า เอนโทรปี (Entropy) โดยนิยามเอนโทรปีของต้นไม้การตัดสินใจในตัวในเซตของตัวอย่าง S คือ $E(S)$ ดังนี้

$$E(S) = - \sum_{j=1}^n p_S(j) \log_2 p_S(j)$$

เมื่อแทนค่า S คือ ตัวอย่างที่ประกอบด้วยชุดของตัวแปรต้นและตัวแปร ตามหลายๆ กรณี $P_S(j)$ คือ อัตราส่วนของกรณีใน S ที่ตัวแปรตามหรือผลลัพธ์มีค่า j โดยสำหรับต้นไม้การตัดสินใจที่มีผลลัพธ์เป็นแค่เพียงค่าตรรกะ (Boolean) ใช่กับไม่ใช่เหมือนกับที่ยกมาตอนต้นของบทความนั้น จะมีเอนโทรปี คือ

$$E(S) = -p_{yes} \log_2(p_{yes}) - p_{no} \log_2(p_{no})$$

เมื่อพิจารณาเอนโทรปีแล้วจะเห็นว่าเอนโทรปีจะมีค่าอยู่ระหว่าง 0 กับ 1 โดยจะมีค่าเป็นศูนย์เมื่อทุกๆ กรณีมีผลลัพธ์เพียงแบบเดียว เช่น ใช่ทั้งหมด หรือ ไม่ใช่ทั้งหมด และจะมีค่ามาก

ขึ้นเมื่อเริ่มมีค่าที่แตกต่างกันมากขึ้น หรือจะพูดอีกนัยหนึ่งก็คือเอนโทรปีจะมีค่า มากขึ้นหาก ข้อมูลไม่บริสุทธิ์ และจะตัดสินใจได้ว่าผลลัพธ์จะเป็นอะไรเมื่อเอนโทรปี เป็น 0 เท่านั้น

เกนความรู้ (Information Gain) ซึ่งจากการนิยามเอนโทรปีข้างต้น ทำให้ เราสามารถนิยาม ลักษณะของตัวแปรต้นที่ดีได้ โดยตัวแปร A จะเป็นตัวแปรต้นที่ดีก็ต่อเมื่อหา กว่าแบ่งข้อมูล ตัวอย่าง (Example) ออกเป็นชุด มีจำนวนชุดตามจำนวนค่าของ A ที่เป็นไปได้ เพื่อให้แต่ละ กรณี (Instance) ในชุดนั้นมีค่า A เพียงค่าเดียวและค่าเฉลี่ยของเอนโทรปีของชุด ข้อมูลที่ถูกแบ่ง ออก (partition) มานั้นต่ำที่สุด เรียกค่าคาดหวังของการลดลงของเอนโทรปี หลังจากข้อมูลถูก แบ่งด้วย A ว่าเกนความรู้ของ A นิยามโดย

$$Gain(S, A) = E(S) - \sum_{v=value(A)} \frac{|S_v|}{|S|} E(S_v)$$

เมื่อแทนค่า S คือตัวอย่างที่ประกอบด้วยชุดของตัวแปรต้นและตัวแปร ตามหลายๆ กรณี

E คือเอนโทรปีของตัวอย่าง

A คือตัวแปรต้นที่พิจารณา

Value(A) คือเซตของค่าของ A ที่เป็นไปได้

S_v คือตัวอย่างที่ A มีค่า v ทั้งหมด

จะเห็นว่าหากเกนความรู้ของ A ยิ่งมากแสดงว่าหลังจากแบ่งตัวอย่าง S ด้วย A แล้วใน แต่ละชุดที่แบ่งได้จะมี Entropy เข้าใกล้ศูนย์มากยิ่งขึ้น ทำให้ใกล้ที่จะตัดสินใจได้ มากขึ้น เกน ความรู้จึงเป็นค่าที่ดีที่จะบอกความดีของตัวแปรต้นที่นำมาพิจารณา

2.5 บทสรุป

จากการทบทวนแนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้องกับการวิเคราะห์ ข้อมูล เพื่อสร้างเครื่องมือช่วยตัดสินใจสำหรับผู้สนใจเข้าศึกษาในคณะบริหารธุรกิจและศิลป ศาสตร์มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ โดยใช้เทคนิค การทำเหมืองข้อมูล ผู้ศึกษาได้เลือกใช้กระบวนการวิเคราะห์ข้อมูลตามขั้นตอนของ Cross-Industry Standard Process for Data Mining (CRISP-DM) ซึ่งเป็นเทคนิคที่ได้รับการยอมรับในวงวิจัย ด้านการทำ เหมืองข้อมูล โดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เพื่อเครื่องมือช่วยตัดสินใจสำหรับผู้สนใจเข้าศึกษาในคณะบริหารธุรกิจและศิลปศาสตร์มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ จากนั้นนำข้อมูลมาทำการแสดงผลแบบ Visualization ในรูปแบบของแดชบอร์ดและ เผยแพร่ผ่านเว็บไซต์