

บทที่ 2

แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงการเรื่อง การวิเคราะห์ข้อมูลปัญหาสุขภาพจิตในการทำงาน ประจำปี 2557 – 2559 ในบทนี้เป็นการนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้องของการวิเคราะห์ข้อมูลปัญหาสุขภาพจิตในการทำงาน ประจำปี 2557 – 2559 และการแสดงผลข้อมูลบนเว็บไซต์ ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ข้อมูล เพื่อใช้เป็นแนวทางการศึกษาประกอบด้วยรายละเอียดตามลำดับ ดังนี้

2.1 แนวคิด

2.2 ทฤษฎี

2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.4 วรรณกรรมที่เกี่ยวข้อง

2.5 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data analytic)

ในการดำเนินงานเรื่องการวิเคราะห์ข้อมูลปัญหาสุขภาพจิตในการทำงาน ประจำปี 2557 – 2559 และการแสดงผลข้อมูลบนเว็บไซต์ ผู้วิเคราะห์ข้อมูลได้ศึกษาหลัก การและทฤษฎีต่าง ๆ องค์ประกอบหนึ่งที่สำคัญคือการวิเคราะห์ข้อมูล ซึ่งมีรายละเอียด ดังนี้

ความหมายของการวิเคราะห์ข้อมูล

รองศาสตราจารย์มลิวัดย์ ผิวคราม (2559) ได้ให้ความหมายการวิเคราะห์ข้อมูลไว้ว่า การวิเคราะห์ข้อมูล (Analysis) เป็นการจัดระเบียบแยกแยะส่วนต่าง ๆ ของหลักฐาน หรือข้อมูลที่ได้ออกเป็นหมวดหมู่ เพื่อหาคำตอบตามความมุ่งหมาย และตามสมมติฐานที่ได้กำหนดไว้ การวิเคราะห์ข้อมูลนี้เป็นขั้นการทำงานที่ต่อเนื่องมาจากการวัด การนับ และจัดเรียงลำดับข้อมูล ส่วนใหญ่เป็นเรื่องเกี่ยวกับการนำเอาวิธีการทางสถิติมาวิเคราะห์หาค่าตัวแปรหรือหาลักษณะของตัวแปร ผู้วิจัยจะต้องวางแผนและเตรียมการล่วงหน้าตั้งแต่เริ่มทำการวิจัยโดยมีข้อเสนอแนะในการวิเคราะห์ดังนี้

1) กลับไปอ่านจุดมุ่งหมายหรือข้อความที่เป็นปัญหาจนแจ่มแจ้งก่อน

- 2) ดูแต่ละหัวข้อปัญหาว่าต้องการข้อมูลประเภทใด และจะใช้วิธีการสถิติอะไร
- 3) สถิติเหล่านั้นหาได้หรือไม่จากข้อมูล เพื่อไปแก้ปัญหามาจากจุดมุ่งหมายแต่ละข้อ
- 4) เลือกข้อมูลที่ได้มา นำมาจัดเป็นหมวดหมู่ แบ่งตามเนื้อหาของปัญหาแต่ละข้อ
- 5) คำนวณค่าสถิติให้ตรงตามหัวข้อปัญหาที่จะตอบ
- 6) พยายามแปลความหมายของข้อมูลเป็นระยะ ๆ ไป
- 7) พยายามนึกถึงรูปร่างของตารางที่จะเสนอ ลักษณะควรย่อ ลั่น แต่บรรยายความได้มาก
- 8) ถ้าข้อมูลจัดเสนอเป็นกราฟชนิดต่าง ๆ ก็ต้องหาวิธีการทำให้เข้าใจได้ง่ายที่สุดอย่าให้ซับซ้อน

เจริญศักดิ์ แซ่จิ่ง (2562) ได้ให้ความหมายการวิเคราะห์ข้อมูลทางธุรกิจไว้ว่า การวิเคราะห์ข้อมูลทางธุรกิจเป็นการใช้ข้อมูล เทคโนโลยีสารสนเทศ การวิเคราะห์ทาสถิติ วิธีการเชิงปริมาณ และตัวแบบทางคณิตศาสตร์หรือตัวแบบที่อาศัยคอมพิวเตอร์เป็นฐาน เพื่อช่วยให้ผู้บริหารได้รับข้อมูลเชิงลึกที่ได้รับการปรับปรุงที่ดีขึ้นเกี่ยวกับการดำเนินงานทางธุรกิจขององค์กร อันนำไปสู่ความสามารถในการตัดสินใจบนพื้นฐานของข้อเท็จจริงที่ดียิ่งขึ้น

องค์กรต่าง ๆ การประยุกต์ใช้การวิเคราะห์ข้อมูลกับข้อมูลทางธุรกิจเพื่ออธิบายทำนาย ตลอดจนทำให้ดำเนินธุรกิจเกิดความเหมาะสมมากที่สุด แรงขับเคลื่อนที่สำคัญในการผลักดันให้การวิเคราะห์ข้อมูลได้รับความสนใจอย่างมากในปัจจุบันทั้งในภาคธุรกิจและวงการวิชาการ ได้แก่ ความต้องการในการตัดสินใจทางธุรกิจที่ดีขึ้น ความพร้อมของข้อมูลและความเป็นไปได้ในการใช้งานฮาร์ดแวร์และซอฟต์แวร์เพื่อการวิเคราะห์ข้อมูล ตลอดจนการเปลี่ยนแปลงวิธีการตัดสินใจขององค์กรไปเป็นการตัดสินใจชนิดพึ่งพาข้อมูล

สมปอง เขียวช่วยพรหม (2560) ได้ให้ความหมายการวิเคราะห์ข้อมูลเชิงคุณภาพไว้ว่า การวิเคราะห์ข้อมูลเชิงคุณภาพ เป็นการนำข้อมูลที่ได้จากการเครื่องมือที่ใช้ในการวิจัยเชิงคุณภาพมาทำการวิเคราะห์ โดยการวิเคราะห์ข้อมูลเชิงคุณภาพมีเทคนิคที่สำคัญ ดังนี้

- 1) การจำแนกและจัดระบบข้อมูล (Typology and Taxonomy) เป็นการนำข้อมูลที่ได้นำมาจำแนกและจัดหมวดหมู่ออกให้เป็นระบบ เช่น ข้อมูลหมวดบุคคลากร ข้อมูลหมวดงบประมาณ ข้อมูลหมวดวัสดุอุปกรณ์ ข้อมูลหมวดงบประมาณ เป็นต้น
- 2) การวิเคราะห์สรุปอุปนัย (Analytic Induction) เป็นการนำข้อมูลที่ได้จากเหตุการณ์ต่าง ๆ ที่เกิดขึ้น มาวิเคราะห์เพื่อหาบทสรุปร่วมกันของเรื่องนั้น

3) การเปรียบเทียบเหตุการณ์ (Constant Comparison) เป็นการนำข้อมูลที่ได้มาไปเทียบเคียงหรือเปรียบเทียบกับเหตุการณ์อื่น เพื่อหาความเหมือนและความแตกต่างกันที่เกิดขึ้น เช่น เปรียบเทียบหน่วยงานหนึ่งกับอีกหน่วยงานหนึ่งที่ประสบผลสำเร็จทางการบริหาร เป็นต้น

4) การวิเคราะห์ส่วนประกอบ (Componential Analysis) เป็นการนำข้อมูลที่ได้มาทำการวิเคราะห์ออกให้เห็นเป็นส่วนๆ เช่น วิเคราะห์การบริหารงานขององค์การออกเป็น 7 หมวด ตามกรอบของ PMQA เป็นต้น

5) การวิเคราะห์ข้อมูลเอกสาร (Content Analysis) เป็นการนำเอกสารหรือหลักฐานต่าง ๆ มาวิเคราะห์ให้เห็นว่า มุ่งพรรณนาและอธิบายปรากฏการณ์ที่เกิดขึ้น เช่น วิเคราะห์การปกครองสมัย พ.ศ.2475 จากหลักฐานทางประวัติศาสตร์ เป็นต้น

6) การวิเคราะห์สาเหตุและผล (Cause and Effect Analysis) เป็นการนำข้อมูลที่ได้มาวิเคราะห์ให้เห็นว่าจากผลมาจากเหตุ คือ วิเคราะห์ผลที่เกิดขึ้น ย้อนกลับมาให้เห็นว่าเกิดมาจากเหตุปัจจัยใดบ้าง หรือวิเคราะห์เหตุไปหาผล คือ วิเคราะห์จากเหตุไปหาผล คือ วิเคราะห์ให้เห็นว่าเมื่อเหตุนี้เกิดขึ้น ได้นำไปสู่ผลที่เกิดขึ้นอะไรบ้าง

7) การสร้างจินตนาการเชิงสังคมวิทยา (Sociology Imaginary) เป็นการนำข้อมูลที่ได้มาวิเคราะห์โดยเปลี่ยนมุมมองการวิเคราะห์ไปยังมุมมองอื่น ๆ เพื่อดูผลการวิเคราะห์ที่เกิดขึ้นว่าเป็นเช่นใด เช่น เปลี่ยนมุมมองการวิเคราะห์จากมุมมองค่านิยม มาเป็นการวิเคราะห์มุมมองด้านวัฒนธรรม เป็นต้น

หัสพร ทองแดง, พิศุทธิภา เมธิกุล และคณะ (2559) ได้ให้ความหมายการวิเคราะห์ข้อมูลเชิงคุณภาพไว้ว่า การวิเคราะห์เป็นการจำแนกข้อมูลให้เป็นระบบ เพื่อทำให้เกิดความเข้าใจในความหลากหลาย ความหมายและความสัมพันธ์ของข้อมูลในบริบทของสังคมและวัฒนธรรมหนึ่ง ๆ ในการวิเคราะห์ข้อมูลเชิงคุณภาพนี้ หลังจากที่มีการแยกข้อมูลออกได้เป็นหมวดหมู่แล้ว นักวิจัยต้องใช้วิธีการนำเสนอโดยการพรรณนาให้เป็นสภาพเงื่อนไข กระบวนการขั้นตอนการสัมพันธ์ต่าง ๆ รวมทั้งพยายามหาความหมายทางวัฒนธรรมของปรากฏการณ์ในทัศนะของบุคคลหรือกลุ่มตัวอย่างซึ่งเป็นผู้ให้ข้อมูล และอาจมีความเข้าใจและตีความแตกต่างกันออกไปหรือแตกต่างไปจากของผู้วิจัยเอง การวิเคราะห์ข้อมูลสำหรับการวิจัยเชิงคุณภาพ จึงเป็นเสมือนการทำความเข้าใจแบบแผนและความหมายของพฤติกรรมมากกว่าที่จะมุ่งหาว่าระดับความมากน้อยหรือระดับความเข้มข้นของความสัมพันธ์ของปรากฏการณ์ จึงเป็นการวิเคราะห์ข้อมูลที่อาศัยเครื่องมือทางความคิด (Conceptual tools) มากกว่าที่จะใช้เครื่องมือทางสถิติ และที่สำคัญต้องอาศัยความสามารถของนักวิจัยที่จะอธิบายให้เห็นความสัมพันธ์และ

ความหมายของปรากฏการณ์บนพื้นฐานความเข้าใจในวัฒนธรรมและสังคมที่เป็นบริบทของปรากฏการณ์ที่ศึกษา

2.1.2 แนวคิดเกี่ยวกับการประยุกต์ใช้ข้อมูลทางสถิติ

ความหมายของการประยุกต์ใช้ข้อมูลทางสถิติ

ดร.สุทิน ชนะบุญ (2560) ได้ให้ความหมายสถิติไว้ว่า โดยทั่วไปสถิติมีความหมาย 2 นัย คือ ความหมายของตัวเลขข้อมูลสถิติ และสถิติที่เป็นศาสตร์ ดังนี้

1) สถิติที่เป็นตัวเลขหรือข้อมูลสถิติ (Statistics Data) หมายถึง ตัวเลขที่ได้จากการรวบรวมข้อมูล นำมาประมวล วิเคราะห์และสรุปเพื่อใช้บรรยายหรืออธิบายเหตุการณ์ปรากฏการณ์หรือ ตัวแปรต่าง ๆ ที่เราสนใจ เช่น สถิติการเจ็บป่วย 10 อันดับแรกของผู้มารับบริการในแต่ละปีของสถานบริการ อัตราการเกิดอุบัติเหตุในช่วง 10 วันอันตรายในช่วงสงกรานต์ เป็นต้น

2) สถิติที่เป็นศาสตร์ (Statistic) หมายถึง วิชาที่ว่าด้วยการจัดกระทำต่าง ๆ เกี่ยวกับข้อมูลเพื่อให้สามารถบรรยายลักษณะของสิ่งที่ศึกษา ได้แก่ การเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล การนำเสนอข้อมูล และการแปลความหมายของข้อมูล รวมถึงการนำเอาข้อมูลที่รวบรวมได้จากตัวอย่างไปใช้ในการคาดคะเนและตัดสินใจต่าง ๆ ที่เกี่ยวกับประชากร

กัลยา สิ้นชัย (2559) ได้ให้ความหมายสถิติไว้ว่า สถิติเป็นการกระทำกับหลักฐานที่เป็นข้อมูลซึ่งอาจจะเป็นข้อมูลเชิงปริมาณหรือเชิงคุณภาพ โดยการเก็บรวบรวมข้อมูล การนำเสนอข้อมูล การวิเคราะห์โดยใช้หลักการทางคณิตศาสตร์ และการนำผลการวิเคราะห์มาสรุป ประเภทของสถิติ แบ่งได้เป็น 2 ประเภทใหญ่ ๆ ได้แก่

1) สถิติเชิงพรรณนา (Descriptive Statistics) เป็นสถิติที่บรรยายถึงคุณลักษณะของสิ่งที่กำลังต้องการศึกษา ได้แก่ ค่าวัดแนวโน้มสู่ส่วนกลาง (ค่าเฉลี่ยเลขคณิต, มัชฌิมฐาน, ฐานนิยม) ค่าวัดการกระจายข้อมูล (ส่วนเบี่ยงเบนมาตรฐาน, พิสัย)

2) สถิติเชิงอนุมาน (Inferential Statistics) เป็นสถิติที่ศึกษาข้อมูลที่เป็นกลุ่มตัวอย่างเพียงกลุ่มเดียว จากข้อมูลของประชากรทั้งหมด

ธีระพงษ์ กระการดี (2562) ได้ให้ความหมายสถิติไว้ว่า สถิติ คือ ตัวเลขที่แสดงข้อเท็จจริงเกี่ยวกับเรื่องใดเรื่องหนึ่ง เช่น สถิติที่แสดงปริมาณน้ำฝน สถิติอุบัติเหตุ สถิตินักเรียนจำนวนผู้ป่วยเป็นเอดส์ของจังหวัดสุโขทัย โดยสถิติ หมายถึง ศาสตร์ หรือหลักการและระเบียบวิธีทางสถิติที่ว่าด้วยการเก็บรวบรวมข้อมูล การนำเสนอข้อมูล การวิเคราะห์ข้อมูล และการตีความหมายข้อมูล

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

ความหมายของการแสดงผลข้อมูล

อาทิตย์ สิทธิบรรเจิด (2562) ได้ให้ความหมายของการแสดงผลข้อมูลไว้ว่า การแสดงผลข้อมูล Visualization เป็นส่วนประกอบสำคัญใน Cognitive System ซึ่งเป็นส่วนในการแสดงข้อมูลหรือผลลัพธ์ต่าง ๆ ในระหว่างคอมพิวเตอร์และผู้ใช้งานในรูปแบบของภาพ โดยผู้ใช้งานสามารถเรียนรู้และจดจำข้อมูลผ่านการมองเห็นได้มากกว่าการใช้ประสาทสัมผัสอื่น ๆ หรือจะกล่าวได้ว่า Visualization ก็คือ การสร้างมโนภาพของสิ่งต่าง ๆ ที่เราสนใจขึ้นมาในใจ ซึ่งต่อมาได้กลายเป็นการนำภาพมาใช้กับการนำเสนอหรือนำมาเป็นกรอบความคิด ซึ่งได้นำไปใช้ในการสนับสนุนการตัดสินใจ ซึ่งข้อดีของ Visualization มีดังนี้

- 1) ช่วยในการแสดงข้อมูลที่มีปริมาณสูง
- 2) ช่วยในการรับรู้หรือคาดคะเนสิ่งที้อาจจะเกิดขึ้นในอนาคต
- 3) Visualization ไม่เพียงแต่แสดงรายละเอียดข้อมูลในตัวเองเท่านั้นยังสามารถแสดงปัญหาที่เกิดขึ้นได้ด้วย
- 4) ช่วยให้ผู้ใช้งานสามารถเข้าใจหรือวิเคราะห์ข้อมูลได้สะดวกขึ้นไม่ว่าขอบเขตข้อมูลนั้นจะมีขนาดใหญ่หรือเล็ก
- 5) ช่วยให้ผู้ใช้งานสามารถสันนิษฐานข้อมูลได้สะดวกขึ้น

ซึ่ง Visualization เป็นมากกว่า วิธีการทาง Computer Visualization เป็นการนำเสนอข้อมูลแบบหนึ่งที่ทำกรแสดงข้อมูลให้อยู่ในรูปแบบของ Visual Form ซึ่งอาจจะเป็นการแสดงในรูปแบบของรูปภาพ, กราฟ หรือ แผนภาพ ซึ่ง ผลลัพธ์ของการทำ Visualization คือ การนำข้อมูลต่าง ๆ ที่ซ่อนอยู่ในตัวของข้อมูลเองออกมาให้ผู้ใช้งานสามารถสัมผัสได้ ซึ่งกระบวนการต่าง ๆ ในการแสดงหรือ นำข้อมูลต่าง ๆ ออกมา อาจจะถูกแอบซ่อนอยู่หรือผู้ใช้ไม่ทันสังเกตแต่อย่างไรก็ตาม Visualization เป็นสิ่งที่จำเป็นในการค้นหาข้อมูล หรือ ใช้ในการวิเคราะห์ข้อมูลโดยจุดมุ่งหมายของ Visualization ก็คือ การถ่ายทอดข้อมูลไปสู่ระบบการรับรู้โดยภาพของผู้ใช้ระบบ เพื่อช่วยในการลดช่องว่างระหว่างผู้ใช้และตัวข้อมูล และยังช่วยผู้ใช้งานสามารถเรียนรู้ ได้อย่างมีประสิทธิภาพ

Buro Brand Asia (2562) ได้ให้ความหมายของการแสดงผลข้อมูลไว้ว่า กระบวนการ Data Visualization เป็นการนำข้อมูลดิบมาเปลี่ยนเป็นกราฟ แผนภูมิ หรือแม้กระทั่งวิดีโอที่ช่วยอธิบายปริมาณ ตัวเลข และช่วยให้คุณได้ข้อมูลเชิงลึกจากข้อมูลเหล่านั้น การใช้ Data Visualization เปลี่ยนแนวคิดของเราเกี่ยวกับการสร้างคุณค่าให้กับข้อมูล ทำให้เราค้นพบรูปแบบใหม่และมองเห็นแนวโน้มของข้อมูลมากขึ้น

Patipol Treerojporn (2562) ได้ให้ความหมายของการแสดงผลข้อมูลไว้ว่า Data Visualization หรือ Information Visualization เป็นเครื่องมือสำคัญที่ช่วยให้เห็นภาพของข้อมูลเชิงลึกที่ผ่านการเก็บข้อมูล การจัดการข้อมูล การวิเคราะห์ข้อมูล แล้วรวบรวมได้เป็นชุดข้อมูล หากต้องการประสบความสำเร็จในโลกธุรกิจที่มีการแข่งขันสูง ความสามารถในการตัดสินใจที่ถูกต้องในเวลาที่เหมาะสม โดยอาศัยข้อมูลเชิงลึกที่รัดกุมเป็นสิ่งสำคัญ ปัจจุบันองค์กรธุรกิจส่วนใหญ่ในประเทศไทย เริ่มใช้ประโยชน์จากการทำ Data Visualization เพราะว่าเครื่องมือนี้เป็นกุญแจสำคัญในการแสดงผลของข้อมูลให้เข้าใจง่าย ซึ่งจะช่วยให้ทุกแผนกในฝ่ายงานได้ใช้ประโยชน์ ในการหาแนวโน้มและความสัมพันธ์เชื่อมโยง อีกทั้งทำให้ทราบข้อมูลได้เร็วขึ้น ส่งผลให้การทำงานขององค์กรมีประสิทธิภาพเพิ่มขึ้น จนดำเนินการถึงการวางกลยุทธ์ภายในองค์กร และสามารถนำไปใช้กับการทำเว็บไซต์ได้ Data Visualization ช่วยให้การสื่อสารข้อมูล มีความรวดเร็วและเข้าใจง่ายมากขึ้น

พื้นฐานของการสร้าง Data Visualization คือ การ Mapping ส่วนข้อมูล กับส่วนของ Graphic เข้าด้วยกัน ซึ่งตอนนี้มีโปรแกรมสำเร็จรูปในการสร้าง Data Visualization หลากหลาย โปรแกรม มีฟังก์ชันการใช้งานที่เข้าใจง่าย เช่น การสร้างฟิลเตอร์ การออกแบบเพื่อให้งานการวิเคราะห์ข้อมูลมีความยืดหยุ่น เป็นต้น ทำให้ข้อเสนอคุณมีคุณค่ามากขึ้น ข้อมูลแบบ Data Visualization นั้นมีพลังทางการตลาดอย่างมาก เพราะมันสามารถแสดงผลลัพธ์และดึงดูดความสนใจของผู้ชมได้เป็นอย่างดี นอกจากนี้ยังทำให้พวกเขาจินตนาการตามข้อมูลที่ได้รับได้อีกด้วย ลองนำการแสดงข้อมูลแบบ Data Visualization มาใช้ บนหน้าแรกของเว็บไซต์หรือบล็อกเพื่อโปรโมทโปรโมชันหรือข้อเสนอทางการตลาดของธุรกิจ แล้วคุณจะพบว่าผู้ชมมีการตอบสนองกับข้อความเหล่านี้มากขึ้น มากกว่าการเสนอขายด้วยการใช้ตัวหนังสือเพียงอย่างเดียว ซึ่งตอนนี้มีโปรแกรมสำเร็จรูปในการสร้าง Data Visualization หลากหลายโปรแกรม มีฟังก์ชันการใช้งานที่เข้าใจง่าย เช่น การสร้างฟิลเตอร์ การออกแบบเพื่อให้งานการวิเคราะห์ข้อมูลมีความยืดหยุ่น เป็นต้น

2.1.3.1 ประโยชน์ของ Data Visualization

1) เล่าเรื่องได้ดีกว่า

หนึ่งในสาเหตุที่สื่อโซเชียลมีเดียได้รับความนิยม เป็นเพราะรูปภาพสามารถบอกเล่าเรื่องราวได้ดีกว่าแผนภูมิ ข้อความบนเว็บไซต์ หรือแม้แต่พรีเซนเทชัน ลองค้นหาแนวคิดที่แท้จริงที่คุณต้องการนำเสนอว่าคุณต้องการพูดถึงเรื่องอะไร สัมผัสการใช้แผนภูมิหรือกราฟไปก่อน แต่หันมาเลือกวิธีแสดงข้อมูลแบบ Data Visualization ที่มีความเป็นพลวัต

(Dynamic) ที่ผู้ชมสามารถรับรู้เรื่องราวที่คุณต้องการเล่าได้ง่ายกว่า เพราะข้อมูลเหล่านี้จะช่วยให้ดวงตาและสมองทำงานอย่างสัมพันธ์กัน

2) ทำให้ข้อเสนอคุณมีคุณค่ามากขึ้น

ข้อมูลแบบ Data Visualization นั้นมีพลังทางการตลาดอย่างมาก เพราะมันสามารถแสดงผลลัพธ์และดึงความสนใจของผู้ชมได้เป็นอย่างดี นอกจากนี้ยังทำให้พวกเขาจินตนาการตามข้อมูลที่ได้รับได้อีกด้วย ลองนำการแสดงข้อมูลแบบ Data Visualization มาใช้บนหน้าแรกของเว็บไซต์หรือบล็อกเพื่อโปรโมทโปรโมชันหรือข้อเสนอทางการตลาดของธุรกิจ แล้วคุณจะพบว่าผู้ชมมีการตอบสนองกับข้อความเหล่านี้มากขึ้น มากกว่าการเสนอขายด้วยการใช้ตัวหนังสือเพียงอย่างเดียว

3) ช่วยให้สื่อสารข้อมูลได้เร็วขึ้น

ยกตัวอย่าง เช่น การติดตามข่าวเรื่องโลกร้อน บางครั้งข้อมูลสถิติด้านการเปลี่ยนแปลงของสภาพอากาศก็เป็นเรื่องยากที่เราจะติดตามและทำความเข้าใจกับตัวเลขที่เปลี่ยนแปลงแทบจะตลอดเวลาและมีค่ามากมายที่สัมพันธ์กัน แต่มีนักวิทยาศาสตร์รายหนึ่งเลือกใช้วิธีนำเสนอข้อมูลเหล่านั้นด้วย Data Visualization ผ่านทวิตเตอร์ ทำให้เราสามารถย่อยข้อมูลที่เป็นสาระสำคัญได้ง่ายขึ้น เข้าใจมากขึ้น ผ่านการแยกสีเส้นที่แตกต่างกันแทนข้อมูลอุณหภูมิของแต่ละเดือน และนี่ก็คือประโยชน์ของข้อมูลภาพที่ช่วยให้เราสามารถประมวลผลสิ่งที่ผู้นำเสนอต้องการได้อย่างรวดเร็ว

4) สร้างแรงจูงใจให้กับทีม

ในการทำงาน เป็นเรื่องง่ายที่ทีมจะไขว่เขวหากไม่มีการวัดความคืบหน้าของงานหรือการกำหนดเป้าหมายดังนั้นทีมจึงต้องระลึกอยู่เสมอว่าองค์กรต้องการอะไรและเป้าหมายของพวกคุณคือที่ไหน ข้อมูลด้านการขายและการนำเสนอผ่านบอร์ดที่ออกแบบเองอาจทำให้พลาดข้อมูลเชิงลึกที่สำคัญหลายอย่าง

ขณะที่การใช้ข้อมูลรูปภาพนั้นมีความครอบคลุมมากกว่า นอกจากนี้มันยังสามารถนำมาใช้ในการวัดผลและติดตามการขายเพื่อเพิ่มประสิทธิภาพพนักงานและสร้างความเติบโตของทีมได้ จากกรณีตัวอย่างกรณีหนึ่งพบว่าการใช้ข้อมูลรูปภาพในการขายและการนำเสนอ ช่วยให้ออดขายเติบโตขึ้นถึง 24%

5) ทำให้สามารถตัดสินใจได้อย่างรวดเร็ว

ข้อมูลรูปภาพช่วยให้สมองประมวลผลได้เร็วขึ้น เมื่ออยู่ในสถานการณ์ที่ต้องใช้ทักษะการตัดสินใจอย่างรวดเร็ว เช่น การบริหารจัดการพนักงาน หรือการจัดการสินค้าคงคลัง ซึ่งข้อมูล Data Visualization จะช่วยให้คุณมองภาพรวมได้ดีกว่าการกวาดตาผ่านข้อมูล

แผนภูมิจำนวนมาก หรือแม้แต่ในช่วงเวลาปกติที่ไม่จำเป็นต้องตัดสินใจแบบเร่งด่วน ข้อมูลรูปภาพก็มีประโยชน์ในการทำการตัดสินใจครั้งสำคัญเช่นกัน

และนี่ก็คือประโยชน์ของข้อมูลแบบ Data Visualization ที่มีต่อธุรกิจและทีม มันคือการนำเสนอข้อมูลรูปแบบใหม่ที่มีความใกล้ชิดกับธรรมชาติการรับรู้ของมนุษย์มากที่สุด เพียงต้องลองนำมาประยุกต์ใช้ จะพบว่าข้อมูลรูปภาพทำอะไรได้มากกว่าที่เราคิด

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) คือ การวิเคราะห์ข้อมูลจากข้อมูลจำนวนมาก (big data) เพื่อหาความสัมพันธ์ของข้อมูลที่ซ่อนอยู่ โดยทำการจำแนกประเภท รูปแบบ เชื่อมโยงข้อมูลที่มีความสัมพันธ์กัน และหาความน่าจะเป็นที่จะเกิดขึ้น เพื่อให้ได้องค์ความรู้ใหม่ที่สามารถนำไปใช้ประกอบการตัดสินใจในด้านต่าง ๆ เช่น ตลาดหลักทรัพย์, ทางธุรกิจ, ทางด้านการแพทย์, ยุทธศาสตร์ทหาร เป็นต้น เปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลอย่างง่าย ๆ มาสู่การจัดเก็บในรูปแบบฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนถึงการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล

2.2.1.1 ขั้นตอนการทำ Data Mining

- 1) Data cleaning เป็นขั้นตอนสำหรับการคัดข้อมูลที่ไม่เกี่ยวข้องออกไป
- 2) Data Integration เป็นขั้นตอนการรวมข้อมูลที่มีหลายแหล่งให้เป็นข้อมูลชุดเดียวกัน
- 3) Data Selection เป็นขั้นตอนการดึงข้อมูลสำหรับการวิเคราะห์จากแหล่งที่บันทึกไว้
- 4) Data Transformation เป็นขั้นตอนการแปลงข้อมูลให้เหมาะสมสำหรับการใช้งาน
- 5) Data mining เป็นขั้นตอนการค้นหารูปแบบที่เป็นประโยชน์จากข้อมูลที่มีอยู่
- 6) Pattern Evaluation เป็นขั้นตอนการประเมินรูปแบบที่ได้จากการทำเหมืองข้อมูล

7) Knowledge Representation เป็นขั้นตอนการนำเสนอความรู้ที่ค้นพบ โดยใช้เทคนิคในการนำเสนอเพื่อให้เข้าใจ



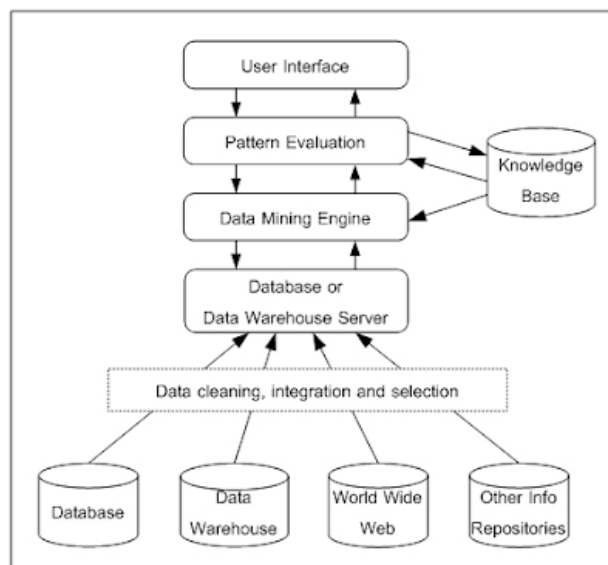
ภาพที่ 2.1 ขั้นตอนการทำ Data Mining

(ที่มา : <https://th.wikipedia.org/wiki/การทำเหมืองข้อมูล>)

2.2.1.2 ส่วนประกอบของการทำเหมืองข้อมูล

สถาปัตยกรรมของระบบการทำเหมืองข้อมูล ประกอบด้วยส่วนประกอบที่

สำคัญ



ภาพที่ 2.2 สถาปัตยกรรมระบบการทำเหมืองข้อมูล

(ที่มา : <https://th.wikipedia.org/wiki/การทำเหมืองข้อมูล>)

- 1) Database, Data Warehouse, World Wide Web และ Other Info Repositories เป็นแหล่งข้อมูลสำหรับการทำเหมืองข้อมูล
- 2) Database หรือ Data Warehouse Server ทำหน้าที่นำเข้าสู่ข้อมูลตามคำขอของผู้ใช้
- 3) Knowledge Base ได้แก่ ความรู้เฉพาะด้านในงานที่ทำจะเป็นประโยชน์ต่อการสืบค้น หรือประเมินความน่าสนใจของรูปแบบผลลัพธ์ที่ได้
- 4) Data Mining Engine เป็นส่วนประกอบหลักประกอบด้วยโมดูลที่รับผิดชอบงานทำเหมืองข้อมูลประเภทต่าง ๆ ได้แก่ การหากฎความสัมพันธ์ การจำแนกประเภท การจัดกลุ่ม
- 5) Pattern Evaluation Module ทำงานร่วมกับ Data Mining Engine โดยใช้มาตรวัดความน่าสนใจในการกลั่นกรองรูปแบบผลลัพธ์ที่ได้ เพื่อให้การค้นหามุ่งเน้นเฉพาะรูปแบบที่น่าสนใจ
- 6) Graphic User Interface ส่วนติดต่อประสานระหว่างผู้ใช้กับระบบการทำเหมืองข้อมูล ช่วยให้ผู้ใช้สามารถระบุงานทำเหมืองข้อมูลที่ต้องการทำ ดูข้อมูลหรือโครงสร้างการจัดเก็บข้อมูล ประเมินผลลัพธ์ที่ได้

2.2.2 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

2.2.2.1 หลักในการออกแบบเว็บไซต์

เว็บไซต์เป็นสื่อที่ได้รับความนิยมอย่างมากบนอินเทอร์เน็ต ซึ่งเว็บไซต์เป็นสื่อที่อยู่ในความควบคุมของผู้ใช้โดยสมบูรณ์ กล่าวคือ ผู้ใช้สามารถตัดสินใจเลือกได้ว่า จะดูเว็บไซต์ใดและจะไม่เลือกดูเว็บไซต์ใด ได้ตามต้องการ จึงทำให้ผู้ใช้ไม่มีความอดทนต่ออุปสรรคและปัญหาที่เกิดจากการออกแบบเว็บไซต์ผิดพลาดถ้าผู้ใช้เห็นว่าเว็บที่กำลังดูอยู่นั้นไม่มีประโยชน์ต่อตัวเขา หรือไม่เข้าใจว่าเว็บไซต์นี้จะใช้งานอย่างไร เขาก็สามารถที่จะเปลี่ยนไปดูเว็บไซต์อื่น ๆ ได้อย่างรวดเร็ว เนื่องจากในปัจจุบันมีเว็บไซต์อยู่มากมาย และยังมีเว็บไซต์ที่เกิดขึ้นใหม่ทุกวันผู้ใช้งานจึงมีทางเลือกมากขึ้น และสามารถเปรียบเทียบคุณภาพของเว็บไซต์ต่าง ๆ ได้เอง

เว็บไซต์ที่ได้รับการออกแบบอย่างสวยงาม มีการใช้งานที่สะดวก ย่อมได้รับความสนใจจากผู้ใช้งานมากกว่าเว็บไซต์ที่ดูสับสนวุ่นวาย มีข้อมูลมากมายแต่หาอะไรไม่เจอ นอกจากนี้ยังใช้เวลาในการแสดงผลแต่ละหน้านานเกินไป ซึ่งปัญหาเหล่านี้ล้วนเป็นผลมาจากการออกแบบเว็บไซต์ไม่ดีทั้งสิ้น

ดังนั้น การออกแบบเว็บไซต์จึงเป็นกระบวนการสำคัญในการสร้างเว็บไซต์ให้ประทับใจผู้ใช้ ทำให้เขาอยากกลับมาเว็บไซต์เดิมอีกในอนาคต ซึ่งนอกจากต้องพัฒนาเว็บไซต์ที่ดีมีประโยชน์แล้ว ยังต้องคำนึงถึงการแข่งขันกับเว็บไซต์อื่น ๆ อีกด้วย

หลักในการออกแบบโครงสร้างเว็บไซต์ ควรพิจารณาดังนี้

1) กำหนดวัตถุประสงค์ โดยพิจารณาว่าเป้าหมายของการสร้างเว็บไซต์นี้ทำเพื่ออะไร

2) ศึกษาคุณลักษณะของผู้ที่เข้ามาใช้ว่ากลุ่มเป้าหมายใดที่ผู้สร้างต้องการสื่อสาร ข้อมูลอะไรที่พวกเขาต้องการโดยขั้นตอนนี้ควรปฏิบัติควบคู่ไปกับขั้นตอนที่หนึ่ง

3) วางแผนเกี่ยวกับการจัดรูปแบบโครงสร้างเนื้อหาสาระ การออกแบบเว็บไซต์ต้องมีการจัดโครงสร้างหรือจัดระเบียบข้อมูลที่ชัดเจน การที่เนื้อหาไม่ต่อเนื่องไปไม่สิ้นสุดหรือกระจายมากเกินไป อาจทำให้เกิดความสับสนต่อผู้ใช้ได้ ฉะนั้นจึงควรออกแบบให้มีลักษณะที่ชัดเจนแยกย่อยออกเป็นส่วนต่าง ๆ จัดหมวดหมู่ในเรื่องที่สัมพันธ์กัน รวมทั้งอาจมีการแสดงให้ผู้ใช้เห็นแผนที่โครงสร้างเพื่อป้องกันความสับสนได้

4) กำหนดรายละเอียดให้กับโครงสร้าง ซึ่งพิจารณาจากวัตถุประสงค์ที่ตั้งไว้ โดยตั้งเกณฑ์ในการใช้ เช่น ผู้ใช้ควรทำอะไรบ้าง จำนวนหน้าควรมีเท่าใด มีการเชื่อมโยงมากน้อยเพียงใด

5) หลังจากนั้น จึงทำการสร้างเว็บไซต์แล้วนำไปทดลองเพื่อหาข้อผิดพลาดและทำการแก้ไขปรับปรุง แล้วจึงนำเข้าสู่เครือข่ายอินเทอร์เน็ตเป็นขั้นสุดท้าย

2.2.2.2 องค์ประกอบในการออกแบบเว็บไซต์

1) ความเรียบง่าย เว็บไซต์ที่ดีควรมีรูปแบบที่เรียบง่ายและไม่ซับซ้อน เพื่อให้ผู้ชมสามารถใช้งานเว็บไซต์ได้อย่างสะดวกมากขึ้น โดยเฉพาะพวกกราฟิกทั้งหลายจะต้องไม่ใช่ตัวอักษรที่เคลื่อนไหวอยู่ตลอดเวลา และไม่มีสีล้นที่ดูแสบตาจนเกินไป

2) ความสม่ำเสมอ คือการเลือกรูปแบบ กราฟิก โทนมสี และการตกแต่งหรือการแสดงผลต่าง ๆ ในเว็บไซต์ให้เป็นรูปแบบเดียวกันหรือคล้ายคลึงกันตลอดทั้งเว็บ

3) ความเป็นเอกลักษณ์ เว็บไซต์ควรมีเอกลักษณ์เฉพาะตัว ที่สามารถบ่งบอกได้ถึงความเป็นบริษัท องค์กรหรือแบรนด์ต่าง ๆ ตัวอย่างเช่น เว็บไซต์ของ 1 Belief จะมีสีฟ้า ที่เป็นเอกลักษณ์ของบริษัทอยู่บนเว็บ

4) เนื้อหา โดยเนื้อหาที่นำมาลงในเว็บ ควรเป็นเนื้อหาที่มีความเกี่ยวข้องกับเว็บ หรืออาจเป็นเนื้อหาที่ได้สาระ มีประโยชน์ สามารถดึงดูดความสนใจของผู้คนได้ดี และที่สำคัญจะต้องมีความถูกต้อง สมบูรณ์และมีความทันสมัย

5) ระบบเนวิเกชัน ควรออกแบบให้สามารถใช้งานได้ง่ายและสะดวก สื่อความหมายต่าง ๆ และอธิบายได้อย่างชัดเจน รวมถึงต้องมีรูปแบบ และลำดับรายการที่มีความสม่ำเสมอ

6) ลักษณะเด่น ส่วนนี้จะถือเป็นหน้าตาของเว็บไซต์เพื่อใช้ในการดึงดูดลูกค้า อาจออกแบบลักษณะเด่นของเว็บให้ตรงกับความชอบส่วนใหญ่ของกลุ่มเป้าหมาย หรือจะออกแบบให้สัมพันธ์ประเภทของเว็บ และคุณภาพขององค์ประกอบต่าง ๆ บนเว็บ

7) การใช้งานที่ไม่จำกัด การทำเว็บไซต์ให้รองรับการเข้าใช้งานจากหลายระบบ ไม่ว่าจะเป็นการเข้าใช้งานจากเครื่อง PC สมาร์ทโฟน หรือการใช้ Browser ต่าง ๆ ในการเข้าใช้งาน

8) คุณภาพในการออกแบบ จำเป็นต้องทำเว็บไซต์ให้มีคุณภาพมากที่สุด ไม่ว่าจะเป็นในเรื่องของการเรียบเรียงเนื้อหาอย่างรอบคอบ การตรวจสอบความถูกต้องและการทำให้เว็บไซต์มีความน่าเชื่อถือ

9) การเชื่อมโยงไปยังลิงค์ต่าง ๆ ซึ่งจะต้องเชื่อมโยงไปยังหน้าเว็บที่มีอยู่จริง และมีเนื้อหาที่เกี่ยวข้องกัน และควรหมั่นตรวจสอบอยู่เสมอ ว่าระบบการเชื่อมโยงยังคงทำงานได้ตามปกติและมีความถูกต้อง แม่นยำ อยู่หรือไม่

2.2.2.3 ส่วนประกอบสำคัญของหน้าเว็บเพจ

บนหน้าเว็บเพจ จะมีส่วนประกอบสำคัญที่จำเป็นต้องมีอยู่ 3 ส่วน ได้แก่

1) ส่วนหัวของหน้า (Header)

อยู่ตอนบนสุดของหน้าและเป็นส่วนที่สำคัญที่สุด โดยจะต้องทำให้สามารถดึงดูดผู้ชมให้รู้สึกอยากติดตามเนื้อหาในเว็บไซด์ต่อไป ซึ่งส่วนใหญ่มักจะมีการใส่ภาพกราฟิกให้ดูสวยงาม สิ่งสำคัญหลัก ๆ เลย ก็คือ โลโก้ ชื่อเว็บไซต์และเมนูหลักที่สามารถลิงค์ไปยังเนื้อหาในหน้าเว็บเพจต่าง ๆ ได้

2) ส่วนของเนื้อหา (Body)

อยู่บริเวณตอนกลางของหน้าเว็บ โดยจะแสดงข้อมูลเกี่ยวกับเนื้อหาบนเว็บแบบคร่าวๆ ซึ่งก็จะมีข้อความ กราฟิก ตารางข้อมูลหรือวิดีโอประกอบอยู่ และหากมีเมนูแบบ

เฉพาะกลุ่มก็จะถูกจัดไว้ในหน้านี้เช่นกัน และที่สำคัญเนื้อหาในส่วนนี้ควรจะต้องมีความกระชับ เข้าใจง่าย มีการใช้รูปแบบตัวอักษรแบบเรียบง่ายและเป็นระเบียบ

3) ส่วนท้ายของหน้า (Footer)

อยู่ล่างสุดของหน้าเว็บ ซึ่งจะมีหรือไม่มีก็ได้ ส่วนนี้จะแสดงถึงข้อมูลต่าง ๆ เพิ่มเติมเข้าไป เช่น ข้อความที่แสดงถึงการเป็นลิขสิทธิ์ ข้อมูลเจ้าของเว็บไซต์ วิธีการติดต่อและ คำแนะนำต่าง ๆ เกี่ยวกับการใช้งานเว็บไซต์อย่างถูกต้อง เป็นต้น

2.2.2.4 การเลือกใช้สีสำหรับการออกแบบเว็บไซต์

การเลือกใช้สีในการออกแบบเว็บไซต์มีความสำคัญเป็นอย่างมาก เพราะสีสามารถกำหนดอารมณ์ ความรู้สึกและกระตุ้นการรับรู้ทางด้านจิตใจของมนุษย์ได้ดี ดังนั้นสีที่ใช้จึงต้องมีความสอดคล้องกับเนื้อหาและจุดประสงค์ของเว็บ ว่าต้องการให้ผู้เข้าชมรู้สึกอย่างไรต่อเนื้อหาที่ได้อ่าน โดยรูปแบบของสีที่สายตาของมนุษย์สามารถมองเห็นได้ก็แบ่งออกเป็น 3 กลุ่มดังต่อไปนี้

1) สีโทนร้อน (Warm Colors) เป็นสีแห่งความอบอุ่น ปลอดภัยและกระตุ้นความสุขได้ดี ซึ่งจะทำให้ผู้เข้าชมรู้สึกมีชีวิตชีวาและมีแรงผลักดันมากขึ้น อีกทั้งยังช่วยดึงดูดให้ผู้ชมรู้สึกอยากติดตามเนื้อหามากขึ้น

2) สีโทนเย็น (Cool Colors) เป็นสีแห่งความสุภาพและความอ่อนโยน ทำให้ผู้ชมรู้สึกผ่อนคลายและเพลิดเพลินมากขึ้น และยังสามารถใช้โน้มน้าวจากในระยะไกลได้อีกด้วย

3) สีโทนกลาง (Neutral Colors) สีเหล่านี้มักจะถูกนำไปผสมกับสีอื่น ๆ เพื่อให้เกิดสีที่เป็นกลางมากขึ้น และให้ความรู้สึกที่เป็นธรรมชาติ

สีสามารถสื่อถึงอารมณ์ความรู้สึกและสื่อความหมายของเว็บไซต์นั้น ๆ ได้อย่างชัดเจน ดังนั้นหากเลือกใช้สีไม่เหมาะสมกับเนื้อหาหรือจุดประสงค์ของเว็บ ก็จะทำให้เว็บดูไม่น่าสนใจ ขาดความน่าเชื่อถือและทำให้ผู้ที่เคยเข้ามาใช้บริการไม่คิดจะกลับมาใช้บริการอีก

2.2.3 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS

CSS ย่อมาจาก Cascading Style Sheet มักเรียกโดยย่อว่า "สไตลชีต" คือภาษาที่ใช้เป็นส่วนของการจัดรูปแบบการแสดงผลเอกสาร HTML โดยที่ CSS กำหนดกฎเกณฑ์ในการระบุรูปแบบ (หรือ "Style") ของเนื้อหาในเอกสาร อันได้แก่ สีของข้อความ สีพื้นหลัง ประเภทตัวอักษร และการจัดวางข้อความ ซึ่งการกำหนดรูปแบบ หรือ Style นี้ใช้หลักการของการแยก

เนื้อหาเอกสาร HTML ออกจากคำสั่งที่ใช้ในการจัดรูปแบบการแสดงผล กำหนดให้รูปแบบของการแสดงผลเอกสาร ไม่ขึ้นอยู่กับเนื้อหาของเอกสาร เพื่อให้ง่ายต่อการจัดรูปแบบการแสดงผล ลัฟร์ของเอกสาร HTML โดยเฉพาะในกรณีที่มีการเปลี่ยนแปลงเนื้อหาเอกสารบ่อยครั้ง หรือต้องการควบคุมให้รูปแบบการแสดงผลเอกสาร HTML มีลักษณะของความสม่ำเสมอทั่วกันทุกหน้าเอกสารภายในเว็บไซต์เดียวกัน โดยกฎเกณฑ์ในการกำหนดรูปแบบ (Style) เอกสาร HTML ถูกเพิ่มเข้ามาครั้งแรกใน HTML 4.0 เมื่อปีพ.ศ. 2539 ในรูปแบบของ CSS level 1 Recommendations ที่กำหนดโดย องค์กร World Wide Web Consortium หรือ W3C

2.2.3.1 ประโยชน์ของ CSS

ภาษา CSS (Cascading Style Sheets) มีประโยชน์หลายอย่างเลยทีเดียวซึ่งทำให้การพัฒนาเว็บเพจด้วยภาษา HTML เป็นเรื่องที่ย่างมากขึ้นและการเขียน CSS ที่ดีมีผลดีต่อการดันอันดับเว็บไซต์ ด้วย SEO อย่างไร

1) ภาษา CSS จะช่วยในการจัดรูปแบบแสดงผลให้กับภาษา HTML ซึ่งจะช่วยลดการใช้ภาษา HTML ให้น้อยลง โดยเหลือเพียงแต่ส่วนที่เป็นเอกสารที่เป็นภาษา HTML เท่านั้นทำให้มีการแก้ไขและทำความเข้าใจได้ง่ายขึ้น

2) ทำให้ขนาดไฟล์ HTML น้อยลงเนื่องจาก ภาษา CSS จะช่วยลดการใช้ภาษา HTML ลงทำให้ขนาดไฟล์นั้นก็เล็กลงไปด้วยเช่นกัน

3) ภาษา CSS เป็นภาษา Style Sheets โดย Style Sheets ชุดเดียวสามารถใช้กำหนดรูปแบบการแสดงผลให้เอกสาร HTML ทั้งหมด หรือทุกหน้ามีผลเหมือนกันได้ จึงทำให้เวลาที่มีการแก้ไขก็จะแก้ไขได้ง่ายขึ้นเพียงแก้ไข Style Sheets ที่ใช้งานเพียงชุดเดียวเท่านั้น

4) ทำให้เว็บไซต์มีมาตรฐานเพราะการใช้งาน CSS นั้นจะทำให้การแสดงผลในสื่อต่าง ๆ ถูกปรับเปลี่ยนไปได้อย่างเหมาะสม เช่น การแสดงผลบนหน้าจอ และการแสดงผลในมือถือ

5) CSS สามารถที่จะใช้งานได้หลากหลาย เว็บเบราว์เซอร์ ทำให้การใช้งานนั้นสะดวกมากยิ่งขึ้น

6) CSS สามารถกำหนดแยกไว้ต่างหากจากไฟล์เอกสาร HTML และสามารถนำมาใช้ร่วมกับเอกสารหลายไฟล์ได้ การแก้ไขก็แก้เพียง จุดเดียวก็มีผลกับเอกสารทั้งหมด

2.2.4 ทฤษฎีเกี่ยวกับการ Visualization

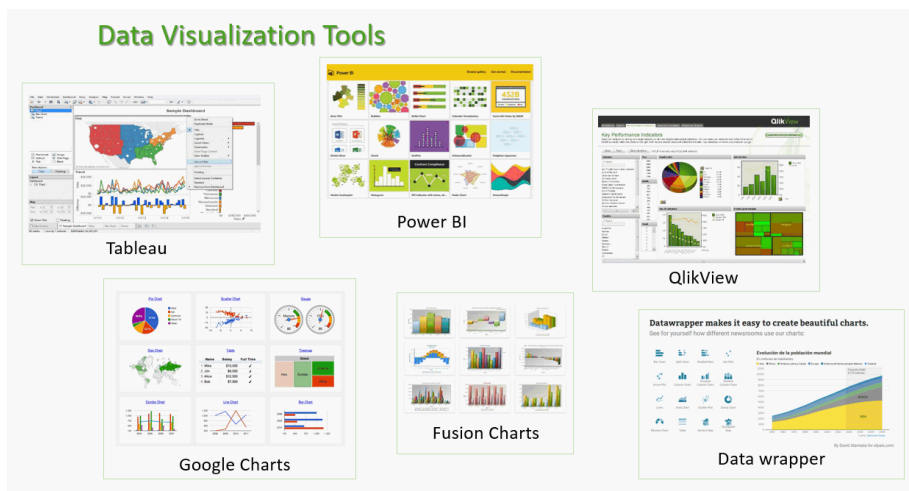
Data Visualization หรือ Information Visualization คือการนำข้อมูลในเชิงปริมาณ ทั้งที่จัดเก็บไว้ในรูปแบบของข้อมูลจำนวนน้อย และข้อมูลจำนวนมาก (Big Data) มา

ประมวลผล จากนั้นจึงนำมาแสดงผลในรูปแบบของกราฟ แผนภูมิอินฟอร์เมชันกราฟิก หรือแม้กระทั่งอินเทอร์แอคทีฟกราฟิก ที่ผู้บริโภครสามารถคลิกหรือมีปฏิสัมพันธ์กับกราฟิกนั้น ๆ ได้ ซึ่งปัจจุบันสื่อหลายสำนักในประเทศไทย เริ่มมีการนำเสนอข่าวโดยการนำข้อมูลแบบประยุกต์ด้วยภาพ (Data Visualization) การถ่ายทอดข้อมูลในเชิงปริมาณที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ ซึ่งคำว่า “ประสิทธิภาพ” ในที่นี้หมายถึงมีความชัดเจน (Clarity), มีความแม่นยำ (Precision), และมีประสิทธิภาพ (Efficiency) หากไม่มีการทำ Data Visualization แล้ว อาจทำให้เราไม่สามารถค้นพบนัยยะของข้อมูลในแง่ของแนวโน้ม, รูปแบบพฤติกรรม, และความสัมพันธ์เชื่อมโยงได้

พื้นฐานของการสร้าง Data Visualization คือ การ Mapping ส่วนข้อมูล กับส่วนของ Graphic เข้าด้วยกัน ซึ่งตอนนี้มีโปรแกรมสำเร็จรูปในการสร้าง Data Visualization หลากหลายโปรแกรม มีฟังก์ชันการใช้งานที่เข้าใจง่าย เช่น การสร้างฟิลเตอร์ การออกแบบเพื่อให้งานการวิเคราะห์ข้อมูลมีความยืดหยุ่น เป็นต้น

2.2.4.1 โปรแกรมสำหรับการสร้างแดชบอร์ด (Data Visualization)

Tools ที่นิยมใช้ในตอนนี้ได้แก่ Tableau, Microsoft Power BI, Qlik View, Google Charts, Fusion Charts, Data wrapper และอื่น ๆ อีกมากมาย



ภาพที่ 2.3 Tools Data Visualization

(ที่มา : www.autosoft.in.th/data-Visualization/การนำเสนอแผนภาพข้อมูล-data-visu/)

2.2.4.2 การนำเสนอแบบทิศทางหรือแนวโน้ม (Trending)

การนำเสนอแบบทิศทางหรือแนวโน้ม (Trending) เราใช้กราฟที่แสดงผลแบบทิศทางหรือแนวโน้ม เพื่อนำเสนอข้อมูลให้เห็นจำนวนข้อมูลที่เกิดขึ้นในแต่ละช่วงเวลา

(period) รวมถึงเน้นข้อมูลที่ต้องการนำเสนอ เช่น Line Chart, Bar Chart, Radar Chart, Area Chart เป็นต้น

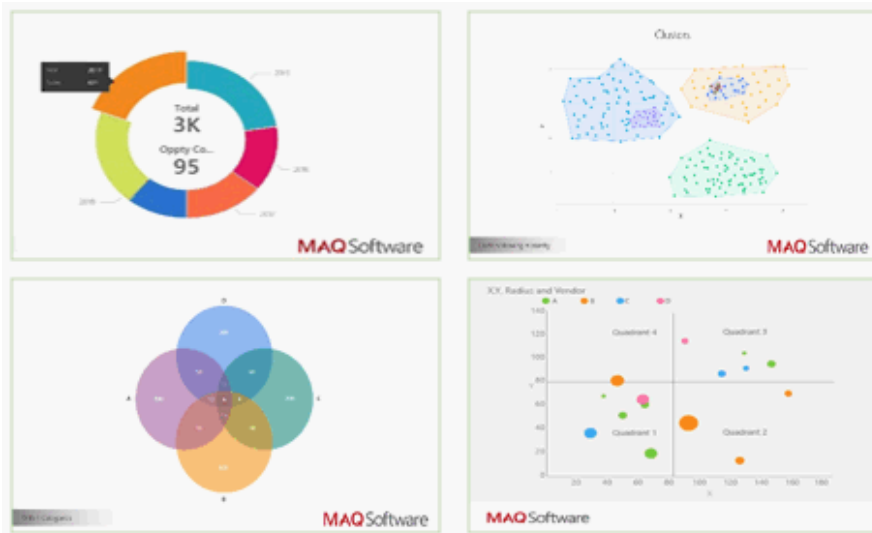


ภาพที่ 2.4 การนำเสนอแบบทิศทางหรือแนวโน้ม (Trending)

(ที่มา : www.autosoft.in.th/data-Visualization/การนำเสนอแผนภาพข้อมูล-data-visu/)

2.2.4.3 การนำเสนอแบบกลุ่มข้อมูล (Classification)

การนำเสนอแบบกลุ่มข้อมูล (Classification) เป็นการนำเสนอโดยนำข้อมูลมาจัดเป็นกลุ่ม ๆ เช่น Donut Chart, Ring Chart, Pie Chart เป็นต้น

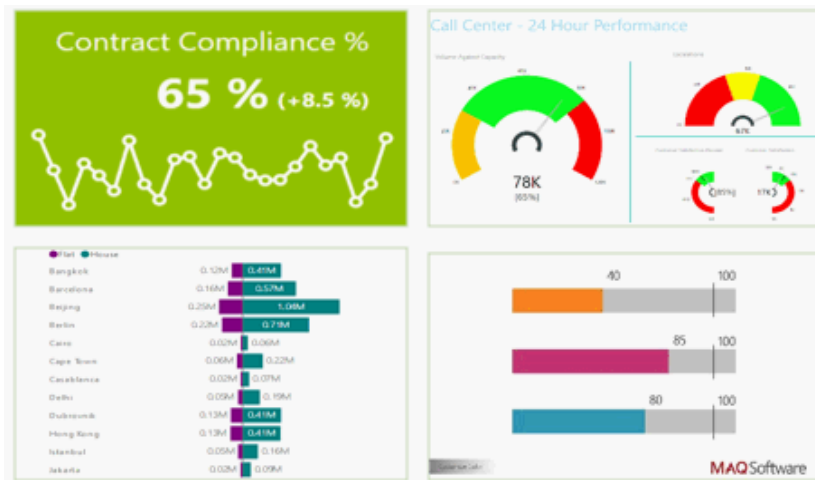


ภาพที่ 2.5 การนำเสนอแบบกลุ่มข้อมูล (Classification)

(ที่มา : www.autosoft.in.th/data-Visualization/การนำเสนอแผนภาพข้อมูล-data-visu/)

2.2.4.4 การนำเสนอเชิงเปรียบเทียบข้อมูล (Comparison)

การนำเสนอเชิงเปรียบเทียบข้อมูล (Comparison) เหมาะสำหรับการนำเสนอที่ต้องการเปรียบเทียบผลการดำเนินงาน เช่น เทียบกับปีที่แล้ว (YoY) เปรียบเทียบกับเป้าที่ตั้งไว้ (Target) ซึ่งกราฟที่เหมาะสมและมักนำมาใช้ เช่น KPI Indicator, Bullet Chart, Power BI Card with state เป็นต้น



ภาพที่ 2.6 การนำเสนอเชิงเปรียบเทียบข้อมูล (Comparison)

(ที่มา : www.autosoft.in.th/data-Visualization/การนำเสนอแผนภาพข้อมูล-data-visu/)

2.2.4.5 การนำเสนอรูปแบบแผนที่ (Geographical)

การนำเสนอรูปแบบแผนที่ (Geographical) เหมาะสำหรับนำเสนอข้อมูลบนแผนที่ โดยสามารถที่จะนำยอดขาย, รายได้, ความหนาแน่นของประชากร เพื่อ Focus กลุ่มลูกค้าในแต่ละพื้นที่ที่เราสนใจ เช่น Globe Map, Google Map, Flow Map เป็นต้น

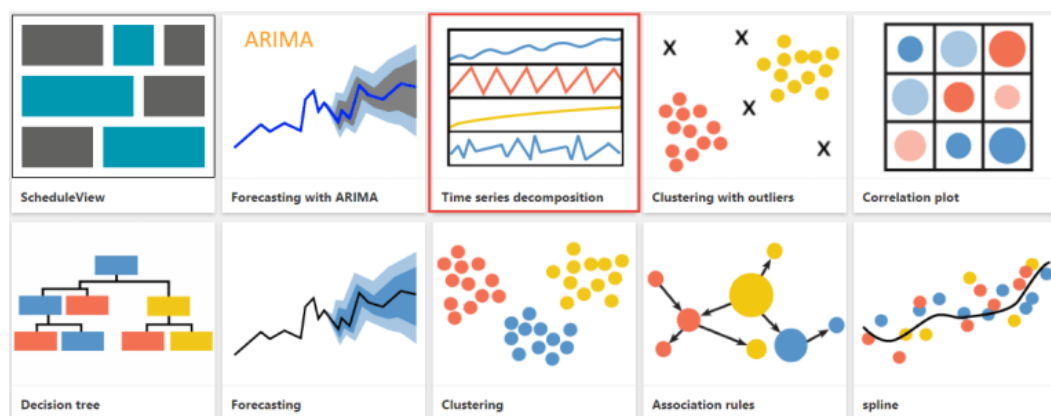


ภาพที่ 2.7 การนำเสนอรูปแบบแผนที่ (Geographical)

(ที่มา : www.autosoft.in.th/data-Visualization/การนำเสนอแผนภาพข้อมูล-data-visu/)

2.2.4.6 กลุ่มที่ต้องพยากรณ์ล่วงหน้าและการวิเคราะห์ข้อมูลที่ซับซ้อน (Analytics)

กลุ่มที่ต้องพยากรณ์ล่วงหน้าและการวิเคราะห์ข้อมูลที่ซับซ้อน (Analytics) เราสามารถที่จะใช้ภาษา R หรือ Python ดึงข้อมูลในอดีตมาเพื่อวิเคราะห์เพื่อพยากรณ์อนาคต และนำเสนอข้อมูลที่ได้ในรูปของกราฟ เช่น Association Rules, Clustering, Forecasting Time series, Calculation plot เป็นต้น



ภาพที่ 2.8 กลุ่มที่ต้องพยากรณ์ล่วงหน้าและการวิเคราะห์ข้อมูลที่ซับซ้อน (Analytics)

(ที่มา : www.autosoft.in.th/data-Visualization/การนำเสนอแผนภาพข้อมูล-data-visu/)

ซึ่งข้อดีของ Data Visualization คือ ข้อมูลที่มีปริมาณมากมาจกมหาศาลยากต่อการเข้าใจ สามารถทำให้เข้าใจง่ายได้ด้วยรูปภาพ ช่วยจัดระเบียบความคิดวิเคราะห์ข้อมูลให้มีความน่าเชื่อถือและประหยัดเวลาในการนำเสนอ

2.2.5 ทฤษฎีเกี่ยวกับต้นไม้ตัดสินใจ

2.2.5.1 ลักษณะต้นไม้การตัดสินใจ

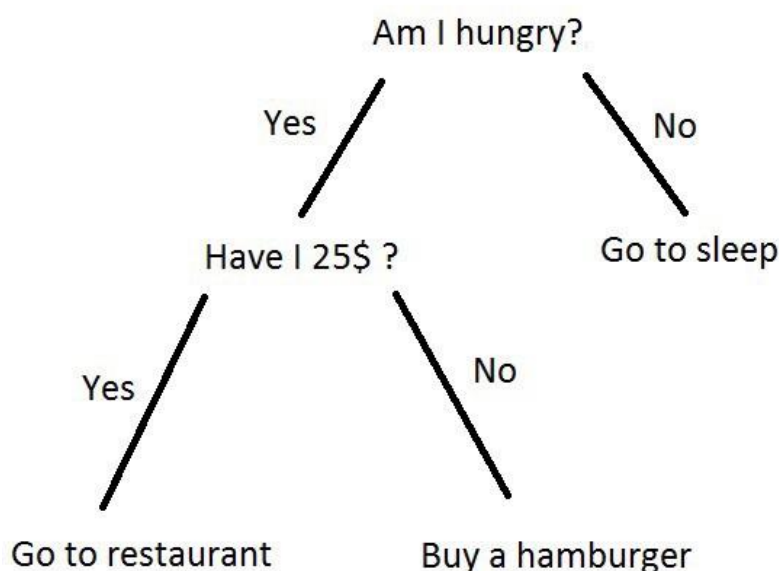
ต้นไม้การตัดสินใจจะทำการจัดกลุ่ม (classify) ชุดข้อมูลนำเข้าในแต่ละกรณี (Instance) แต่ละบัพ (node) ของต้นไม้การตัดสินใจคือตัวแปร (attribute) ต่าง ๆ ของชุดข้อมูล เช่นหากต้องการตัดสินใจว่าจะไปเล่นกีฬาหรือไม่ก็จะมีตัวแปรต้นที่จะต้องพิจารณาคือทัศนียภาพ ลม ความชื้น อุณหภูมิ เป็นต้น และมีตัวแปรตามซึ่งเป็นผลลัพธ์จากต้นไม้คือการตัดสินใจว่าจะไปเล่นกีฬารึเปล่า ซึ่งแต่ละตัวแปรนั้นก็จะมีค่าของตัวเอง (value) เกิดเป็นชุดของตัวแปร-ค่าของตัวแปร (attribute-value pair) เช่น ทัศนียภาพเป็นตัวแปร ก็อาจมีค่าได้เป็น ฝนตก แดดออก หรือการตัดสินใจว่าจะไปเล่นกีฬารึเปล่านั้นก็อาจมีค่าได้เป็นใช่ กับ ไม่ใช่ เป็นต้น การทำนายประเภทด้วยต้นไม้ตัดสินใจ จะเริ่มจากบัพราก โดยทดสอบค่าตัวแปรของบัพ แล้วจึงตามกิ่งของต้นไม้ที่กำหนดค่า เพื่อไปยังบัพลูกถัดไป การทดสอบนี้จะกระทำไปจนกระทั่งเจอบัพใบซึ่งจะแสดงผลการทำนาย

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 เป็นแผนผังต้นไม้ช่วยในการตัดสินใจ โดยแสดงถึงมูลค่าของทรัพยากรที่จะใช้ ความเสี่ยงในการลงทุนและผลลัพธ์ที่มีโอกาสเกิดขึ้น ต้นไม้ตัดสินใจสร้างขึ้นเพื่อช่วยการตัดสินใจเพื่อใช้ในการสร้างแผนงาน นิยมใช้มากในการบริหารความเสี่ยง (Risk Management) ต้นไม้ตัดสินใจเป็นส่วนหนึ่งของทฤษฎีการตัดสินใจ (Decision Theory) และ ทฤษฎีกราฟ ต้นไม้ตัดสินใจเป็นวิธีการพื้นฐานอย่างหนึ่งสำหรับการทำเหมืองข้อมูล



ภาพที่ 2.9 ลักษณะต้นไม้การตัดสินใจ

(ที่มา : <https://th.wikipedia.org/wiki/ต้นไม้ตัดสินใจ>)

2.2.6 แนวคิดปัจจัยเสี่ยงและปัจจัยปกป้องโรคจิตเวช

การป้องกันโรคจิตเวชมุ่งเป้าที่ปัจจัยกำหนดทั้งหลายที่เป็นสาเหตุ หรือชักนำให้เกิดโรคจิตเวช ปัจจัยเสี่ยง (Risk Factors) จะเพิ่มโอกาสเกิดปัญหาสุขภาพที่สำคัญหรือทำให้ปัญหารุนแรงขึ้นหรือยาวนานขึ้น ปัจจัยปกป้อง (Protective Factors) หมายถึงการเพิ่มความต้านทานของบุคคลต่อปัจจัยเสี่ยงและความเจ็บป่วย ซึ่งปัจจัยดังกล่าวสามารถแก้ไข ลด หรือปรับเปลี่ยนการตอบสนองของคนต่อบางสถานการณ์เสี่ยงที่ชักนำไปสู่ผลที่ไม่เหมาะสมได้ (Rutter, 1985) ส่วนใหญ่แล้วปัจจัยปกป้องระดับบุคคลจะเหมือนกับลักษณะของสุขภาพจิตที่ดี เช่น ความภาคภูมิใจในตนเอง ความสามารถในการปรับอารมณ์ ความคิดเชิงบวก ทักษะการแก้ปัญหาและทักษะทางสังคม ทักษะการจัดการกับความเครียด และความรู้ลึกควมคุมตัวเองได้ ด้วยเหตุนี้ มาตรการป้องกันที่มีวัตถุประสงค์เพื่อเสริมสร้างปัจจัยปกป้องส่วนใหญ่จึงทับซ้อนกันกับการส่งเสริมสุขภาพจิต

หลักฐานปัจจัยเสี่ยงและปัจจัยปกป้องกันและความเชื่อมโยงที่นำไปสู่การเกิดโรคจิตเวชนั้นชัดเจนมาก (เช่น Coie และคณะ, 1993; Ingram & Price, 2000) ปัจจัยเสี่ยงและปัจจัยปกป้องมีทั้งระดับบุคคล ครอบครัว สังคม เศรษฐกิจ และสิ่งแวดล้อมส่วนใหญ่แล้วผลจากการมีหลายปัจจัยเสี่ยงร่วมกัน หรือขาดปัจจัยปกป้อง หรือมีทั้งปัจจัยเสี่ยงและปัจจัยปกป้องผสมผสานกันและชักนำให้บุคคลนั้นเปลี่ยนจากสุขภาพจิตดีเป็นการเพิ่มโอกาสป่วย จนเกิดปัญหาสุขภาพจิตและเกิดโรคจิตเวชขึ้นในที่สุด

มาตรการเพื่อป้องกันความเจ็บป่วยทางจิตใจมีจุดมุ่งหมายที่จะรับมือกับปัจจัยเสี่ยงและเสริมสร้างปัจจัยปกป้องตลอดทุกช่วงวัย เพื่อขัดขวางกระบวนการที่จะนำไปสู่ความผิดปกติทางจิตใจ ปัจจัยบุคคลใดมีอิทธิพลต่อพัฒนาการสุขภาพจิตและการเกิดโรคจิตเวชมาก หากจัดการปัจจัยดังกล่าวได้สำเร็จจะยิ่งได้ผลในการป้องกันสูงมาก

ปัจจัยกำหนดในมาตรการป้องกันจะต้องเป็นปัจจัยที่สามารถปรับเปลี่ยนได้และมีทั้งปัจจัยเฉพาะโรค ปัจจัยเสี่ยงและปัจจัยปกป้องทั่วไป โดยปัจจัยเสี่ยงและปัจจัยปกป้องทั่วไป (Generic risk and protective factors) นั้นเป็นปัจจัยที่พบร่วมได้บ่อยในปัญหาสุขภาพจิตและโรคจิตเวชต่าง ๆ มาตรการจัดการกับปัจจัยทั่วไปได้ดีนั้นจะส่งผลกระทบต่อการป้องกันได้ในวงกว้าง ตัวอย่างเช่น ความยากจนและการล่วงละเมิดเด็กซึ่งพบได้บ่อยในโรคซึมเศร้า โรควิตกกังวลและความผิดปกติการใช้สารเสพติด ดังนั้นมาตรการที่จะจัดการกับความยากจนและการล่วงละเมิดเด็กได้ดี คาดได้ว่าจะส่งผลกระทบต่อทั้งสามโรคนี้ ส่วนปัจจัยเสี่ยงและปัจจัยปกป้องเฉพาะโรค (Disease-specific risk and preventive factors) ส่วนใหญ่จะเกี่ยวข้องกับการ

เกิดโรคนั้น ๆ ตัวอย่างเช่น ความคิดเชิงลบจะมีความเฉพาะกับโรคซึมเศร้า และโรคซึมเศร้ามีความเฉพาะกับการฆ่าตัวตาย เป็นต้น

นอกจากนี้ยังมีความสัมพันธ์ระหว่างสุขภาพกายและสุขภาพจิต ยกตัวอย่างเช่น โรคหัวใจและหลอดเลือดสามารถนำไปสู่โรคซึมเศร้าหรือในทางกลับกันโรคซึมเศร้าสามารถนำไปสู่โรคหัวใจและหลอดเลือดได้ สุขภาพจิตและสุขภาพกายมีความเกี่ยวข้องผ่านปัจจัยเสี่ยงที่พบบ่อยร่วมกัน เช่น ที่อยู่อาศัยยากจนนำไปสู่ทั้งสุขภาพจิตไม่ดีและสุขภาพกายไม่ดี จึงจำเป็นต้องทำความเข้าใจระหว่างโรคจิตเวชต่าง ๆ ระหว่างสุขภาพจิตและสุขภาพกาย และการที่ปัจจัยเสี่ยงทั่วไปและปัจจัยเสี่ยงเฉพาะโรคจะเป็นหนทางนำไปสู่ความเจ็บป่วยทางจิตเวช

องค์ความรู้ที่มีหลักฐานเกี่ยวกับปัจจัยเสี่ยงและปัจจัยปกป้องมีอยู่เพียงพอแล้วที่จะรับประกันการลงทุนโดยภาครัฐและเอกชนในการที่จะพัฒนา เผยแพร่ และดำเนินการตามโปรแกรมและนโยบายที่อิงหลักฐานเชิงประจักษ์ มาตรการที่มีผลต่อปัจจัยเสี่ยงและปัจจัยปกป้องที่ส่งผลกระทบต่อขนาดใหญ่หรือมาตรการที่มีผลต่อปัญหาที่พบบ่อยร่วมกัน รวมทั้งปัญหาสังคมและเศรษฐกิจจะมีความคุ้มค่าประสิทธิผลที่สุดและน่าสนใจที่สุดสำหรับผู้กำหนดนโยบายและผู้มีส่วนได้ส่วนเสียอื่น ๆ

ผู้กำหนดนโยบายและผู้ออกแบบโปรแกรมต้องคำนึงว่าโรคจิตเวชเฉพาะโรคอาจทำให้เกิดผลลัพธ์ที่แตกต่างกันในกลุ่มประชากรที่มีความเสี่ยงหลากหลาย ดังนั้นนโยบายสาธารณสุขที่มีประสิทธิผลควรครอบคลุมมาตรการป้องกันหลายมาตรการที่จัดการกันสาเหตุที่หลากหลายในประชากรกลุ่มเสี่ยงที่เกี่ยวข้อง

2.2.6.1 ปัจจัยกำหนดทางสังคม สิ่งแวดล้อม และเศรษฐกิจ (Social, environmental and economic determinants)

ปัจจัยทางเศรษฐกิจสังคมและสิ่งแวดล้อมที่สำคัญสำหรับสุขภาพจิตมักเป็นประเด็นใหญ่เช่น ความยากจน สงคราม และความไม่เท่าเทียมกัน ตัวอย่างเช่น คนยากจนมักจะมีชีวิตอยู่โดยขาดเสรีภาพพื้นฐานทั้งด้านความปลอดภัย การกระทำ หรือทางเลือกที่จะทำให้ชีวิตดีขึ้น โดยมักจะไม่มีความรู้ที่พอก หรือไม่ได้รับการศึกษาหรือการดูแลสุขภาพอย่างเหมาะสมความขาดแคลนนี้จำกัดการใช้ชีวิตที่แตกต่างจากชีวิตที่มีคุณค่าเช่นคนทั่วไป ประชากรที่อาศัยอยู่ในสภาพเศรษฐกิจและสังคมที่ยากจนเพิ่มความเสี่ยงของปัญหาสุขภาพจิต โรคซึมเศร้า และความอยู่ดีมีสุข ปัจจัยมหภาคอื่น ๆ เช่น การขยายสังคมเมือง สงคราม และการย้ายถิ่น การเหยียดเชื้อชาติ และความไม่มั่นคงทางเศรษฐกิจจะเชื่อมโยงกับการเพิ่มขึ้นของอาการและความเจ็บป่วยทางจิตเวช ยกตัวอย่าง เช่น สงครามและบาดแผลทางใจจากสงครามเป็นสาเหตุของโรคเครียดหลังเหตุการณ์สะเทือนใจ โรคซึมเศร้า โรควิตกกังวล และความ

ผิดปกติพฤติกรรมดื่มสุรา นอกจากนี้บาดแผลทางใจดังกล่าวยังเพิ่มความเสี่ยงของปัญหาสุขภาพจิตในลูกหลานที่พ่อแม่เผชิญเหตุการณ์สะเทือนใจหรือโรคซึมเศร้าอีกด้วย

ตารางที่ 2.1 ปัจจัยสังคม สิ่งแวดล้อม และเศรษฐกิจของสุขภาพจิต

ปัจจัยเสี่ยง	ปัจจัยปกป้อง
1) การเข้าถึงสุราและยาเสพติด	1) การเสริมความสามารถ
2) การย้ายถิ่น	2) การรวมกลุ่มชนกลุ่มน้อย
3) การแยกตัวเองและความห่างเหิน	3) สัมพันธภาพเชิงบวก
4) ไม่ได้รับการศึกษา การคมนาคมลำบาก ไร้ที่อยู่อาศัย	4) การมีส่วนร่วมทางสังคม
5) ชุมชนเพื่อนบ้านวุ่นวาย	5) ความรับผิดชอบต่อสังคมและความ อดทน
6) การถูกปฏิเสธจากกลุ่มเพื่อน	6) การบริการสังคม
7) สภาพสังคมที่ขาดแคลน	7) การสนับสนุนทางสังคมและเครือข่ายใน ชุมชน
8) โภชนาการไม่ดี	
9) ความยากจน	
10) ความอยุติธรรมและการเลือกปฏิบัติ ทางเชื้อชาติ	
11) การเสียเปรียบทางสังคม	
12) สังคมเมือง	
13) ความรุนแรงและการกระทำผิด กฎหมาย	
14) สงคราม	
15) ความเครียดจากการทำงาน	
16) การว่างงาน	

2.2.6.2 ปัจจัยกำหนดด้านบุคคลและครอบครัว (Individual and family-related determinants)

ปัจจัยเสี่ยงและปัจจัยปกป้องของบุคคลและครอบครัวเป็นได้ทั้งชีวภาพ อารมณ์ ความคิด พฤติกรรม สัมพันธภาพระหว่างบุคคล หรือเกี่ยวกับบริบทครอบครัว โดยอาจเกิดผลกระทบต่อสุขภาพจิตที่สำคัญเฉพาะ ณ บางช่วงอายุของชีวิต และสามารถส่งผลข้ามจากรุ่นสู่รุ่นได้ ยกตัวอย่างเช่น การล่วงละเมิดเด็กและความเจ็บป่วยทางจิตเวชของพ่อแม่

ในช่วงวัยทารกหรือวัยเด็กตอนต้นสามารถนำไปสู่โรคซึมเศร้าหรือโรควิตกกังวลของบุคคลนั้น เมื่อโตขึ้นและยังส่งผ่านไปสู่วัยรุ่นได้ ในขณะที่ความผูกพันที่มั่นคงและสังคมครอบครัวที่เกื้อหนุนสามารถลดความเสี่ยงดังกล่าวได้ ส่วนพฤติกรรมเสี่ยงของมารดาาระหว่างตั้งครรภ์หรือการเลี้ยงดูไม่ดีตั้งแต่ช่วงต้นของชีวิตเป็นเหตุของความเครียดทางจิตเวชและระบบประสาทได้ ความขัดแย้งในชีวิตสมรสนำไปสู่ความผิดปกติพฤติกรรมเกร (conduct disorders) ในเด็ก โรคซึมเศร้าในผู้หญิงและความผิดปกติพฤติกรรมติ่มสุราในทั้งพ่อและแม่ได้

ผู้สูงอายุที่เจ็บป่วยทางกายอาจมีปัจจัยเสี่ยงหรือปัญหาที่ตามมา เช่น นอนไม่หลับเรื้อรัง ปัญหาการติ่มสุรา การทารุณกรรมผู้สูงอายุ ความโคกเศร้าจากการสูญเสียบุคคลอันเป็นที่รัก หรือปัจจัยเสี่ยงอื่น ๆ ที่เกี่ยวกับความเจ็บป่วยหรือปัญหาพฤติกรรมในอดีต เช่น เคยมีโรคซึมเศร้ามาก่อน เป็นต้น โรควิตกกังวลเองเพิ่มความเสี่ยงของโรคซึมเศร้า โรคซึมเศร้าเพิ่มความเสี่ยงของโรคหัวใจและหลอดเลือดตามมา บางครั้งอาจเกิดจากหลายเหตุต่อเนื่องกัน เช่น โรคสมาธิสั้นในวัยเด็ก ความผิดปกติพฤติกรรมเกรในวัยรุ่น และความผิดปกติพฤติกรรมติ่มสุราและโรคซึมเศร้าในวัยผู้ใหญ่ เหล่านี้เป็นเพียงตัวอย่างบางส่วนของปัจจัยเสี่ยงและปัจจัยปกป้องที่มีบทบาทสำคัญต่อบุคคลและครอบครัว คำอธิบายที่ 5 แสดงให้เห็นถึงปัจจัยหลักอิงตามหลักฐานเชิงประจักษ์หลักที่พบว่าเกี่ยวข้องกับการเกิดโรคจิตเวช

ผู้กำหนดนโยบายและผู้ปฏิบัติงานควรได้รับรู้ถึงองค์ความรู้ปัจจัยกำหนดต่อสุขภาพจิตที่มีหลักฐานเชิงประจักษ์และแก้ไขได้ รวมถึงความเชื่อมโยงกับความเจ็บป่วยทางจิต หลักฐานของปัจจัยระดับบุคคล ครอบครัว สังคม เศรษฐกิจ และสิ่งแวดล้อมต่อสุขภาพจิตนั้นชัดเจน และพบว่าไม่เพียงนำไปสู่ปัญหาสุขภาพจิตและโรคจิตเวชเท่านั้นยังสัมพันธ์กับปัญหาสุขภาพกายอีกด้วย ตัวอย่างเช่น กะโหลกศีรษะร้าว บาดเจ็บที่ศีรษะ โรคหัวใจและหลอดเลือด โรคมะเร็ง และโรคตับแข็ง ในบทยัดไปจะกล่าวถึงหลักฐานที่แสดงว่าปัญหาสุขภาพจิตและโรคจิตเวชเหล่านั้นสามารถแก้ไขด้วยการจัดการปัจจัยเสี่ยงและปัจจัยปกป้องผ่านมาตรการการป้องกันและส่งเสริมสุขภาพจิตได้อย่างไร

ตารางที่ 2.2 ปัจจัยเสี่ยงและปัจจัยปกป้องสำหรับโรคจิตเวช

ปัจจัยเสี่ยง	ปัจจัยปกป้อง
1) ไม่ประสบความสำเร็จในการเรียนและขาดการเรียนรู้ทางศิลปกรรม	1) ความสามารถในการรับมือกับความเครียด
2) สมาธิสั้น	2) ความสามารถในการเผชิญกับความทุกข์ยาก
3) การดูแลผู้ป่วยโรคเรื้อรังหรือสมองเสื่อม	3) การปรับตัว
4) การล่วงละเมิดเด็กหรือทอดทิ้งเด็ก	

5) นอนไม่หลับเรื้อรัง	4) ความเป็นตัวของตัวเอง
6) ปวดเรื้อรัง	5) การกระตุ้นความคิดอ่านตั้งแต่อายุน้อย
7) การสื่อสารไม่ตรงไปตรงมา	6) การออกกำลังกาย
8) การตั้งครมร์ภในวัยรุน	7) ความรู้สึกลบลดภย
9) การทารุณกรรมผู้สูงอายุ	8) ความรู้สึกรว่าจัดการและควบคุมได้
10) วุฒิภาวะอารมณ์ไม่สมวัยและไม่สามารถควบคุมได้	9) การอบรมเลี้ยงดูที่ดี
11) ใช้สารเสพติดมากเกินไป	10) การรู้หนังสือ
12) เผลอญกับความก้าวร้าว รุนแรง และบาดแผลทางใจ	11) ความผูกพันที่ดีและความผูกพันในช่วงต้นของชีวิต
13) ความขัดแย้งในครอบครัวหรือครอบครัวแตกแยก	12) ปฏิสัมพันธ์พ่อแม่และลูกในเชิงบวก
14) ความเหงา	13) ทักษะการแก้ปัญหา
15) น้ำหนักแรกคลอดต่ำกว่าเกณฑ์	14) พฤติกรรมเตรียมเข้าสังคม
16) ระดับชั้นทางสังคมต่ำ	15) ความภาคภูมิใจในตนเอง
17) ความเจ็บป่วยทางกาย	16) ทักษะชีวิต
18) ความไม่สมดุลของสารสื่อประสาทในสมอง	17) ทักษะการจัดการทางสังคมและจัดการกับความขัดแย้ง
19) พ่อแม่ป่วยด้วยโรคจิตเวช	18) การพัฒนาทางสังคมอารมณ์
20) พ่อแม่ใช้สารเสพติด	19) การจัดการกับความเครียด
21) ภาวะแทรกซ้อนระหว่างคลอด	20) การเกื้อหนุนทางสังคมโดยครอบครัวและเพื่อน
22) ความโศกเศร้าจากการสูญเสียบุคคลอันเป็นที่รัก	
23) ทักษะและความสามารถการทำงานไม่ดี	
24) ความบกพร่องด้านการอ่าน	
25) ความพิการทางประสาทสัมผัสหรือพิการทางร่างกาย	
26) ทักษะทางสังคมไม่ดี	
27) เหตุการณ์ที่เครียดในชีวิต	
28) การใช้สารเสพติดในระหว่างตั้งครมร์ภ	

2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.3.1 แบบจำลองต้นไม้ตัดสินใจ (Decision Tree)

2.3.1.1 เทคนิคต้นไม้ตัดสินใจ (Decision Tree)

เป็นหนึ่งในวิธีการเรียนรู้ซึ่งใช้ในสถิติ, การเรียนรู้ของเครื่อง และการทำเหมืองข้อมูล โดยพิจารณาการสังเกตการแบ่งแยกข้อมูลโดยพิจารณาข้อมูลในการเรียนรู้ของเครื่อง (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 เป็นแผนผังต้นไม้ช่วยในการตัดสินใจ โดยแสดงถึงมูลค่าของทรัพยากรที่จะใช้ ความเสี่ยงในการลงทุนและผลลัพธ์ที่มีโอกาสเกิดขึ้น ต้นไม้ตัดสินใจสร้างขึ้นเพื่อช่วยการตัดสินใจเพื่อใช้ในการสร้างแผนงาน นิยมใช้มากในการบริหารความเสี่ยง (Risk Management) ต้นไม้ตัดสินใจเป็นส่วนหนึ่งของทฤษฎีการตัดสินใจ (Decision Theory) และ ทฤษฎีกราฟ ต้นไม้ตัดสินใจเป็นวิธีการพื้นฐานอย่างหนึ่งสำหรับการทำเหมืองข้อมูล

2.3.1.2 ปัญหาที่เหมาะสมสำหรับต้นไม้การตัดสินใจ

เนื่องจากต้นไม้การตัดสินใจเป็นต้นไม้ที่แต่ละกิ่งที่ออกมาจากบัพแทนค่าของข้อมูลที่เป็นไปได้ในบัพนั้น เนื่องจากต้นไม้มีจำนวนกิ่งที่จำกัด ดังนั้นค่าของตัวแปรที่เป็นไปได้จึงต้องจำกัดด้วย จึงต้องมีจำนวนตัวแปรที่จำกัด และนอกจากนั้นยังบังคับว่าค่าของตัวแปรนั้นต้องไม่ต่อเนื่องด้วย โดยข้อมูลที่เข้ามานั้นอาจมีความผิดพลาดได้บ้าง โดยต้นไม้การตัดสินใจจะมีกระบวนการที่จะไม่นำความผิดพลาดนั้นมาพิจารณาเรียกว่าการตัดแต่งกิ่ง (Post-Pruning)

2.3.1.3 ขั้นตอนวิธีการสร้างต้นไม้การตัดสินใจ เพื่อจำแนกข้อมูล มีดังนี้

- 1) เลือก Attribute ที่ทำหน้าที่เป็น Root Node
- 2) จาก Root Node สร้างเส้นเชื่อมโยงไปยังโหนดลูก จำนวนเส้นเชื่อมโยงจะเท่ากับจำนวนค่าที่เป็นไปได้ทั้งหมดของ Attribute ที่เลือกเป็น Root Node
- 3) ถ้าโหนดลูกเป็นกลุ่มของข้อมูลที่อยู่ในคลาสเดียวกันทั้งหมด ให้หยุด

สร้างต้นไม้ แต่ถ้าโหนดมีข้อมูลของหลายคลาสปะปนกันอยู่ ต้องสร้าง Subtree เพื่อจำแนกข้อมูลต่อไป โดยเลือก subtree มาทำหน้าที่เป็น Root Node ของ Subtree มาทำซ้ำในขั้นตอนที่ 2.) , 3.) ซึ่งการคำนวณมีดังนี้

เอนโทรปี (Entropy) การสร้างต้นไม้ตัดสินใจจากบนลงล่างด้วยการถามว่าลักษณะใด (ขอใช้คำว่าลักษณะแทนตัวแปรต้น) ควรจะเป็นรากของต้นไม้การตัดสินใจต้นนี้ และถามซ้ำ ๆ ไปเรื่อย ๆ เพื่อหาต้นไม้ทั้งต้นด้วยการเขียนโปรแกรมด้วยความสัมพันธ์แบบเวียนเกิด (อังกฤษ: recursion) โดยในการเลือกว่าลักษณะใดดีที่สุดนั้นดูจากค่าของลักษณะเรียกว่าเกนความรู้ (Information gain) ก่อนที่จะรู้จักเกนความรู้จะต้องนิยามค่าหนึ่งที่ใช้บอกความไม่บริสุทธิ์ของข้อมูลก่อน เรียกว่าเอนโทรปี (Entropy) โดยนิยามเอนโทรปีของต้นไม้การตัดสินใจในตัวในเซตของตัวอย่าง S คือ $E(S)$ ดังนี้

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

เมื่อพิจารณาเอนโทรปีแล้วจะเห็นว่าเอนโทรปีจะมีค่าอยู่ระหว่าง 0 กับ 1 โดยจะมีค่าเป็นศูนย์เมื่อทุก ๆ กรณีมีผลลัพธ์เพียงแบบเดียว เช่น ใช้ทั้งหมด หรือ ไม่ใช่ทั้งหมด และจะมีค่ามากขึ้นเมื่อเริ่มมีค่าที่แตกต่างกันมากขึ้น หรือจะพูดอีกนัยหนึ่งก็คือเอนโทรปีจะมีค่ามากขึ้นหากข้อมูลไม่บริสุทธิ์ และจะตัดสินใจได้ว่าผลลัพธ์จะเป็นอะไรเมื่อเอนโทรปีเป็น 0 เท่านั้น

เกนความรู้ (Information Gain) ซึ่งจากการนิยามเอนโทรปีข้างต้น ทำให้เราสามารถนิยามลักษณะของตัวแปรต้นที่ดีได้ โดยตัวแปร A จะเป็นตัวแปรต้นที่ดีก็ต่อเมื่อหากว่าแบ่งข้อมูลตัวอย่าง (Example) ออกเป็นชุด ๆ มีจำนวนชุดตามจำนวนค่าของ A ที่เป็นไปได้ เพื่อให้แต่ละกรณี (Instance) ในชุดนั้นมีค่า A เพียงค่าเดียวและค่าเฉลี่ยของเอนโทรปีของชุดข้อมูลที่ถูกแบ่งออก (partition) มานั้นต่ำที่สุด เรียกค่าคาดหวังของการลดลงของเอนโทรปีหลังจากข้อมูลถูกแบ่งด้วย A ว่าเกนความรู้ของ A นิยามโดย

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

เมื่อแทนค่า S คือ ตัวอย่างที่ประกอบด้วยชุดของตัวแปรต้นและตัวแปรตามหลายๆกรณี

E คือ เอนโทรปีของตัวอย่าง

A คือ ตัวแปรต้นที่พิจารณา

value (A) คือ เซตของค่าของ A ที่เป็นไปได้

Sv คือ ตัวอย่างที่ A มีค่า v ทั้งหมด

จะเห็นว่าหากเอนโทรปีของ A ยิ่งมากแสดงว่าหลังจากแบ่งตัวอย่าง S ด้วย A แล้วในแต่ละชุดที่แบ่งได้จะมี Entropy เข้าใกล้ศูนย์มากยิ่งขึ้น ทำให้ใกล้ที่จะตัดสินใจได้มากขึ้น เอนโทรปีจึงเป็นค่าที่ดีที่จะบอกความดีของตัวแปรต้นที่นำมาพิจารณา

2.3.2 การประเมินผลโมเดล (Decision Tree)

สถิติในการทดสอบหาความแม่นยำของอัลกอริทึม และทดสอบความถูกต้อง และความน่าเชื่อถือของโมเดลโดยใช้การประเมินค่าความแม่นยำ Confusion Matrix

Confusion Matrix คือ การประเมินผลลัพธ์การทำนาย (หรือผลลัพธ์จากโปรแกรม) เปรียบเทียบกับผลลัพธ์จริงโดยที่

True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าจริง และคนบอกว่ามันจริง

True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าไม่จริง และคนบอกว่ามันไม่จริง

False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าจริง แต่คนบอกว่าไม่จริง

False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าไม่จริง แต่คนบอกว่าจริง

		Actual Value (as confirmed by experiment)	
		positives	negatives
Predicted Value (predicted by the test)	positives	TP True Positive	FP False Positive
	negatives	FN False Negative	TN True Negative

ภาพที่ 2.10 การประเมินค่าความแม่นยำ Confusion Matrix

(ที่มา : <https://medium.com/@natratanonkanraweekultana/confusion-matrix-d6146b275faa>)

แสดงเป็นสมการได้ ดังนี้

Accuracy คือ ค่าที่บอกว่าโปรแกรมสามารถทำนายได้แม่นยำขนาดไหน

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

สมการ accuracy = total correct predictions / total predictions made * 100

$$\text{accuracy} = 7 / 10 * 100$$

Expected,	Predicted
man,	woman
man,	man
woman,	woman
man,	man
woman,	man
woman,	woman
woman,	woman
man,	man
man,	woman
woman,	woman

จากข้อมูลที่ติด label ถูกต้อง คือ Expected และข้อมูลที่ทำนายออกมา คือ Prediction

Precision คือ ค่าที่บอกว่าโปรแกรมทำนายว่าจริง ถูกต้องเท่าไร

$$\text{precision} = \frac{TP}{TP + FP}$$

สมการ Precision = TP/(TP+FP) ใช้ในการทำนายว่าถ้าทำนายผิดว่าเป็นแล้วเสียหาย (ถ้าทำนายผิดแล้วเกิดความเสียหาย) ถ้าค่า Precision มีค่าน้อยแสดงว่าโมเดลไม่เหมาะสม เช่นอีเมลไม่ได้เป็น spam แต่ทำนายว่าเป็น spam

Recall (True Positive Rate) คือ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าจริง เป็นอัตราส่วนเท่าไรของจริงทั้งหมด

$$\text{recall} = \frac{TP}{TP + FN}$$

สมการ Recall = TP/(TP+FN) ใช้ในการทำนายที่ผิดแบบ negative แล้วเสียหาย เช่น เป็น HIV แล้วบอกว่าไม่เป็น

F1-score คือ การนำค่า Precision กับ Recall มาคำนวณค่าเฉลี่ยกันสำหรับพิจารณาร่วมกันเพื่อความแม่นยำ และความครบถ้วน

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

สมการ $F1\text{-score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$ มีบางโมเดลที่จำเป็นต้องใช้ทั้ง 2 สูตร จึงต้องใช้หาค่าเฉลี่ยระหว่าง Precision * Recall แต่หน่วยหารมันไม่เท่ากัน จึงต้องใช้สูตร F1-score

2.3.3 การแสดงผลแบบ Visualization

2.3.3.1 การแสดงผลแบบ Visualization ด้วย Power BI

Power BI Desktop คือ โปรแกรมที่ใช้ช่วยวิเคราะห์ สรุปผลข้อมูลจำนวนมากจากหลายแหล่งข้อมูล อาทิ Excel File, Microsoft Access Database, SQL Server, Oracle เป็นต้น ซึ่ง Power BI Desktop สามารถทำงานได้อย่างรวดเร็ว โดยสามารถแสดงผลได้ทั้งรูปแบบ ตารางสรุปผลด้วย Visualization กราฟในรูปแบบต่าง ๆ และยังติดตั้งเพิ่มเติมได้จาก Marketplace อีกทั้งยังสามารถแสดงผลในรูปแบบแผนที่ได้ โดยสามารถแสดงผลผ่านเว็บไซต์ และอุปกรณ์ Mobile และ Tablet ทำให้ได้ข้อมูลที่ถูกต้องและรวดเร็ว มีความพร้อมในการกำหนดกลยุทธ์ และตัดสินใจได้อย่างถูกต้องและแม่นยำในทางธุรกิจ

2.3.3.2 Power BI Desktop สามารถเลือกแหล่งข้อมูลใดได้บ้าง

- นำเข้าจากไฟล์ข้อมูล ได้แก่ Excel, CSV, XML, Text เป็นต้น
- นำเข้าจากฐานข้อมูล ได้แก่ SQL Server, Access Database, SQL Server Analysis Services Database, Oracle, MySQL, DB2, Sybase, PostgreSQL
- Azure ได้แก่ Microsoft Azure SQL Database, Azure SQL Data Warehouse, Azure Marketplace, Azure Blob Storage, Azure Table Storage เป็นต้น

2.3.3.2 ลักษณะเด่นของ Power BI

- มีระบบ Dashboard สำหรับการสร้างรายงานที่ใช้งานได้ง่าย อีกทั้งตัวซอฟต์แวร์ยังมีชุดรายงานสำเร็จรูปที่สนับสนุนให้การทำรายงานเป็นไปอย่างรวดเร็วและง่ายดายยิ่งขึ้น

- Power BI สามารถรองรับข้อมูลที่ผู้ใช้ต้องการนำมาวิเคราะห์ได้ค่อนข้างหลากหลายโดยสามารถแบ่งชนิดข้อมูลที่จะนำมาใช้งาน ได้ดังนี้

- 1) ข้อมูลในลักษณะ Text file อันได้แก่ Text, Excel, CSV และ XML เป็นต้น

- 2) ข้อมูลในลักษณะฐานข้อมูล เช่น PostgreSQL, MySQL, DB2, Sybase, Oracle, SQL Server และ Access Database เป็นต้น

- 3) ข้อมูลที่อยู่คลาวด์ (Cloud System) ซึ่งโดยมากจะเป็นบริการของ Microsoft Azure นั่นเอง

- มุมมองในการแสดงผลมีความสวยงาม โดยเฉพาะการนำเสนอรายงานในรูปแบบของกราฟ เนื่องจาก Power BI มีเครื่องมือ (Tools) ที่ใช้ในการสร้างรายงานให้ออกมาน่าสนใจอยู่มากมาย เช่น Area charts: Basic (Layered) and Stacked, Bar and column charts, Cards: Multi row, Cards: Single number, Combo charts, Doughnut charts, Funnel charts, Gauge charts, KPIs, Line charts และ Pie charts เป็นต้น
- Power BI สามารถแสดงผลข้อมูลออกมาในรูปแบบของแผนที่ได้ โดยจะใช้ระบบบริการแผนที่ของ Bing มาช่วยในการนำเสนอข้อมูลในรูปแบบของรายงานได้อย่างสวยงามอีกด้วย
- เมื่อมีการอัปเดตของมูลต้นทาง เช่น มีการปรับเปลี่ยนแก้ไขฐานข้อมูล การแสดงรายงานก็มีการปรับเปลี่ยนตามไปด้วยเช่นเดียวกัน ทำให้รายงานที่นำเสนอมีความเป็นปัจจุบันมากยิ่งขึ้น

2.4 วรรณกรรมที่เกี่ยวข้อง

ณัฐวดี หงษ์บุญมี และคณะ (2562) ภาวะซึมเศร้าเป็นโรคทางจิตเวชที่มีผลกระทบต่อการใช้ชีวิตประจำวันและสภาวะสังคมในปัจจุบันพบว่ามีจำนวนผู้ป่วยเพิ่มมากขึ้น งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อสร้างแบบจำลองการคัดกรองผู้ป่วยภาวะซึมเศร้าด้วยเทคนิคเหมืองข้อมูล เพื่อให้ได้แบบจำลองที่มีประสิทธิภาพสูงสุด ที่จะนำมาวิเคราะห์หาปัจจัยที่ส่งผลกระทบต่อความเสี่ยงภาวะซึมเศร้าและพัฒนาระบบคัดกรองภาวะซึมเศร้าผ่านสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ ข้อมูลที่นำมาใช้ในการทดลองเป็นข้อมูลจากการเก็บข้อมูลแบบสุ่มกลุ่มตัวอย่างจำนวน 505 ชุดข้อมูล เทคนิคเหมืองข้อมูลที่นำมาประยุกต์ใช้ในการสร้างแบบจำลองมี 3 เทคนิค คือ ต้นไม้ตัดสินใจ โครงข่ายประสาทเทียม และนาอิวเบย์ วัดประสิทธิภาพของแบบจำลองด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกลับ ค่าความถ่วงดุล และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE) ผลการทดลองพบว่า แบบจำลองจากเทคนิคโครงข่ายประสาทเทียมมีค่าประสิทธิภาพที่ดีที่สุด โดยมีค่าความถูกต้อง 97.43% ค่าความแม่นยำ 97.50% ค่าความระลึกลับ 97.40% ค่าความถ่วงดุล 97.40% และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง 0.1091 ผลการวิเคราะห์ปัจจัยที่มีความสำคัญ โดยการลดการนำเข้าที่ละปัจจัยพบว่า ปัจจัยด้านการนอนไม่หลับ การเบื่ออาหาร และความรู้สึกไม่ดีต่อตนเอง เป็นปัจจัยที่ส่งผลกระทบต่อความเสี่ยงภาวะซึมเศร้า จากนั้นนำแบบจำลองที่ได้ไปพัฒนาระบบในรูปแบบแอปพลิเคชันบนสมาร์ตโฟน เครื่องมือในการพัฒนาใช้โปรแกรม Android Studio และภาษา Java ผลการประเมินคุณภาพของระบบจากกลุ่มตัวอย่าง 2 กลุ่ม คือ ผู้เชี่ยวชาญ จำนวน 3 คน และผู้ใช้ทั่วไป จำนวน 30

คน พบว่าผู้เชี่ยวชาญมีค่าเฉลี่ยของความพึงพอใจเท่ากับ 4.25 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.44 ผู้ใช้ทั่วไปมีค่าเฉลี่ยของความพึงพอใจเท่ากับ 4.20 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.58 แสดงให้เห็นว่าระบบที่พัฒนาขึ้นนี้มีประสิทธิภาพอยู่ในเกณฑ์ดีสามารถนำไปประยุกต์ใช้งานได้จริง

ปวีณา ชัยวนารมย์ (2558) งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาหาแบบจำลองที่เหมาะสมสำหรับพยากรณ์ความเครียดด้วยเทคนิคการทำเหมืองข้อมูล กระบวนการวิจัยเริ่มต้นด้วยให้กลุ่มตัวอย่างจำนวน 300 คนทำแบบทดสอบประเมินความเครียดเป็นจำนวน 4 รอบ โดยแต่ละรอบการ

ประเมินมีระยะห่างของการทำแบบทดสอบเป็นระยะเวลา 2 เดือน จากนั้น ผลการทดสอบที่ได้ทั้งหมดถูกป้อนเข้าสู่โปรแกรม WEKA เพื่อทำการสร้างแบบจำลองด้วยอัลกอริทึมทางด้านเหมืองข้อมูลจำนวน 6 อัลกอริทึม คือ Bayesian Network, Naïve Bayesian, Decision Tree:4.5, Decision Table, Partial Rules (PART) และ Multilayer Perceptron (MLP) ในการสร้างแบบจำลองและการทดสอบ แบบจำลองนั้น วิธี 10-fold cross-validation ได้ถูกนำมาใช้ในการแบ่งข้อมูลออกเป็นสองชุด ได้แก่ ชุดข้อมูลสอน และชุดข้อมูลทดสอบ

จากการทดสอบแบบจำลองทั้งหมดพบว่า แบบจำลองที่เหมาะสมที่สุดในการนำมาใช้เพื่อพยากรณ์ความเครียดคือ แบบจำลองที่สร้างจากอัลกอริทึม MLP ที่ใช้กับข้อมูลย้อนหลัง 6 เดือน โดยแบบจำลองนี้ มีค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก และค่าความเหวี่ยง เท่ากับ 81%, 0.81, 0.81 และ 0.81 ตามลำดับ

กิตติศักดิ์ สุมามาลย์ (2558) งานวิจัยนี้ได้นำเสนอการวิเคราะห์ข้อมูลจากแบบสำรวจสภาวะสุขภาพประชาชนโดยใช้ข้อมูลประชาชนในเขตเทศบาลเมืองจังหวัดบุรีรัมย์ ปี พ.ศ. 2555 จำนวน 1,071 เรคคอร์ด ด้วยเทคนิคการทำเหมืองข้อมูล โดยแบ่งการศึกษาเป็น 2 ประเด็น คือ การศึกษาความสัมพันธ์ของการเป็นโรคต่าง ๆ โดยใช้เทคนิค Association Rule และศึกษาการคัดกรองสุขภาพเบื้องต้นของประชาชน โดยแบ่งออกเป็น กลุ่มปกติ กลุ่มเสี่ยง และกลุ่มป่วย ด้วยเทคนิค Classification และนำผลลัพธ์ที่ได้ไปพัฒนาเป็นระบบคัดกรองสุขภาพเบื้องต้นเพื่อช่วยลดภาระงานของผู้เชี่ยวชาญในการวิเคราะห์ข้อมูล

ผลการทดลองพบว่า การศึกษาความสัมพันธ์ของการเป็นโรคต่าง ๆ อัลกอริทึม FP Growth ให้คำตอบที่เหมาะสมกว่าอัลกอริทึม Apriori และในการศึกษาคัดกรองสุขภาพเบื้องต้นของประชาชน พบว่าอัลกอริทึม Partial Rules ให้ความถูกต้องมากกว่าอัลกอริทึมต้นไม้ตัดสินใจ C4.5 โดยมีค่าความถูกต้อง (Precision) เท่ากับ 88.60% ให้ค่าระลึก (Recall) เท่ากับ 89.20% แลค่าความเหวี่ยง (F-measure) เท่ากับ 88.80% เมื่อนำกฎที่ได้ไปให้ผู้เชี่ยวชาญตรวจสอบ

ผู้เชี่ยวชาญได้ตัดข้อมูลบางแอททริบิวต์และกฎบางกฎที่ไม่เหมาะสมออกไป สุดท้ายจะได้ผลลัพธ์เป็นกฎจำนวน 42 กฎ และถูกนำไปใช้สร้างระบบคัดกรองสุขภาพอัตโนมัติสำหรับให้ประชาชนทั่วไปเข้ามาตรวจสอบสุขภาพเบื้องต้นของตนเอง

วิฑิตา คำผล (2558) การพยากรณ์กระแสไฟฟ้าขัดข้องโดยใช้เทคนิคการทำเหมืองข้อมูล กรณีศึกษา การไฟฟ้าส่วนภูมิภาค เขต 3 ภาคกลาง มีวัตถุประสงค์เพื่อศึกษาพฤติกรรมและวิเคราะห์ตัวแปรที่มีผลต่อความสูญเสีย เมื่อเกิดกระแสไฟฟ้าขัดข้องแล้วนำมาวางแผนเพื่อป้องกันการเกิดกระแสไฟฟ้าขัดข้อง โดยดำเนินการเลือกวิเคราะห์การเกิดกระแสไฟฟ้าขัดข้องของการไฟฟ้าส่วนภูมิภาคเขต 3 ภาคกลาง จากการศึกษาได้ทำการพยากรณ์การเกิดกระแสไฟฟ้าขัดข้อง โดยผู้ศึกษาวิจัยสามารถสรุปผลการวิจัยจากข้อมูลสถิติกระแสไฟฟ้าขัดข้องระหว่างปี 2555-2557 ซึ่งมีการวิเคราะห์ข้อมูลร่วมกันโดยใช้เทคนิคการจัดกลุ่ม (Cluster Analysis) ด้วยวิธี Kmean ในงานวิจัยนี้สามารถแบ่งกลุ่มตัวแบบพยากรณ์ได้ 7 กลุ่ม เนื่องมาจากการแบ่งกลุ่มข้อมูล ตัวอย่างมีขนาดเล็ก หรือไม่หลากหลายเกินไป ทำให้ผลจากการวิเคราะห์ดังกล่าวสามารถนำไปใช้วางแผนระบบไฟฟ้าส่วนภูมิภาคได้อย่างแม่นยำขึ้น และการจำแนกประเภท (Classification) ด้วยวิธีต้นไม้การตัดสินใจ (Decision tree) เพื่อให้ตัวแบบพยากรณ์มีประสิทธิภาพสูงสุด อีกทั้งยังรวบรวมข้อมูลตามกรอบการพัฒนาเหมืองข้อมูลแบบ CRISP-DM ซึ่งจะทำให้ได้ผลลัพธ์ข้อมูลกระแสไฟฟ้าขัดข้องมีความน่าเชื่อถือมากขึ้น

ศุภามณ จันทร์สกุล (2560) คณะพยาบาลศาสตร์ มหาวิทยาลัยอัสสัมชัญเอเซีย มีแนวคิดในการนำข้อมูล และฐานข้อมูลที่เกี่ยวข้องในในวิชาชีพการพยาบาลมาใช้ประโยชน์โดยค้นหารูปแบบความสัมพันธ์ของข้อมูลซึ่งเรียกว่า “การวิเคราะห์เหมืองข้อมูล” เพื่อนำเสนอความรู้เบื้องต้นในการทำเหมืองข้อมูลสำหรับพยาบาล การวิเคราะห์เหมืองข้อมูลแบ่งออกเป็นสองกลุ่มใหญ่คือ (1) การวิเคราะห์เหมืองข้อมูลเพื่อการทำนายมีการเรียนรู้แบบมีการสอน และ (2) การวิเคราะห์เหมืองข้อมูลเพื่อการอธิบายมีการเรียนรู้แบบไม่มีการสอน เทคนิคเหมืองข้อมูล 3 เทคนิคที่ได้รับความนิยมได้แก่ การจำแนกประเภท (ต้นไม้ตัดสินใจ และเครือข่ายประสาท) การจัดกลุ่มข้อมูล และการค้นหาหาความสัมพันธ์ของข้อมูล ผลการวิเคราะห์ที่ได้จากเหมืองข้อมูลสร้างองค์ความรู้ และเป็นข้อค้นพบที่เป็นประโยชน์ในการนำไปใช้เพื่อการตัดสินใจ

ผลการวิจัยพบว่า การวิเคราะห์เครือข่ายประสาทแบบ MLP สามารถระบุตัวแปรสำคัญที่ใช้ในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนได้คิดเป็นร้อยละ 50 และตัวแปรที่มีความสำคัญลำดับมากพิจารณาจากค่าความสำคัญที่ใช้ในการพยากรณ์พบว่าตัวแปรสไตส์การเรียนรู้มีค่าความสำคัญ (normalized important) สูงสุดคือร้อยละ 100 ผลการวิเคราะห์

สอดคล้องกับการวิเคราะห์ต้นไม้ตัดสินใจด้วยวิธี CRT ซึ่งพบว่าตัวแปรสามารถจัดประเภทบุคคลด้วยการทำนายผลสัมฤทธิ์ทางการเรียนได้ถูกต้องคิดเป็นร้อยละ 59.4 โดยข้อมูลตารางต้นไม้แสดงให้เห็นว่าตัวแปรสโตร์สการเรียนรู้แบบ VARK เป็นตัวแปรต้นหลักในการทำนายผลสัมฤทธิ์ทางการเรียน และพบกฎในการจัดประเภทมีจำนวน 2 กฎคือ (1) ถ้าผู้เรียนมีสโตร์สการเรียนแบบ VAR ในรูปแบบ trimodal หรือ VARK/multimodal จะมีผลสัมฤทธิ์ทางการเรียนดี คิดเป็นร้อยละ 57 และ (2) ถ้าผู้เรียนมีสโตร์สการเรียนแบบ VARK ในรูปแบบ unimodal หรือ bimodal จะมีผลสัมฤทธิ์ทางการเรียนไม่ดีคิดเป็นร้อยละ 62.4

ฉัฐติมา ช่วงชัย (2559) การวิจัยนี้มีวัตถุประสงค์เพื่อจัดการเรียนรู้ให้เหมาะกะกับรูปแบบการเรียนรู้ของผู้เรียน เป็นปัจจัยที่ทำให้ผู้วิจัยมีความสนใจต่อการหารูปแบบการเรียนรู้โดยยึดหลักทฤษฎีของเดวิด คอส์บ มาศศึกษากับนักศึกษาระดับปริญญาตรี สาขาการจัดการเทคโนโลยีการผลิตและสารสนเทศ โดยวิเคราะห์ด้วยการทำเหมืองข้อมูลจากแบบสอบถาม ใช้โปรแกรมเวก้าเป็นเครื่องมือในการวิเคราะห์ เลือกวิธีวิเคราะห์ผลด้วยรูปแบบ Rule Based Classification ด้วยวิธี Decision Table, Jrip และ PART และรูปแบบ Decision Tree Classification ด้วยวิธี LMT, J48 และ Random Tree พร้อมกับวิเคราะห์ตัวแปรด้านเพศ และชั้นปีด้วยวิธีRandom Tree เพื่อดูผลการตัดแยกจากโมเดลแผนภาพต้นไม้ จากการวิเคราะห์ผลทั้งหมดพบว่ารูปแบบของ Decision Tree ด้วยวิธีการ Random Tree ให้ค่าความถูกต้องสูงสุด (100%) ทำนายรูปแบบการเรียนรู้เป็นแบบเอกนัย (Converger) ส่วนรูปแบบของ Rule Based ด้วยวิธี PART ให้ค่าความถูกต้องสูงสุด (84.12%) ทำนายรูปแบบการเรียนรู้ของนักศึกษาทั้งหมดเป็นแบบเอกนัยเช่นกัน แต่เมื่อแบ่งตามเพศและชั้นปี พบว่าเพศหญิงจะมีรูปแบบการเรียนรู้แบบเอกนัย ส่วนเพศชายมีการเรียนรู้แบบบอเนกนัย (Diverger) และเมื่อพิจารณาจากชั้นปีเลือกดูในชั้นปีที่ 3 ซึ่งเป็นชั้นปีที่ต้องจัดทำปริญญาานิพนธ์ พบว่าเพศชายมีเป็นรูปแบบการเรียนรู้แบบบอเนกนัยเช่นกัน จึงเหมาะกะกับการเลือกทำปริญญาานิพนธ์ทางด้านการผลิตและสารสนเทศ เน้นเชิงรูปธรรมด้วยการคิดหลายด้านสร้างเครื่องอุปกรณ์จำลองการผลิต สร้างซอฟต์แวร์ ส่วนเพศหญิงมีรูปแบบการเรียนรู้แบบเอกนัย จึงเหมาะกะกับการเลือกทำทางด้านการจัดการ เน้นเชิงนามธรรม ปฏิบัติเกี่ยวกับแนวคิด ทฤษฎี ผสมกับการทดลอง รูปแบบจึงเป็นการนำทฤษฎีไปวางแผนหรือทดลองเก็บข้อมูลในสถานประกอบการเพื่อให้เกิดประสิทธิภาพดียิ่งขึ้น

เล็กฤทัย ชันทองชัย และคณะ (2558) การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) วิเคราะห์พฤติกรรมผู้เรียนที่เรียนด้วยระบบอีเลิรน์นิงในรายวิชาการเตรียมฝึกประสบการณ์วิชาชีพ โดยใช้เทคนิคเหมืองข้อมูล 2) ค้นหากฎความสัมพันธ์ระหว่างข้อมูลในระบบฐานข้อมูลอีเลิรน์นิง โดยใช้อัลกอริธึม Apriori เครื่องมือที่ใช้ในการวิจัยคือ โปรแกรม WEKA กลุ่มตัวอย่างที่ใช้ในการ

วิเคราะห์ข้อมูลการเข้าเรียนระบบอีเลิร์นนิ่งของนักศึกษาจำนวน 270 คน คณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏจันทรเกษม ปีการศึกษา 1/2557 และใช้สถิติเปอร์เซ็นต์ไทล์ในการวิเคราะห์ข้อมูลด้วยอัลกอริทึม Apriori เพื่อการหาความสัมพันธ์ และหาประสิทธิภาพโดยใช้ค่าสนับสนุน และค่าความเชื่อมั่น

ผลการวิจัยพบว่า 1) พฤติกรรมผู้เรียนที่เรียนด้วยระบบอีเลิร์นนิ่ง วิชาการเตรียมฝึกประสบการณ์วิชาชีพ มีความสนใจใฝ่เรียนรู้สม่ำเสมอในการเข้าเรียนออนไลน์ พฤติกรรมผู้เรียนส่วนใหญ่มี การเตรียมความพร้อมก่อนไปฝึกประสบการณ์ จากข้อมูลการเข้าเรียนเรื่องการเขียนหนังสือราชการ การเขียนจดหมายสมัครงานและประวัติส่วนตัว และเรื่องภาษาอังกฤษเพื่อการสมัครงาน มากที่สุด 2) กฎความสัมพันธ์ระหว่างข้อมูลในระบบฐานข้อมูลอีเลิร์นนิ่ง พบว่าเมื่อผู้เรียนเข้าเรียนด้วยระบบอีเลิร์นนิ่ง เรื่องการเขียนหนังสือราชการ เป็นลำดับที่ 1 แล้วจะเข้าเรียนเรื่องการเขียนจดหมายสมัครงานและประวัติส่วนตัว เป็นลำดับที่ 2 และเรื่องภาษาอังกฤษเพื่อเตรียมความพร้อมในการสมัครงาน เป็นลำดับที่ 3 เสมอ โดยมีความเชื่อมั่นที่ 100เปอร์เซ็นต์ สรุปได้ว่ารูปแบบความสัมพันธ์ของพฤติกรรมผู้เรียนที่เรียนด้วยระบบอีเลิร์นนิ่ง สามารถใช้เป็นแนวทางในการปรับปรุงการจัดลำดับหัวข้อเนื้อหาในการเข้าเรียนบทเรียนให้มีความสัมพันธ์และสอดคล้องกับความต้องการของผู้เรียนอย่างแท้จริง อีกทั้งยังได้ประโยชน์ในการพัฒนาระบบจัดการเรียนการสอนด้วยระบบอีเลิร์นนิ่งให้มีประสิทธิภาพมากยิ่งขึ้น

วิรัตน์ดี พองเงิน และคณะ (2560) การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาเทคนิคเหมืองข้อมูลที่เหมาะสมในการพยากรณ์ปริมาณน้ำในเขื่อนและ 2) เปรียบเทียบผลการพยากรณ์ปริมาณน้ำรายเดือนในเขื่อนก๊วลม จังหวัดลำปาง โดยใช้เทคนิคเหมืองข้อมูล งานวิจัยนี้ได้นำข้อมูลที่เป็นปัจจัยที่มีผลต่อการเปลี่ยนแปลงระดับน้ำประกอบด้วย ปริมาณน้ำไหลเข้าเขื่อนปริมาณน้ำในเขื่อน ปริมาณการปล่อยน้ำและอัตราการระเหย โดยรวบรวมข้อมูลรายวันตั้งแต่ปี พ.ศ. 2535 – พ.ศ.2559 รวม 25 ปี จำนวน 9,300 รายการ โดยมีการแยกข้อมูลรายเดือนเพื่อนำมาพยากรณ์ด้วยเทคนิคการพยากรณ์

ผลการวิจัยพบว่า 1) เทคนิคเหมืองข้อมูลที่เหมาะสมในการพยากรณ์ปริมาณน้ำในเขื่อนประกอบด้วย 4 เทคนิค ได้แก่ เทคนิควิธีการวิเคราะห์การถดถอย (Regression Analysis) วิธีและโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) วิธีแบบจำลองต้นไม้เอ็มไพร์ฟี (Model Tree: M5P) และ วิธีเทคนิคซัพพอร์ตเวกเตอร์แมชชีนส์ (SVM) และ 2) ผลการเปรียบเทียบการพยากรณ์ปริมาณน้ำรายเดือนในเขื่อนก๊วลม จังหวัดลำปาง โดยใช้เทคนิคเหมืองข้อมูลทั้ง 4 เทคนิค พบว่า วิธีแบบจำลองต้นไม้เอ็มไพร์ฟี มีค่าสัมบูรณ์ของความ

คลาดเคลื่อนต่ำสุด ที่ 10.56 และเป็นวิธีที่เหมาะสมที่สุดสำหรับนำไปพัฒนาระบบพยากรณ์น้ำในเขื่อน ทั้งนี้เมื่อพิจารณาค่าสัมบูรณ์ของความคลาดเคลื่อนแต่ละเทคนิค พบว่า วิธีแบบจำลองต้นไม้เอมไพร์ฟิรวิธีชัพพอร์ตเวกเตอร์แมชชีนส์วิธีวิเคราะห์การถดถอย และวิธีโครงข่ายประสาทเทียม มีค่าความคลาดเคลื่อนเท่ากับ 10.56, 10.84, 11.12 และ 12.53 ตามลำดับ

ถาวร ทองเพชร (2560) อุบัติเหตุทางการจราจรถือเป็นปัญหาสำคัญทางสาธารณสุขของประเทศไทย ก่อให้เกิดการเสียชีวิตและ ทูพพลภาพ การใช้ข้อมูลทางระบาดวิทยาในการหาปัจจัยที่เกี่ยวข้องกับความรุนแรงของโรคถือเป็นข้อมูลสำคัญในการรักษาและป้องกันโรงพยาบาลศูนย์สุราษฎร์ธานีถือเป็นโรงพยาบาลหนึ่งที่ต้องรับกับปัญหาผู้ที่ได้รับบาดเจ็บจากอุบัติเหตุทางการจราจร การวิจัยครั้งนี้เป็นการศึกษาวิเคราะห์ข้อมูลย้อนหลังของผู้บาดเจ็บทางกระดูกและข้อจาก อุบัติเหตุทางการจราจรที่มารับบริการที่โรงพยาบาลศูนย์สุราษฎร์ธานี ตั้งแต่ วันที่ 1 มกราคม พ.ศ.2551 ถึง 31 ธันวาคม พ.ศ. 2558 โดยใช้วิธีการทางสถิติ Logistic regression เปรียบเทียบกับการใช้วิธีใหม่ การเรียนรู้ของเครื่อง (Machine Learning) และเปรียบเทียบประสิทธิภาพว่าวิธีใดสามารถทำนายการเสียชีวิตของผู้ป่วยกระดูกและข้อที่ได้รับการบาดเจ็บทางการจราจร ได้แม่นยำที่สุด โดยใช้เทคนิควิธีการวิเคราะห์ข้อมูลดังต่อไปนี้ 1) Decision Trees or Recursive Partitioning 2) Random Forests 3) Artificial Neural Network

จากผลการวิจัยพบว่ามีจำนวนผู้บาดเจ็บจากอุบัติเหตุการจราจรทั้งสิ้น 69,716 ราย และผู้ที่มีการบาดเจ็บทางกระดูกและข้อมีจำนวน 57,779 ราย คิดเป็นร้อยละ 82.87 และที่เข้ารับพักรักษาในโรงพยาบาลมีจำนวน 27,656 คน ซึ่งเสียชีวิตจำนวน 1,454 คน คิดเป็นร้อยละ 5.26 ของจำนวนผู้ที่เข้ารับการรักษาในโรงพยาบาล

ศิริกาญจน พิลาบุตร (2561) การทำเหมืองข้อมูล (data mining) หรือการค้นหาความรู้ในฐานข้อมูล (knowledge discovery in databases: KDD) เป็นกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักสถิติการรู้จากการเรียนรู้ของเครื่อง และหลักคณิตศาสตร์ซึ่งรูปแบบของข้อมูลที่นำมาทำเหมืองข้อมูลโดยส่วนใหญ่จะจัดเก็บในรูปแบบของระบบฐานข้อมูลเชิงสัมพันธ์ โดยผลลัพธ์ที่ได้จากการทำเหมืองข้อมูลมีลักษณะเป็นกฎสามารถนำไปทำนายผลได้งานวิจัยนี้เสนอแนวคิดในการทำนายผลการศึกษา โดยใช้กฎที่ได้จากการทำเหมืองข้อมูลประเภทการจำแนกมาเป็นตัวสร้างกฎในการทำนายผลการศึกษา เพื่อให้ทั้งอาจารย์และนักศึกษาตระหนักถึงปัจจัยที่มีผลต่อการสำเร็จการศึกษาในระดับปริญญาตรีวิทยาลัยนครราชสีมา

ในการทดลองได้ใช้ Training Data จำนวน 8,261 เร็คคอร์ดจากรายละเอียด Attribute ทั้งหมด เลือกใช้เพียง 3 Attribute ในการทำเหมืองข้อมูลคือ ประเภทการเข้าศึกษา, ภูมิภาค, สถานภาพของนักศึกษา ซึ่งจะพิจารณาประเภทการเข้าศึกษาและภูมิภาคของนักศึกษาว่ามีผลต่อการไม่สำเร็จการศึกษาของนักศึกษาระดับปริญญาตรีหรือไม่ จากการทดสอบความถูกต้องของ Rules โดยการนำไปใช้งานเพื่อทดสอบความถูกต้องของ Test Data พบว่ามีประโยชน์ในการทำนายผลการศึกษาวาปัจจุบันดังกล่าวมีผลต่อการสำเร็จการศึกษาของนักศึกษาหรือไม่ร้อยละ 58.85

จิระนันต์ เจริญรัตน์ (2560) ร้านอาหารบ้านฟ้าโปร่ง จ.สกลนคร เป็นร้านที่ให้บริการเบเกอรี่ กาแฟสด น้ำผลไม้ และอาหาร ซึ่งถือว่าเป็นร้านอาหารที่ได้มีการนำเอาระบบคอมพิวเตอร์มาใช้ในการจัดเก็บข้อมูลการซื้อขายสินค้าที่ลูกค้าซื้อในแต่ละครั้งไว้อย่างเป็นระบบ ส่วนด้านการตลาดส่วนใหญ่จะเป็นลักษณะลูกค้าบอกต่อแบบปากต่อปาก และลูกค้าบางส่วนได้นำเสนอภาพ และเมนูอาหารของร้านผ่านโซเชียลมีเดียและแชร์ต่อกันในกลุ่มคนที่รู้จัก ทางร้านยังไม่ได้มีการวางแผนการตลาดอย่างชัดเจน ไม่ได้มีการทำตลาดเชิงรุก ซึ่งในสถานการณ์ปัจจุบันที่มีการแข่งขันสูงเป็นสิ่งที่จำเป็นอย่างยิ่งที่จะต้องหาแนวทางในการวางแผนการตลาดให้ชัดเจนและสอดคล้องกับความต้องการของลูกค้ากลุ่มเป้าหมาย ดังนั้นผู้วิจัยจึงสนใจที่จะนำข้อมูลการสั่งซื้ออาหารของลูกค้าซึ่งเก็บไว้ในระบบฐานข้อมูลของร้าน มาวิเคราะห์หากฎความสัมพันธ์ของการสั่งอาหารด้วยเทคนิคเหมืองข้อมูล ทั้งนี้สามารถนำกฎความสัมพันธ์ที่ได้มาใช้ในการวางแผนการตลาด หรือจัดกิจกรรมส่งเสริมการขายที่มีความสอดคล้องกับความต้องการของลูกค้าและเหมาะสมกับพฤติกรรมการสั่งอาหารเพื่อให้ลูกค้าเกิดความพึงพอใจและกลับมาใช้บริการซ้ำรวมถึงการพัฒนาระบบแนะนำการสั่งอาหารต่อไปได้อีกด้วย

เทคนิคเหมืองข้อมูลได้ถูกนำมาประยุกต์ใช้ในงานธุรกิจอย่างแพร่หลาย ทั้งธุรกิจค้าส่ง ค้าปลีก ธุรกิจให้บริการ เป็นต้น งานวิจัยนี้มีวัตถุประสงค์เพื่อ วิเคราะห์กฎความสัมพันธ์ของการสั่งอาหารด้วยเทคนิคเหมืองข้อมูล ตามกรอบแนวคิด คริปส์-ดีเอ็ม (CRIPS-DM) ในการวิเคราะห์ข้อมูลใช้ข้อมูลรายการใบสั่งอาหารเฉพาะเบเกอรี่และเครื่องดื่มจากร้านอาหารบ้านฟ้าโปร่ง จ.สกลนคร ที่บันทึกไว้ในระบบฐานข้อมูล ระหว่างปี พ.ศ. 2557-2558 จำนวน 37,020รายการ ผู้วิจัยได้เลือกใช้โปรแกรม RapidMiner Studio 6.5 และใช้อัลกอริทึม FP-Growth เพื่อหากฎความสัมพันธ์ ผลการวิเคราะห์ แบ่งเป็น 2 กลุ่ม คือ 1) สั่งรับประทานที่ร้าน พบกฎที่น่าสนใจ 5 กฎ มีค่าความเชื่อมั่น 48, 45, 42, 42 และ 41 เปอร์เซนต์ตามลำดับ 2) สั่งกลับบ้าน พบกฎที่น่าสนใจ 9 กฎ มีค่าความเชื่อมั่น 59, 55, 55, 55, 54, 54, 53, 53 และ 51

เปอร์เซ็นต์ตามลำดับ จากผลการวิจัยพบว่ากฎความสัมพันธ์ที่ดีที่สุด คือ กลุ่มที่สั่งกลับบ้าน มีค่าความเชื่อมั่น 59 เปอร์เซ็นต์ นั่นคือเมื่อลูกค้าสั่งเค้กนมสด และ สั่งเค้กชาเขียวนมสด ก็จะสั่งเค้กมะพร้าวด้วย แสดงว่าจากกฎที่ได้ทำให้รู้ว่าอาหารอะไรเป็นจุดเด่นของร้านนี้ซึ่งทำให้สามารถจัดกิจกรรมส่งเสริมการตลาดที่สอดคล้องกับพฤติกรรมการสั่งซื้อของลูกค้าได้อย่างเหมาะสม

2.5 บทสรุป

จากแนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้องที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น ผู้วิเคราะห์ข้อมูลได้เลือกใช้ขั้นตอนกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบ Classification ด้วยการสร้างโมเดล Decision Tree เพื่อจัดกลุ่มข้อมูลที่มีความสัมพันธ์กัน โดยใช้โปรแกรมที่ใช้ทำเหมืองข้อมูล คือ โปรแกรม Weka 3.8.4 และโปรแกรม RapidMiner Studio 9.5.1 ในการสร้างโมเดล Decision Tree เพื่อทำการเปรียบเทียบ และเลือกผลลัพธ์ที่ดีที่สุดมานำเสนอ จากนั้นนำข้อมูลสารสนเทศมาทำการแสดงผลแบบ Visualization ในรูปแบบของภาพโดยใช้โปรแกรม Power BI เผยแพร่บน Web Browser ที่เป็นที่นิยมในยุคอินเทอร์เน็ตคือการเผยแพร่ทางสื่อออนไลน์ โดยใช้ภาษา HTML และ CSS ในการเขียนเว็บไซต์ขึ้นมา