บทที่ 2
แนวคิดทฤษฎีเครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงงานเรื่องการวิเคราะห์ข้อมูลเพื่อสนับสนุนการตัดสินใจการเพิ่มอัตราบุคลากรสายวิชาการ โดยใช้การนำเทคนิคเหมืองข้อมูล (Data mining) การจำแนกข้อมูล (Data classification) และรูปแบบของแผนภูมิต้นไม้ (Decision tree) เพื่อวิเคราะห์หารูปแบบที่เกิดขึ้นจากข้อมูล อาจารย์ผู้รับผิดชอบหลักสูตร อาจารย์ผู้สอน จำนวนภาระการสอนในหลักสูตร จำนวนภาระการสอนเฉลี่ยต่อคนอาจารย์เกษียณในปีถัดไป ซึ่งเก็บอยู่ในฐานข้อมูลบุคลากร และข้อมูลอีกส่วนเป็นข้อมูลนักศึกษาประกอบด้วย จำนวนนักศึกษาคงอยู่ จำนวนผู้สมัคร จำนวนรับสมัครนักศึกษา สัดส่วนนักศึกษาต่ออาจารย์ อัตราการแข่งขัน โดยมีการบันทึกข้อมูลในระบบทะเบียนกลางในส่วนงานรับสมัครนักศึกษา สามารถนำรูปแบบ (Model) ที่ได้มาวางแผนการเพิ่มอัตราบุคลากรสายวิชาการของคณะบริหารธุรกิจและศิลปะศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ และนำไปสู่การการเพิ่มอัตราบุคลากรสายวิชาการต่อไปในอนาคต

ในบทนี้เป็นการนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่ เกี่ยวข้องของการวิเคราะห์ข้อมูลเพื่อสนับสนุนการตัดสินใจการเพิ่มอัตราบุคลากรสายวิชาการนำไปสู่การแสดงผลข้อมูลบนเว็บไซต์ ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง กับการวิเคราะห์ข้อมูล เพื่อใช้เป็นแนวทางการศึกษาประกอบด้วยรายละเอียดตามลำดับ ดังนี้

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data analytic)

2.1.2 แนวคิดเกี่ยวกับการประยุกต์ใช้ข้อมูลทางสถิติ

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data visualization)

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

2.2.2 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

2.2.3 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS และ Bootstrap

2.2.4 ทฤษฎีเกี่ยวกับการ Visualization

2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.3.1 ระบบสนับสนุนการตัดสินใจ (Decision Support System)

2.3.2 การตัดสินใจแบบโครงสร้างต้นไม้ (Decision Tree)

2.3.3 ซอฟต์แวร์ Weka

2.3.4 ซอฟต์แวร์ KNIME

2.3.5 ซอฟต์แวร์ Power BI

2.4 วรรณกรรมที่เกี่ยวข้อง

2.5 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data analytic)

ความหมายของการวิเคราะห์ข้อมูล เอื้อมพร หลินเจริญ (2555) ได้ให้ความหมายการวิเคราะห์ข้อมูลเชิงคุณภาพไว้ว่า การวิเคราะห์ข้อมูลเชิงคุณภาพ นับเป็นขั้นตอนที่สำคัญในกระบวนการวิจัย วิธีการหลักที่ใช้ใน การวิเคราะห์ข้อมูลเชิงคุณภาพ เป็นวิธีการสร้างข้อสรุปจากการศึกษาจากข้อมูลจำนวนหนึ่ง ซึ่งมักไม่ใช้สถิติในการวิเคราะห์หรือถ้าใช้สถิติก็ไม่ได้ถือว่าสถิติเป็นวิธีการวิเคราะห์หลัก แต่จะ ถือเป็นข้อมูลเสริมในการวิเคราะห์ ข้อมูลเชิงคุณภาพนั้น คณะผู้จัดทำนับเป็นผู้ที่มีบทบาทสำคัญยิ่งและควรมีความรอบรู้ในเรื่องแนวคิดทฤษฎีอย่างกว้างขวาง มีความเป็นสหวิทยากร อยู่ในตัวเอง มีความสามารถทางภาษาสามารถเชื่อมโยงข้อความและสร้างข้อสรุปเป็นกรอบแนวคิดและตีความหมายของข้อมูลได้หลายๆแบบทั้งนี้การวิเคราะห์ข้อมูลเชิงคุณภาพนั้น อาจไม่จำเป็นต้องใช้เฉพาะการวิจัยเชิงคุณภาพเท่านั้น แต่อาจใช้กับการวิจัยเชิงปริมาณที่มีการ เก็บรวบรวมข้อมูลเชิงคุณภาพ เช่น แบบสอบถามปลายเปิดการสัมภาษณ์ การสังเกต จดบันทึกมาทำการวิเคราะห์เพื่อให้ข้อมูลที่รวบรวมมามีความหมายและตอบคำถามหรือจุดมุ่งหมายของการวิจัยสำหรับสาระในการวิเคราะห์ข้อมูลเชิงคุณภาพโดยเทคนิคการ วิเคราะห์

ข้อมูลเชิงคุณภาพที่สำคัญ ๆ และนักวิจัย นิยมใช้ ได้แก่ การจำแนกหรือการจัดกลุ่ม ข้อมูล การเปรียบเทียบเหตุการณ์ การวิเคราะห์ส่วนประกอบ การวิเคราะห์แบบอุปนัย และการวิเคราะห์ข้อมูลที่เป็นเอกสาร

หัสพร ทองแดง, พิศุทธิภา เมธีกุล, สมคิด ทุมวงศ์ และแพรว สมบัติใหม่ (2559) ได้ให้ความหมายการวิเคราะห์ข้อมูลเชิงคุณภาพไว้ว่า การวิเคราะห์ข้อมูลเป็นการจำแนกข้อมูล ให้เป็นระบบ เพื่อทำให้เกิดความเข้าใจในความหลากหลาย ความหมายและความสัมพันธ์ของข้อมูลในบริบทของสังคมและวัฒนธรรมหนึ่ง ๆ ในการวิเคราะห์ข้อมูลเชิงคุณภาพนี้หลังจากที่มีการแยกข้อมูลออกได้เป็นหมวดหมู่แล้ว นักวิจัยต้องใช้วิธีการนำเสนอโดยการพรรณนาให้เห็นสภาพเงื่อนไขกระบวนการ ขั้นตอนการสัมพันธ์ต่าง ๆ รวมทั้งพยายามหาความหมายทางวัฒนธรรมของปรากฏการณ์ในทัศนะของบุคคลหรือกลุ่มตัวอย่างซึ่งเป็นผู้ให้ข้อมูล และอาจมีความเข้าใจและตีความแตกต่างกันออกไปหรือแตกต่างไปจากของผู้วิจัยเอง การวิเคราะห์ข้อมูลสำหรับการวิจัยเชิงคุณภาพ จึงเป็นเสมือนการทำความเข้าใจในแบบแผนและความหมายของพฤติกรรมมากกว่าที่จะมุ่งหาระดับความมากน้อยหรือระดับความเข้มข้นของความสัมพันธ์ของปรากฏการณ์จึงเป็นการวิเคราะห์ข้อมูลที่อาศัยเครื่องมือทางความคิด (Conceptual tools)มากกว่าที่จะใช้เครื่องมือทางสถิติและที่สำคัญต้องอาศัยความสามารถของนักวิจัยที่จะอธิบายให้เห็นความสัมพันธ์และความหมายของปรากฏการณ์บนพื้นฐานความเข้าใจในวัฒนธรรมและสังคมที่เป็นบริบทของปรากฏการณ์ที่ศึกษา

จุมพล หนิมพานิช และวรวลัญช์ โรจนพล (2561) ได้ให้ความหมายการวิเคราะห์ข้อมูลเชิงคุณภาพไว้ว่า การวิเคราะห์ข้อมูลเชิงคุณภาพโดยทั่วไปหมายถึง การสรุปเนื้อหาสาระของข้อมูลและหาแบบแผนของความสัมพันธ์ระหว่างข้อมูลหรือตัวแปร หรือหมายถึงการสรุปพรรณนาและการวิเคราะห์ความสัมพันธ์ระหว่างข้อมูล หรือหมายถึง การนำเอาข้อมูลที่ได้จากการค้นคว้าวิจัยมาจัดกระทำให้เป็นระบบและหาความหมาย แยกแยะ องค์ประกอบ รวมทั้งเชื่อมโยงและหาความสัมพันธ์ของข้อมูลเพื่อให้สามารถนาไปสู่ความเข้าใจต่อการดำรงอยู่และการเปลี่ยนแปลงของปรากฏการณ์ที่ศึกษา หรือหมายถึง การแยกแยะ การตีความหมาย การเปรียบเทียบ การหาความสัมพันธ์เกี่ยวกับปรากฏการณ์ การหาแบบแผนการอธิบายและการสรุปเกี่ยวกับกิจกรรมหรือพฤติกรรมในสภาพสังคมหรือปรากฏการณ์ทางสังคมที่ศึกษา

2.1.2 แนวคิดเกี่ยวกับการประยุกต์ใช้ข้อมูลทางสถิต

จากเอกสารประกอบการสอนการวิจัยสำหรับครู ได้ให้ความหมายเกี่ยวกับสถิติไว้ว่า คำว่า สถิติ (Statistics) มาจากภาษาเยอรมันว่า Statistics มีรากศัพท์มาจาก Stat หมายถึงข้อมูลหรือสารสนเทศ ซึ่งจะอำนวยประโยชน์ต่อการบริหารประเทศในด้านต่าง ๆ เช่น การทำสำมะโนครัวเพื่อจะทราบจำนวนพลเมืองในประเทศทั้งหมด ในสมัยต่อมา คำว่า สถิติ ได้หมายถึง ตัวเลขหรือข้อมูลที่ได้จากการเก็บรวบรวม เช่น จำนวนผู้ประสบอุบัติเหตุบนท้องถนนอัตราการเกิดของเด็กทารก ปริมาณน้ำฝนในแต่ละปี เป็นต้น สถิติในความหมายที่กล่าวมานี้เรียกอีกอย่างหนึ่งว่า ข้อมูลทางสถิติ (Statistical data)

1) ค่ากึ่งกลาง (Median) เป็นการวัดแนวโน้มเข้าสู่ส่วนกลางวิธีหนึ่งที่ใช้การเรียงค่าการสะท้อนของจุดภาพจากค่าน้อยที่สุดไปหาค่ามากที่สุด โดยค่ากึ่งกลางเป็นค่าที่อยู่ในตำแหน่งกึ่งกลางของข้อมูลทั้งชุด ค่ากึ่งกลางจึงเป็นตัวแทนค่าการสะท้อนของจำนวนจุดภาพทั้งหมดในช่วงคลื่นหนึ่ง ๆ ที่แสดงให้ทราบว่ามีจำนวนจุดภาพที่มีค่าการสะท้อนมากกว่าและน้อยกว่าค่ากึ่งกลางอยู่ประมาณร้อยละ 50

2) ค่าฐานนิยม (Mode) เป็นการวัดแนวโน้มเข้าสู่ส่วนกลางอีกวิธีหนึ่ง โดยดูจากจำนวนความถี่ของค่าการสะท้อนซึ่งมีความถี่สูงที่สุด นิยมนำมาใช้กับข้อมูลที่เป็นนามบัญญัติ เช่น ค่าของประเภทข้อมูลหลังจากการจำแนกประเภทแล้ว ถือเป็นค่าการสะท้อนที่แสดงการใช้ที่ดินประเภทต่าง ๆ ไม่ใช่ค่าการสะท้อนของวัตถุอีกต่อไป

3) ความเบี่ยงเบนมาตรฐาน (Standard Deviation : S.D.) เป็นการวัดการกระจายที่นิยมใช้มากที่สุดการคำนวณใช้วิธียกกำลังสองของผลต่างระหว่างค่าการสะท้อนของทุกจุดภาพในแต่ละช่วงคลื่นกับค่าเฉลี่ยเลขคณิตของช่วงคลื่นนั้น

ภัทรธิดา ผลงาม (2558) ได้ให้ความหมายเกี่ยวกับสถิติไว้ว่า สถิติ คือ ตัวเลขต่าง ๆ ที่ได้มีการรวบรวมขึ้นเพื่อบอกข้อเท็จจริงเกี่ยวกับคุณสมบัติหรือลักษณะบางสิ่งบางอย่างที่สามารถแสดงออกเป็นตัวเลขได้สถิติแบ่งเป็นประเภทใหญ่ ๆ ได้ 2 ประเภท ดังนี้ คือ

1) สถิติพรรณนา (Descriptive statistics) คือสถิติที่บรรยายถึงลักษณะของข้อมูลเฉพาะกลุ่มนั้น ๆ โดยไม่สรุปอ้างอิงไปยังประชากรกลุ่มอื่น ๆ สถิติประเภทนี้นิยมศึกษาในกลุ่มเล็กหรือกลุ่มใหญ่ก็ได้ สถิติประเภทนี้เป็นสถิติที่บรรยายลักษณะของข้อมูล เช่น ค่าเฉลี่ย ร้อยละ มัธยฐาน พิสัย ส่วนเบี่ยงเบนมาตรฐาน และการวัดความสัมพันธ์ต่าง ๆ เช่น สหสัมพันธ์

2) สถิติอ้างอิงหรือสถิติอนุมาน (Inferential or inductive statistics) คือ สถิติที่นำค่าสถิติพรรณนามาสรุปอ้างอิงไปยังประชากร หรือเป็นสถิติที่ศึกษาจากกลุ่มตัวอย่างแล้วนำผลอ้างอิงไปยังกลุ่มประชากร เช่น การศึกษาความคิดเห็นต่อการเลือกตั้งของประชาชนไม่จำเป็นต้องศึกษาจากประชาชนทุกคน แต่สามารถเลือกศึกษาจากประชาชนบางกลุ่มซึ่งจะเป็นตัวแทนของประชาชนทั้งหมด แล้วจึงสรุปว่าประชาชนมีความคิดเห็นอย่างไรต่อการเลือกตั้งได้ดังนั้นในการใช้สถิติอ้างอิงนี้จึงจำเป็นต้องเลือกลุ่มตัวอย่างที่เหมาะสม สถิติอ้างอิงนี้ก่อนนำไปอ้างอิงกลุ่มประชากรต้องมีการทดสอบทางสถิติก่อนทุกครั้งจึงสามารถอ้างอิงประชากรได้อย่างมีประสิทธิภาพ

สุทิน ชนะบุญ (2555) ได้ให้ความหมายเกี่ยวกับสถิติไว้ว่า สถิติประกอบด้วย 2 องค์ประกอบ ดังนี้ สถิติเชิงพรรณนา (Descriptive Statistics) เป็นวิธีการทางสถิติเพื่อใช้ในการพรรณนาหรือบรรยายลักษณะของสิ่งที่ศึกษา เพื่อให้เข้าใจถึงลักษณะของข้อมูลที่เก็บรวบรวมมาได้ จะพรรณนาภายในขอบเขตของข้อมูลที่เก็บรวบรวมมาเท่านั้น ไม่สามารถจะคาดคะเนลักษณะต่าง ๆ ออกไปนอกเหนือจากข้อมูลที่มีอยู่ได้ หรือไม่มีการอ้างอิงหรืออนุมานไปถึงกลุ่มอื่น หากผู้วิจัยสามารถศึกษาทุกหน่วยของประชากรได้ ก็จะใช้สรุปหรือบรรยายลักษณะของประชากรที่ศึกษา หากผู้วิจัยไม่สามารถศึกษาสมาชิกทุกหน่วยของประชากรได้ สุ่มสมาชิกเพียงบางส่วนมาศึกษา (ศึกษาจากกลุ่มตัวอย่าง) ก็จะใช้สรุปหรือบรรยายเฉพาะกลุ่มตัวอย่างที่สุ่มมาศึกษาเท่านั้น สถิติเชิงอ้างอิงหรือสถิติอนุมาน (Inferential Statistics) เป็นวิธีการทางสถิติที่ใช้ในการสรุปลักษณะของประชากร จากผลการศึกษาข้อมูลในกลุ่ม ตัวอย่าง โดยอาศัยทฤษฎีความน่าจะเป็นในการอนุมานลักษณะประชากร โดยศึกษาจากกลุ่มตัวอย่าง แต่อ้างอิง หรืออนุมานไปถึงประชากร ประกอบด้วย 2 องค์ประกอบ คือ การประมาณค่าพารามิเตอร์ (Parametric Estimation) และการทดสอบสมมติฐาน (Testing Hypothesis)ทฤษฎีความน่าจะเป็นในการอนุมานลักษณะประชากร โดยศึกษาจากกลุ่มตัวอย่าง แต่อ้างอิง หรืออนุมานไปถึงประชากร ประกอบด้วย 2 องค์ประกอบ คือ การประมาณค่าพารามิเตอร์ (Parametric Estimation) และการทดสอบสมมติฐาน (Testing Hypothesis)

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data visualization)

ความหมายการแสดงผลข้อมูล

กานต์ ยงศิริวิทย์ และภาคภูมิ ชัยศิริประเสริฐ (2560) ได้ให้ความหมายของการแสดงผลข้อมูลไว้ว่า การแสดงผลข้อมูล (Data Visualization) ผลลัพธ์ที่ได้จากการวิเคราะห์ ข้อมูลควร

นำเสนอในรูปแบบที่น่าสนใจ และเข้าใจง่าย ดังนั้นเครื่องมือควรที่จะสามารถนำเสนอข้อมูลในรูปแบบของกราฟและแสดงผลลัพธ์บนแผนที่ในขั้นพื้นฐานได้ โดยไม่ควรที่จะต้องมีการติดตั้งซอฟต์แวร์หรือตัวช่วยเพิ่มเติม

อาทิตย์ สิทธิบรรเจิด (2552) ได้ให้ความหมายของการแสดงผลข้อมูลไว้ว่า การแสดงผลข้อมูล (Data Visualization) เป็นส่วนประกอบสำคัญใน Cognitive System ซึ่งเป็นส่วนในการแสดงข้อมูลหรือผลลัพธ์ต่าง ๆ ในระหว่างคอมพิวเตอร์และผู้ใช้งานในรูปแบบของภาพโดยผู้ใช้สามารถเรียนรู้และจดจำข้อมูลผ่านการมองได้มากกว่าการใช้ประสาทสัมผัสอื่น ๆ หรือจะกล่าวว่า Visualization ก็คือ การสร้างมโนภาพของสิ่งต่าง ๆ ที่เราสนใจขึ้นมาในใจ ซึ่งต่อมาได้กลายเป็นการนำภาพมาใช้กับการนำเสนอหรือนามาเป็นกรอบความคิด ซึ่งได้นำไปใช้ในการสนับสนุนการตัดสินใจ

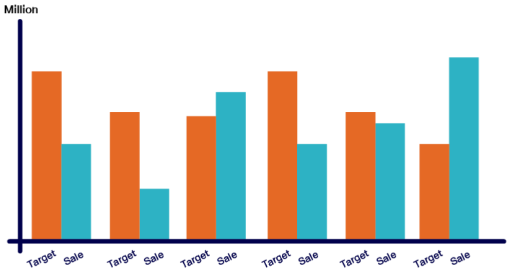
ชนาธิป ชื่นมนัส (2552) ได้ให้ความหมายของการแสดงผลข้อมูลไว้ว่า การ แสดงผลข้อมูล (Data Visualization) คือวิธีการที่ใช้ในการนำข้อมูลที่เป็นนามธรรมมาแสดงให้ เป็นภาพเพื่อเพิ่มประสิทธิผลในการติดต่อสื่อสาร หรือทำให้เข้าใจข้อมูลได้ง่ายขึ้น การใช้ เทคนิคVisualization มีหลายวิธี ซึ่ง แต่ละวิธีเหมาะสมที่จะใช้ในการแสดงข้อมูลที่แตกต่างกัน การเลือกใช้เทคนิคใดนั้นขึ้นกับองค์ประกอบหลายอย่าง เช่น ชนิดของข้อมูลที่จะนำมาแสดงวิธีที่ต้องการแสดงข้อมูล เป็นต้น ในปัจจุบันมีงานวิจัยที่นำเอาเทคนิค Visualization มาใช้โดยมีจุดประสงค์เพื่อเพิ่มประสิทธิผลในการแสดงข้อมูลการติดต่อสื่อสารภายในเครือข่ายคอมพิวเตอร์ และช่วยในการวิเคราะห์ความผิดปกติที่เกิดขึ้นในเครือข่าย

การเลือกรูปแบบ Visualization ให้เหมาะสมกับข้อมูล ในปัจจุบันเป็นยุคเทคโนโลยีเข้าถึงทุกคน ทำให้การรับรู้ข่าวสาร ข้อมูลต่าง ๆ เป็นไปได้ง่าย และรวดเร็วมากขึ้น คนที่นำเสนอข้อมูลจึงต้องนำเสนอข้อมูลให้น่าสนใจ เข้าใจง่าย และรวดเร็ว จึงเกิดการสร้าง Data Visualization ขึ้นมา Data Visualization เป็นการใช้ภาพเพื่อแสดงข้อมูลในเชิงปริมาณที่วัดได้ ซึ่งอาจนำเสนอออกมาในรูปแบบ แผนภูมิ กราฟ กราฟิก และอื่น ๆ อีกมากมาย เพื่อให้เข้าใจได้

1) แผนภูมิแท่ง (Bar Charts) เป็นแผนภูมิที่ประกอบด้วยแกนนอน แกนตั้ง ที่นิยมแสดงออกมาในรูปแท่งสี่เหลี่ยมที่สามารถบอกความสูงได้ เหมาะสำหรับใช้การเปรียบเทียบจำนวนของข้อมูลในแต่ละชุด เช่น รายรับในแต่ละเดือน, ยอดขายที่ขายได้จริงเปรียบเทียบกับเป้าหมายยอดขายที่ตั้งไว้ เป็นต้น ซึ่งแผนภูมิแท่งยังสามารถแบ่งออกเป็นหลายประเภท

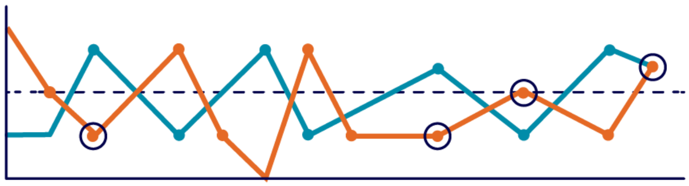
– แผนภูมิแท่งแบบจัดกลุ่ม แผนภูมิแท่งแบบจัดกลุ่ม เหมาะสมกับการนำเสนอข้อมูลที่มีข้อมูลย่อย ๆ อยู่ภายใต้ข้อมูลใหญ่ เป็นการเน้นให้เห็นข้อมูลย่อยนั้น ๆ

– แผนภูมิแท่งแบบวางซ้อนกัน แผนภูมิแท่งแบบวางซ้อนกัน เหมาะสมกับการนำเสนอข้อมูลที่ต้องการให้เห็นข้อมูลย่อยในแต่ละข้อมูลใหญ่และยังแสดงให้เห็นสัดส่วนของข้อมูลย่อยต่าง ๆ เหล่านั้นได้ด้วย ใช้แผนภูมินี้เมื่อมีชุดข้อมูล หลายชุดและต้องการเน้นผลรวมทั้งหมด



ภาพที่ 2.1 แผนภูมิแท่ง (Bar Chart)
ที่มา : https://www.9experttraining.com/

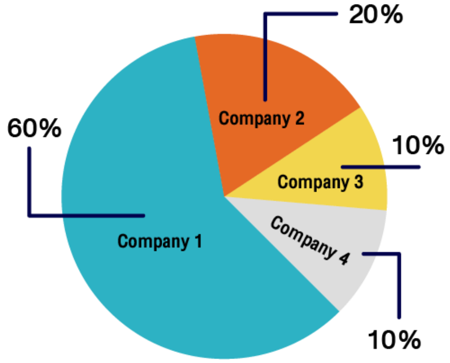
2) แผนภูมิเส้น (Line Charts) แผนภูมิเส้น มีลักษณะคล้ายแผนภูมิแท่ง ซึ่งประกอบด้วยแกนตั้งและนอน เพียงแต่เปลี่ยนจากแท่งข้อมูลเป็นจุดบนแผนภูมิแผนภูมิประเภทนี้เหมาะกับการนำเสนอข้อมูลตัวเลขที่มีความสัมพันธ์กับข้อมูลที่มีลักษณะเป็นช่วง ใช้แสดงให้เห็นการเปลี่ยนแปลงที่เกิดขึ้นตามเวลาเพื่อดูแนวโน้ม รวมถึงสามารถใช้พยากรณ์แนวโน้มในอนาคตได้ เช่น ข้อมูลของยอดขายในแต่ละปี หรือไตรมาส และนำมาวิเคราะห์เพื่อดูแนวโน้ม เป็นต้น



ภาพที่ 2.2 แผนภูมิเส้น (Line Charts)
ที่มา : https://www.9experttraining.com/

3) แผนภูมิวงกลม (Pie Charts) แผนภูมิวงกลมเหมาะกับการนำเสนอข้อมูลที่มีส่วนประกอบย่อยที่รวมกันเป็นส่วนใหญ่ มีการแบ่งส่วนให้ดูง่าย และสวยงามแต่ในทางกลับกัน

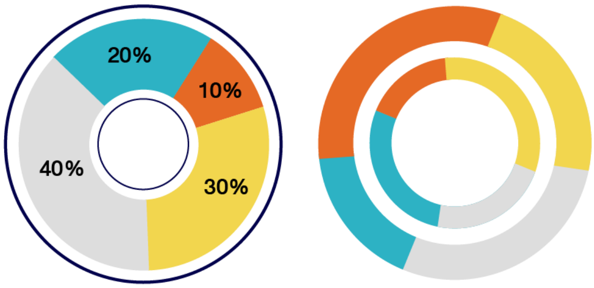
อาจจะดูยากในเรื่องของการประมาณขนาดของแต่ละชิ้น ยิ่งถ้ามีจำนวนชิ้นมาก จะยิ่งแยกยาก เพราะต้องใช้หลายสี ในการนำเสนอข้อมูล เช่น ส่วนแบ่งทางการตลาด (Market Share), ข้อมูลแสดงส่วนผสมต่าง ๆ เป็นต้น



ภาพที่ 2.3 แผนภูมิวงกลม (Pie Charts)

ที่มา : https://www.9experttraining.com/

4) แผนภูมิโดนัท (Doughnut Charts) แผนภูมิโดนัทมีหลักการออกแบบเช่นเดียวกับแผนภูมิวงกลมแต่สามารถแสดงชุดข้อมูลได้มากกว่า 1 ชุด โดยนำเสนอข้อมูลเป็นวงกลมซ้อนกันหลาย ๆ ชั้น

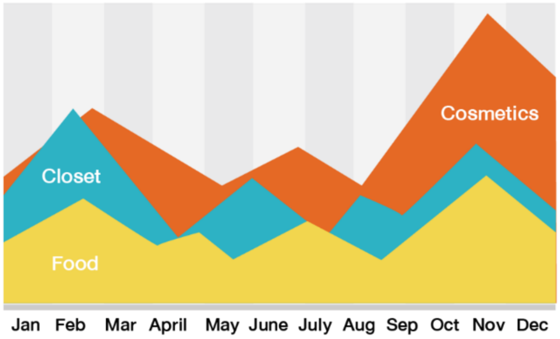


ภาพที่ 2.4 แผนภูมิโดนัท (Doughnut Charts)

ที่มา : https://www.9experttraining.com/

5) แผนภูมิพื้นที่ (Area Charts) มีหน้าตาคล้ายแผนภูมิเส้น แต่มีการแรเงาพื้นที่ใต้เส้นข้อมูลหรือระหว่าง 2 เส้นเพื่อแสดงให้เห็นปริมาณความแตกต่างระหว่างเส้น เหมาะสำหรับเน้นความสำคัญของการเปลี่ยนแปลงในช่วงเวลาแสดงให้เห็นผลรวมของความแตกต่างระหว่างข้อมูล

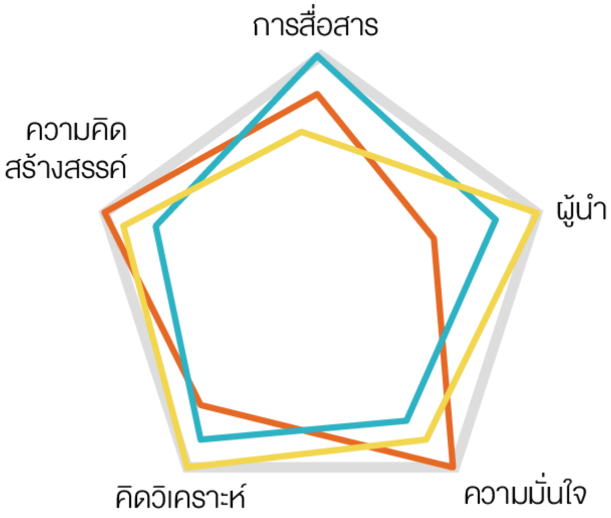
เช่น ข้อมูลของการซื้อสินค้าในห้างสรรพสินค้าในแต่เดือน ตามหมวดหมู่ต่าง ๆ จนไปถึงเครื่องสำอาง เสื้อผ้าแฟชั่น อาหาร ตามลำดับ



ภาพที่ 2.5 แผนภูมิพื้นที่ (Area Charts)

ที่มา : https://www.9experttraining.com/

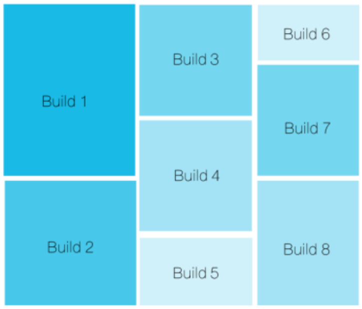
6) แผนภูมิเรดาร์ (Radar Charts) มีลักษณะคล้ายแผนภูมิเส้นที่มีการแสดงผลแบบวงกลม จำนวนเหลี่ยมของเรดาร์เท่ากับจำนวนหัวข้อของข้อมูล แผนภูมินี้ไม่ได้บอกถึงความต่อเนื่องของข้อมูล แต่เหมาะสำหรับการนำเสนอข้อมูลเป็นหัวข้อ แล้วนำมาวิเคราะห์หาจุดอ่อน จุดแข็งของข้อมูล เช่น นำเสนอข้อมูลเกี่ยวกับคุณสมบัติของการรับพนักงานใหม่ เพื่อดูจุดอ่อนจุดแข็งของแต่ละคน เป็นต้น



ภาพที่ 2.6 แผนภูมิเรดาร์ (Radar Charts)

ที่มา : https://www.9experttraining.com/

7) แผนภูมิต้นไม้ (Tree Maps) คือการนำเสนอข้อมูลแบบแสดงให้เห็นพื้นที่ แสดงผลได้ในแบบลำดับชั้น เหมือนแบบโครงสร้างต้นไม้ อาจจะนำเสนอข้อมูลที่ต้องการให้เห็นถึงเขตพื้นที่ แสดงพื้นที่สีที่แตกต่างกันได้

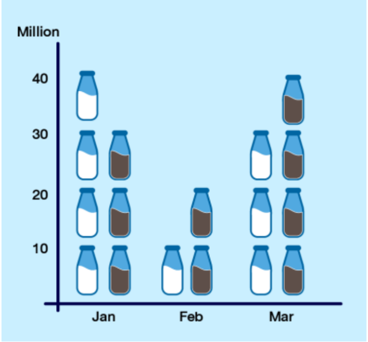


ภาพที่ 2.7 แผนภูมิต้นไม้ (Tree Maps)

ที่มา : https://www.9experttraining.com/

8) แผนภูมิรูปภาพ (Picture Graph) เป็นแผนภูมิที่ประกอบไปด้วยแกนนอน และแกนตั้ง แต่เลือกใช้รูปภาพ หรือไอคอนแทนจำนวนของสิ่งของนั้น ๆ เช่น การแสดงผลจำนวนของนมที่ขายได้ในแต่ละเดือน โดยนำเสนอทั้งนมรสจืด รสช็อกโกแลต เปรียบเทียบในแต่ละเดือนซึ่งมีการนำเสนอลักษณะคล้ายกับกราฟแท่ง แต่เปลี่ยนจากแท่งเป็นรูปกราฟิกของนม 2 รสชาติแทน ก็ทำให้การนำเสนอข้อมูลน่าสนใจมากยิ่งขึ้น

ซึ่งแนวทางการนำเสนอข้อมูลลักษณะนี้ต้องอาศัยความคุ้นชินของคนดู เพื่อแทนสัญลักษณ์ภาพลงไป เช่น เมื่อพูดถึงจำนวนคน อาจจะแทนด้วยภาพไอคอนคนหรือเมื่อพูดถึงจำนวนเงิน ควรแทนภาพเป็นเหรียญเงิน หรือแบงก์แทน ก็จะทำให้คนดูเข้าใจง่ายจากสัญลักษณ์ภาพที่คุ้นเคยอยู่แล้วและยังดึงดูดความสนใจได้มากกว่าการใช้กราฟแท่งสี่เหลี่ยม



ภาพที่ 2.8 แผนภูมิรูปภาพ (Picture Graph)

ที่มา : https://www.9experttraining.com/

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

เหมืองข้อมูล คือกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น ในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ใน งานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และการแพทย์รวมทั้งในด้านเศรษฐกิจและสังคม

2.2.2.1 การทำเหมืองข้อมูล (Data Mining) เปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลอย่างง่ายๆ มาสู่การจัดเก็บในรูปแบบฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนถึงการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล หรือจะแยกเป็นข้อ ๆ ได้ดังนี้

1) กระบวนการหรือการเรียงลำดับของการค้นข้อมูลจำนวนมากและเก็บข้อมูลที่เกี่ยวข้อง

2) การนำมาใช้โดยหน่วยงานทางธุรกิจและนักวิเคราะห์ทางการเงินหรือ การนำมาใช้งานในด้านวิทยาศาสตร์เพื่อเอาข้อมูลขนาดใหญ่ที่สร้างโดยวิธีการทดลองและการสังเกตการณ์ที่ทันสมัย

3) การสกัดหรือแยกข้อมูลที่เป็นประโยชน์จากข้อมูลขนาดใหญ่หรือฐานข้อมูล

4) การวางแผนทรัพยากรขององค์กรโดยสามารถวิเคราะห์ทางสถิติและ ตรรกะของข้อมูลขนาดใหญ่เป็นการมองหารูปแบบที่สามารถช่วยการตัดสินใจได้

2.2.2.2 ขั้นตอนการทำเหมืองข้อมูล

ประกอบด้วยขั้นตอนการทำงานย่อยที่จะเปลี่ยนข้อมูลดิบให้กลายเป็นความรู้ ประกอบด้วยขั้นตอนดังนี้

1) Business Understanding เน้นไปที่การทำความเข้าใจในงาน ระบุโอกาส และหาปัญหาที่จะเกิดขึ้น กำหนดขอบเขตของข้อมูลที่จะนำวิเคราะห์ ซึ่งต้องสามารถระบุ ผลลัพธ์ที่มีได้

2) Data Understanding ทำความเข้าใจข้อมูลโดยการรวบรวมข้อมูลที่เกี่ยวข้อง คัดเลือกให้เหลือเพียงข้อมูลที่มีความถูกต้องและสำคัญต่องานมาทำการวิเคราะห์

3) Data Preparation ทำการแปลงข้อมูล (Raw Data) ให้กลายเป็นข้อมูล ที่สามารถนำมาช่วยในการวิเคราะห์ต่อไปได้ ขั้นตอนนี้จะใช้เวลามากที่สุดในทุกขั้นตอนเพราะ

คุณภาพของงานที่ได้จะดีเพียงใดขึ้นอยู่กับคุณภาพข้อมูลที่จัดเตรียมในขั้นนี้ การเตรียมข้อมูลประกอบด้วย การคัดเลือกข้อมูล การกลั่นกรองข้อมูล และแปลงรูปแบบของข้อมูล

4) Modeling การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูลที่ได้จากขั้นตอนที่ 3 พร้อมทดสอบผลลัพธ์แบบจำลองเพื่อให้ได้คำตอบที่ดีที่สุด บางครั้งอาจมีการย้อนกลับไปปรับ การเตรียมข้อมูลเพื่อให้ได้แบบจำลองที่เหมาะสมที่สุด

5) Evaluation การประเมินผลลัพธ์ที่ได้ก่อนที่จะนำไปใช้จริง ว่าตรงกับวัตถุประสงค์หรือเป้าหมายที่ได้ตั้งไว้หรือมีความน่าเชื่อถือมากน้อยเพียงใด หากไม่ได้ผลลัพธ์ตามวัตถุประสงค์ต้องย้อนกลับไปปรับปรุงแก้ไขการดำเนินงานในขั้นตอนก่อนหน้า

6) Deployment การนำเอาข้อมูลที่เป็นผลลัพธ์จากทั้งหมด มาใช้ประโยชน์ตามวัตถุประสงค์ และทำการประเมินผลลัพธ์ที่ได้ว่ามีประสิทธิภาพตรงตามที่ตั้งไว้หรือไม่

2.2.3 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

2.2.3.1 หลักในการออกแบบเว็บไซต์ หน้าเว็บเป็นสิ่งแรกที่ผู้ใช้จะได้เห็นขณะที่เปิดเข้าสู่เว็บไซต์ และยังเป็นสิ่ง แรกที่แสดงถึงประสิทธิภาพในการออกแบบเว็บไซต์อีกด้วย หน้าเว็บจึงเป็นสิ่งสำคัญมากเพราะเป็นสื่อกลางให้ผู้ชมสามารถใช้ประโยชน์จากข้อมูลของระบบงานของเว็บไซต์นั้นได้ โดยปกติหน้าเว็บจะประกอบด้วย รูปภาพ ตัวอักษร สีพื้น ระบบเนวิเกชัน และองค์ประกอบอื่น ๆ ที่ช่วยสื่อความหมายของเนื้อหาและอำนวยความสะดวกต่อการใช้งานหลักสำคัญในการออกแบบหน้าเว็บก็คือ การใช้รูปภาพและองค์ประกอบต่าง ๆ ร่วมกันเพื่อสื่อความหมายเกี่ยวกับเนื้อหาหรือลักษณะสำคัญของเว็บไซต์ โดยมีเป้าหมายสำคัญเพื่อการสื่อความหมายที่ชัดเจนและน่าสนใจ บนพื้นฐานของความเรียบง่ายและความสะดวกของผู้ใช้ การออกแบบเว็บไซต์ ต้องคำนึงถึง

1) ความเรียบง่าย ได้แก่ มีรูปแบบที่เรียบง่าย ไม่ซับซ้อน และใช้งานได้สะดวก ไม่มีกราฟิกหรือตัวอักษรที่เคลื่อนไหวอยู่ตลอดเวลา ชนิดและสีของตัวอักษรไม่มากจนเกินไปทำให้วุ่นวาย

2) ความสม่ำเสมอ ได้แก่ ใช้รูปแบบเดียวกันตลอดทั้งเว็บไซต์ เช่น รูปแบบของหน้าสไตล์ของกราฟิก ระบบเนวิเกชันและโทนสี ควรมีความคล้ายคลึงกันตลอดทั้งเว็บไซต์

3) ความเป็นเอกลักษณ์ การออกแบบเว็บไซต์ควรคำนึงถึงลักษณะของ องค์กรเพราะรูปแบบของเว็บไซต์จะสะท้อนถึงเอกลักษณ์และลักษณะขององค์กรนั้น ๆ เช่น ถ้าเป็นเว็บไซต์ของทาง ราชการ จะต้องดูน่าเชื่อถือไม่เหมือนสวนสนุก ฯลฯ

4) เนื้อหาที่มีประโยชน์ เนื้อหาเป็นสิ่งที่สำคัญที่สุดในเว็บไซต์ ดังนั้นควร จัดเตรียมเนื้อหาและข้อมูลที่ผู้ใช้ต้องการให้ถูกต้อง และสมบูรณ์ มีการปรับปรุงและเพิ่มเติมให้ ทันเหตุการณ์อยู่เสมอ เนื้อหาไม่ควรซ้ำกับเว็บไซต์อื่น จึงจะดึงดูดความสนใจ

5) ระบบเนวิเกชันที่ใช้งานง่าย ต้องออกแบบให้ผู้ใช้เข้าใจง่ายและใช้งาน สะดวก ใช้กราฟิกที่สื่อความหมายร่วมกับคำอธิบายที่ชัดเจน มีรูปแบบและลำดับของรายการที่ สม่ำเสมอ เช่น วางไว้ ตำแหน่งเดียวกันของทุกหน้า

6) ลักษณะที่น่าสนใจ หน้าตาของเว็บไซต์จะต้องมีความสัมพันธ์กับคุณภาพขององค์ประกอบต่าง ๆ เช่น คุณภาพของกราฟิกที่จะต้องสมบูรณ์ การใช้สี การใช้ ตัวอักษรที่อ่านง่ายสบายตา การใช้โทนสีที่เข้ากันลักษณะหน้าตาที่น่าสนใจนั้นขึ้นอยู่กับความชอบของแต่ละบุคคล

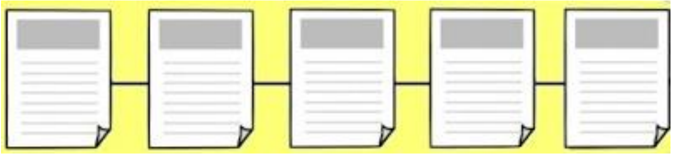
7) การใช้งานอย่างไม่จำกัด ผู้ใช้ส่วนใหญ่สามารถเข้าถึงได้มากที่สุดเลือกใช้บราวเซอร์ชนิดใดก็ได้ในการเข้าถึงเนื้อหาสามารถแสดงผลได้ทุกระบบปฏิบัติการและความละเอียดหน้าจอต่าง ๆ กันอย่างไม่มีปัญหาเป็นลักษณะสำคัญสำหรับผู้ใช้ที่มีจำนวนมาก

8) คุณภาพในการออกแบบ การออกแบบและเรียบเรียงเนื้อหาอย่างรอบคอบสร้างความรู้สึกว่าเว็บไซต์มีคุณภาพ ถูกต้อง และเชื่อถือได้

9) ลิงค์ต่าง ๆ จะต้องเชื่อมโยงไปหน้าที่มีอยู่จริงและถูกต้อง ระบบการทำงานต่าง ๆ ในเว็บไซต์จะต้องมีความแน่นอนและทำหน้าที่ได้อย่างถูก

2.2.3.2 โครงสร้างเว็บไซต์

1) เว็บที่มีโครงสร้างแบบเรียงลำดับ (Sequential Structure) ลักษณะโครงสร้างแบบนี้เป็นโครงสร้างแบบธรรมดาที่ใช้กันมากที่สุดเนื่องจากง่ายต่อการจัดระบบข้อมูล นิยมใช้กับข้อมูลที่มีลักษณะเป็นเรื่องราวตามลำดับของเวลา เช่น การเรียงลำดับตามตัวอักษร ดรรชนี สารานุกรม หรืออภิธานศัพท์ โครงสร้างแบบนี้ เหมาะกับเว็บไซต์ที่มีขนาดเล็กเนื้อหาไม่ซับซ้อนใช้การเชื่อมโยง (Link) ไปทีละหน้า ทิศทางของการเข้าสู่เนื้อหา (Navigation) ภายในเว็บจะเป็นการดำเนินเรื่องในลักษณะเส้นตรง โดยมีปุ่มเดินหน้า–ถอยหลังเป็นเครื่องมือหลักในการกำหนดทิศทาง เหมาะสำหรับเว็บไซต์ที่มีจำนวนเว็บเพจไม่มากนัก หรือเว็บไซต์ที่มีการนำเสนอข้อมูลแบบทีละขั้นตอน ข้อเสียของโครงสร้างแบบนี้คือ ผู้ใช้ไม่สามารถกำหนดทิศทางการเข้าสู่เนื้อหาของตนเองได้ ทำให้เสียเวลาในการเข้าถึงเนื้อหาที่สนใจหรือที่ต้องการได้



ภาพที่ 2.9 เว็บที่มีโครงสร้างแบบเรียงลำดับ

ที่มา : https://sites.google.com/site/pornnipaask11

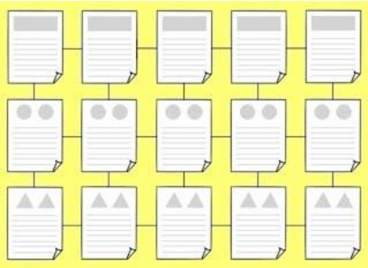
2) เว็บที่มีโครงสร้างแบบลำดับขั้น (Hierarchical Structure) ลักษณะโครงสร้างแบบนี้เป็นพื้นฐานของโครงสร้างระบบข้อมูลที่ดีที่สุด ซึ่งโดยส่วนใหญ่เว็บไซต์จะจัดอยู่รูปแบบลำดับชั้น เนื่องจากจะมีการแบ่งแยกกลุ่มของเนื้อหาอย่างชัดเจน โดยแบ่งออกเป็นส่วนต่าง ๆ และมีรายละเอียดย่อยลดหลั่นกันมา และมีความสัมพันธ์ระหว่างชั้นข้อมูลซึ่งไม่ยากเกินจะเข้าใจเพราะอยู่ในลักษณะแนวเดียวกับแผนภูมิองค์และถือเป็นลักษณะเด่นเฉพาะของเว็บประเภทนี้ที่มีจุดเริ่มต้นที่จุดเดียว นั่นคือ โฮมเพจ(Homepage) และเชื่อมโยงไปสู่เนื้อหา ในลักษณะเป็นลำดับจากบนลงล่าง ข้อดีของโครงสร้างรูปแบบนี้ คือ ง่ายต่อการแยกแยะเนื้อหาของผู้ใช้และจัดรบบข้อมูลของผู้ออกแบบเข้าถึงเนื้อหาได้อย่างรวดเร็วข้อเสียโครงสร้างรูปแบบนี้ คือ ในส่วนของการออกแบบโครงสร้างต้องระวังโครงสร้างที่ไม่สมดุล นั่นคือที่มีลักษณะที่ลึกหรือตื้นเกินไป โครงสร้างที่ลึกเกินไปลักษณะของโครงสร้างที่เนื้อหาในแต่ละส่วนมากเกินไปทำ ให้ผู้ใช้ต้องเสียเวลานานในการเข้าสู่เนื้อหาที่ต้องการ

ภาพที่ 2.10 เว็บที่มีโครงสร้างแบบลำดับขั้น

ที่มา : https://sites.google.com/site/pornnipaask11

3) เว็บที่มีโครงสร้างแบบตาราง (Grid Structure) โครงสร้างรูปแบบนี้มีความซับซ้อนมากกว่ารูปแบบที่ผ่านมา การออกแบบเพิ่มความยืดหยุ่น ให้แก่การเข้าสู่เนื้อหาของผู้ใช้ โดยเพิ่มการเชื่อมโยงซึ่งกันและกันระหว่างเนื้อหาแต่ละส่วน เหมาะแก่ การแสดงให้เห็นความสัมพันธ์กันของเนื้อหา การเข้าสู่เนื้อหาของผู้ใช้จะไม่เป็นลักษณะเชิงเส้นตรงเนื่องจากผู้ใช้สามารถเปลี่ยนทิศทางการเข้าสู่เนื้อหาของตนเองได้ในการจัดระบบโครงสร้างแบบนี้ เนื้อหาที่

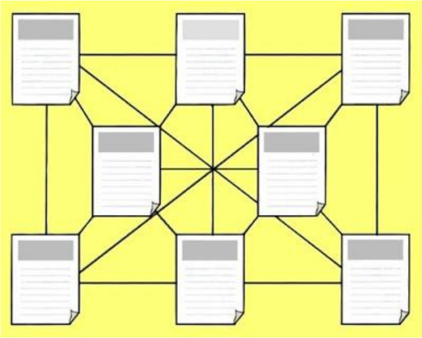
นำมาใช้แต่ละส่วนควรมีลักษณะที่เหมือนกัน และ สามารถใช้รูปแบบร่วมกัน หลักการออกแบบคือ นำหัวข้อทั้งหมดมาบรรจุลงในที่เดียวกันซึ่งโดยทั่วไป จะเป็นหน้าแผนภาพ (Map Page) ที่แสดงในลักษณะเดียวกับ โครงสร้างของเว็บ เมื่อผู้ใช้คลิกเลือก หัวข้อใด ก็จะเข้าไปสู่หน้าเนื้อหา (Topic Page) ที่แสดงรายละเอียดของหัวข้อนั้น ๆ และภายในหน้านั้น ก็จะมีการเชื่อมโยงไปยังหน้ารายละเอียดของหัวข้ออื่นที่เป็นเรื่องเดียวกัน นอกจากนี้ยังสามารถนำโครงสร้างแบบเรียงลำดับและแบบลำดับขั้นมาใช้ร่วมกันได้อีกด้วย ถึงแม้โครงสร้างแบบนี้ อาจจะสร้างความยุ่งยากในการเข้าใจได้ และอาจเกิดปัญหาการคงค้าง ของหัวข้อ (Cognitive Overhead) ได้แต่จะเป็นประโยชน์ที่สุดเมื่อผู้ใช้ได้เข้าใจถึงความสัมพันธ์ ระหว่างเนื้อหา ในส่วนของการออกแบบจำเป็นจะต้องมีการวางแผนที่ดี เนื่องจากมีการเชื่อมโยงที่เกิดขึ้น ได้หลายทิศทาง นอกจากนี้การปรับปรุงแก้ไขอาจเกิดความยุ่งยากเมื่อต้องเพิ่มเนื้อหาในภายหลัง



ภาพที่ 2.11 เว็บที่มีโครงสร้างแบบตาราง

ที่มา : https://sites.google.com/site/pornnipaask11

4) เว็บที่มีโครงสร้างแบบใยแมงมุม (Web Structure) โครงสร้างประเภทนี้จะมีความยืดหยุ่นมากที่สุด ทุกหน้าในเว็บสามารถจะเชื่อมโยงไปถึงกัน ได้หมด เป็นการสร้างรูปแบบการเข้าสู่เนื้อหาที่เป็นอิสระ ผู้ใช้สามารถกำหนดวิธีการเข้าสู่เนื้อหาได้ด้วย ตนเอง การเชื่อมโยงเนื้อหาแต่ละหน้าอาศัยการโยงใยข้อความที่มีมโนทัศน์ (Concept) เหมือนกัน ของแต่ละหน้าในลักษณะของไฮเปอร์เท็กซ์หรือไฮเปอร์มีเดีย โครงสร้างลักษณะนี้จัดเป็นรูปแบบที่ ไม่มีโครงสร้างที่แน่นนอนตายตัว (Unstructured) นอกจากนี้การเชื่อมโยงไม่ได้จำกัดเฉพาะเนื้อหาภายในเว็บนั้น ๆ แต่สามารถเชื่อมโยงออกไปสู่เนื้อหาจากเว็บภายนอก



ภาพที่ 2.12 เว็บที่มีโครงสร้างแบบใยแมงมุม
ที่มา : https://sites.google.com/site/pornnipaask11

2.2.3.3 การใช้สีในการออกแบบเว็บไซต์

การสร้างสีสันบนหน้าเว็บเป็นสิ่งที่สื่อความหมายของเว็บไซต์ได้อย่างชัดเจน การเลือกใช้สีให้เหมาะสม กลมกลืน ไม่เพียงแต่จะสร้างความพึงพอใจให้กับผู้ใช้ แต่ยังสามารถทำให้เห็นถึงความแตกต่างระหว่างเว็บไซต์ได้ สีเป็นองค์ประกอบหลักสำหรับการตกแต่งเว็บ จึงจำเป็นอย่างยิ่งที่จะต้องทำความเข้าใจเกี่ยวกับการใช้สีระบบสีที่แสดงบนจอคอมพิวเตอร์ มีระบบการแสดงผลผ่านหลอดลำแสงที่เรียกว่า CRT (Cathode ray tube) โดยมีลักษณะระบบสีแบบบวก อาศัยการผสมของของแสงสีแดง สีเขียว และสีน้ำเงิน หรือระบบสีRGB สามารถกำหนดค่าสีจาก 0 ถึง 255 ได้ จากการรวมสีของแม่สีหลักจะทำให้เกิดแสงสีขาวมีลักษณะเป็นจุดเล็ก ๆ บนหน้าจอไม่สามารถมองเห็นด้วยตาเปล่าได้ จะมองเห็นเป็นสีที่ถูกผสมเป็นเนื้อสีเดียวกันแล้ว จุดแต่ละจุดหรือพิกเซล (Pixel) เป็นส่วนประกอบของภาพบนหน้าจอคอมพิวเตอร์ โดยจำนวนบิตที่ใช้ในการกำหนดความสามารถของการแสดงสีต่าง ๆ เพื่อสร้างภาพบนจอนั้นเรียกว่า บิตเด็ป (Bit depth) ในภาษา HTML มีการกำหนดสีด้วยระบบเลขฐานสิบหก ซึ่งมีเครื่องหมาย (#) อยู่ด้านหน้าและตามด้วยเลขฐานสิบหกจำนวนอักษรอีก 6 หลัก โดยแต่ละไบต์ (byte) จะมีตัวอักษรสองตัวแบ่งออกเป็น 3 กลุ่ม เช่น #FF12AC การใช้ตัวอักษรแต่ละไบต์นี้เพื่อกำหนดระดับความเข้มของแม่สีแต่ละสีของชุดสี RGB โดย 2 หลักแรกแสดงถึงความเข้มของสีแดง 2 หลักต่อมา แสดงถึงความเข้มของสีเขียว 2 หลักสุดท้ายแสดงถึงความเข้มของสีน้ำเงิน

สีมีอิทธิพลในเรื่องของอารมณ์การสื่อความหมายที่เด่นชัด กระตุ้นการรับรู้ทางด้านจิตใจมนุษย์ สีแต่ละสีให้ความรู้สึก อารมณ์ที่ไม่เหมือนกัน สีบางสีให้ความรู้สึกสงบ บางสีให้ความรู้สึกตื่นเต้นรุนแรง สีจึงเป็นปัจจัยสำคัญอย่างยิ่งต่อการออกแบบเว็บไซต์ดังนั้นการ

เลือกใช้โทนสีภายในเว็บไซต์เป็นการแสดงถึงความแตกต่างของสีที่แสดงออกทาง อารมณ์ มีชีวิตชีวาหรือเศร้าโศก รูปแบบของสีที่สายตาของมนุษย์มองเห็น สามารถแบ่งออกเป็น 3 กลุ่ม คือ

1) สีโทนร้อน (Warm Colors) เป็นกลุ่มสีที่แสดงถึงความสุข ความปลอบโยน ความอบอุ่น และดึงดูดใจ สีกลุ่มนี้เป็นกลุ่มสีที่ช่วยให้หายจากความเฉื่อยชา มีชีวิตชีวามากยิ่งขึ้น

2) สีโทนกลาง (Neutral Colors) สีที่เป็นกลาง ประกอบด้วย สีดำ สีขาว สีเทา และสีน้ำตาล กลุ่มสีเหล่านี้คือ สีกลางที่สามารถนำไปผสมกับสีอื่น ๆ เพื่อให้เกิดสีกลางขึ้นมา

3) สีโทนเย็น (Cool Colors) แสดงถึงความที่ดูสุภาพ อ่อนโยน เรียบร้อย เป็นกลุ่มสีที่มีคนชอบมากที่สุด สามารถโน้มนาวในระยะไกลได้

2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่งภาษาของการออกแบบเว็บไซต์

2.2.4.1 CSS (ย่อมาจาก Cascading Style Sheet) มักเรียกโดยย่อว่า "สไตล์ชีท" คือภาษาที่ใช้เป็นส่วนของการจัดรูปแบบการแสดงผลเอกสาร HTML โดยที่ CSS กำหนดกฎเกณฑ์ในการระบุรูปแบบ (หรือ "Style") ของเนื้อหาในเอกสาร อันได้แก่ สีของข้อความ สีพื้นหลัง ประเภทตัวอักษร และการจัดวางข้อความ ซึ่งการกำหนดรูปแบบ หรือ Style นี้ใช้หลักการของการแยกเนื้อหาเอกสาร HTML ออกจากคำสั่งที่ใช้ในการจัดรูปแบบการแสดงผล กำหนดให้รูปแบบของการแสดงผลเอกสาร ไม่ขึ้นอยู่กับเนื้อหาของเอกสาร เพื่อให้ง่ายต่อการจัดรูปแบบการแสดงผลลัพธ์ของเอกสาร HTML โดยเฉพาะในกรณีที่มีการเปลี่ยนแปลงเนื้อหาเอกสารบ่อยครั้ง หรือต้องการควบคุมให้รูปแบบการแสดงผลเอกสาร HTML มีลักษณะของความสม่ำเสมอทั่วกันทุกหน้าเอกสารภายในเว็บไซต์เดียวกัน โดยกฎเกณฑ์ในการกำหนดรูปแบบ(Style) เอกสาร HTML ถูกเพิ่มเข้ามาครั้งแรกใน HTML 4.0 เมื่อปีพ.ศ. 2539 ในรูปแบบของCSS level 1 Recommendations ที่กำหนดโดย องค์กร World Wide Web Consortium หรือ W3C

2.2.4.2 Bootstrap คือชุดคำสั่งที่ประกอบด้วยภาษา CSS, HTML และJavaScript เป็นชุดคำสั่งที่ถูกพัฒนาขึ้นมาเพื่อกำหนดกรอบหรือรูปแบบการพัฒนาเว็บไซต์ในส่วนของการปฏิสัมพันธ์กับผู้ใช้งานเว็บไซต์ (User Interface) เราจึงสามารถเรียก Bootstrapว่าเป็น Front-end framework คือใช้สำหรับ พัฒนาเว็บไซต์ส่วนการแสดงผล ซึ่งแตกต่างจากภาษาประเภท Server Side Script อย่าง PHP, Python หรือภาษาอื่น ๆ

2.2.5 ทฤษฎีเกี่ยวกับการ visualization

Data Visualization หรือ Information Visualization คือ การนำข้อมูลในเชิงปริมาณทั้งที่จัดเก็บไว้ในรูปแบบของข้อมูลจำนวนน้อย และข้อมูลจำนวนมหาศาล (Big Data)มาประมวลผล จากนั้นจึงนำมาแสดงผลในรูปแบบของกราฟ แผนภูมิอินฟอร์เมชันกราฟิกหรือแม้กระทั่งอินเทอร์แอคทีฟกราฟิก ที่ผู้บริโภคสามารถคลิกหรือมีปฏิสัมพันธ์กับกราฟิกนั้น ๆ ได้ ซึ่งปัจจุบันสื่อหลายสำนักในประเทศไทย เริ่มมีการนำเสนอข่าวโดยการนำข้อมูลแบบประยุกต์ ด้วยภาพ (Data Visualization) การถ่ายทอดข้อมูลในเชิงปริมาณที่มีความซับซ้อนได้อย่างมี ประสิทธิผล ซึ่งคำว่า “ประสิทธิผล” ในที่นี้หมายถึงมีความชัดเจน (Clarity), มีความแม่นยำ (Precision), และมีประสิทธิภาพ (Efficiency) หากไม่มีการทำ Data Visualization แล้วอาจทำให้ เราไม่สามารถค้นพบนัยยะของข้อมูลในแง่ของแนวโน้ม, รูปแบบพฤติกรรม, และความสัมพันธ์ เชื่อมโยงได้

Visualization คือ การจินตนาการ หรือสร้างภาพขึ้นในความคิด ซึ่งเป็นกระบวนการ ทำให้เกิดความรู้ความเข้าใจสำหรับเรื่องที่จินตนาการยาก เข้าใจยาก วิธีการที่เป็นทางลัดก็คือ การสร้าง ภาพ ให้เป็น บันไดความคิด ไปสู่ การใช้ความคิดอีกระดับ บันไดนี้จะช่วยตัดปริมาณ ข้อมูล ช่วยลดภาระการคำนวณหรือการนำไปผ่านหลากหลายกระบวนการความคิด เพื่อนำไปสู่คำตอบที่ต้องการได้โดยเร็วและถูกต้อง

Visualization System คือระบบ ซึ่งประกอบด้วยกระบวนการต่าง ๆ ที่ออกแบบมาเพื่อสร้าง รักษา นำไปใช้ และปรับปรุงทัศนสนเทศ เพื่อทำให้เกิดความรู้ความเข้าใจอย่างถูกต้องรวดเร็ว และได้ผลเป็นอย่างสูง

Data Visualization เป็นการนำข้อมูลมาผสมผสานกับจินตนาการ เพื่อสร้างภาพในความคิดขึ้นมา ซึ่งมีกระบวนการนำเสนอข้อมูลที่มีความซับซ้อนหรือข้อมูลเชิงปริมาณให้ สามารถเข้าใจได้ง่าย ในแบบของ กราฟ แผนภูมิ

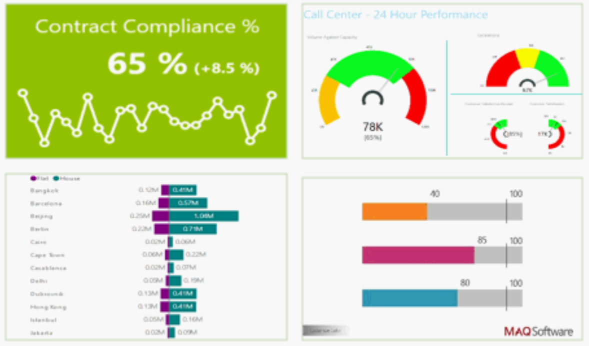
2.2.5.1 โปรแกรมสำหรับการสร้างแดชบอร์ด (Data Visualization) Tools ที่นิยมใช้ในตอนนี้ได้แก่ Tableau, Microsoft Power BI, Qlik View,Google Charts, Fusion Charts,Data wrapper และอื่น ๆ อีกมากมาย

ภาพที่ 2.13 Tools Data Visualization

ที่มา : autosoft.in.th(ม.ป.ป.)

2.2.5.2 รูปแบบในการใช้ Data Visualization

การนำเสนอแบบทิศทางหรือแนวโน้ม (Trending) เราใช้กราฟที่แสดงผลแบบทิศทางหรือแนวโน้ม เพื่อนำเสนอข้อมูลให้เห็นจำนวนข้อมูลที่เกิดขึ้นในแต่ละช่วงเวลา(period) รวมถึงเน้นข้อมูลที่ต้องการนำเสนอ เช่น Line Chart, Bar Chart, RadarChart, Area Chart เป็นต้น



ภาพที่ 2.14 การนำเสนอแบบทิศทางหรือแนวโน้ม (Trending)

ที่มา : autosoft.in.th (ม.ป.ป.)

2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.3.1 ระบบสนับสนุนการตัดสินใจ (Decision Sopport System)

ระบบสนับสนุนการตัดสินใจ (Decision Support System) เป็นระบบย่อยหนึ่งในระบบสารสนเทศเพื่อการจัดการ โดยที่ระบบสนับสนุนการตัดสินใจจะช่วยผู้บริหารในเรื่องการตัดสินใจในเหตุการณ์หรือกิจกรรมทางธุรกิจที่ไม่มีโครงสร้างแน่นอน หรือกึ่งโครงสร้าง ระบบสนับสนุนการตัดสินใจอาจจะใช้กับบุคคลเดียวหรือช่วยสนับสนุนการตัดสินใจเป็นกลุ่ม นอกจากนั้นยังมีระบบสนับสนุนผู้บริหารเพื่อช่วยผู้บริหารในการตัดสินใจเชิงกลยุทธ์

2.3.1.1 ระบบสนับสนุนการตัดสินใจ คือ DSS เป็นซอฟแวร์ที่ช่วยในการตัดสินใจเกี่ยวกับการจัดการ การรวบรวมข้อมูล การวิเคราะห์ข้อมูล และการสร้างตัวแบบที่ซับซ้อน ภายใต้ซอฟต์แวร์เดียวกัน นอกจากนั้น DSS ยังเป็นการประสานการทำงานระหว่างบุคลากรกับเทคโนโลยีทางด้านซอฟต์แวร์ โดยเป็นการกระทำโต้ตอบกัน เพื่อแก้ปัญหาแบบไม่มีโครงสร้าง และอยู่ภายใต้การควบคุมของผู้ใช้ตั้งแต่เริ่มต้นถึงสิ้นสุดขั้นตอนหรืออาจกล่าวได้ว่า DSS เป็นระบบที่โต้ตอบกันโดยใช้คอมพิวเตอร์ เพื่อหาคำตอบที่ง่าย สะดวก รวดเร็วจากปัญหาที่ไม่มีโครงสร้างที่แน่นอน ดังนั้นระบบการสนับสนุนการตัดสินใจ จึงประกอบด้วยชุดเครื่องมือ ข้อมูล ตัวแบบ (Model) และทรัพยากรอื่น ๆ ที่ผู้ใช้หรือนักวิเคราะห์นำมาใช้ในการประเมินผลและแก้ไขปัญหา ดังนั้นหลักการของ DSS จึงเป็นการให้เครื่องมือที่จำเป็นแก่ผู้บริหาร ในการวิเคราะห์ข้อมูลที่มีรูปแบบที่ซับซ้อนแต่มีวิธีการปฏิบัติที่ยืดหยุ่น DSS จึงถูกออกแบบเพื่อเพิ่มประสิทธิภาพในการทำงาน ไม่เพียงแต่การตอบสนองในเรื่องความต้องการของข้อมูลเท่านั้น

2.3.1.2 กระบวนการตัดสินใจประกอบด้วย 4 ขั้นตอน คือ

1. การใช้ความคิดประกอบเหตุผล (Intelligence) เป็นขั้นตอนที่รับรู้และตระหนักถึงปัญหาหรือโอกาสที่เกิดขึ้น ทำการรวบรวมข้อมูลที่เกี่ยวข้องกับปัญหา นำข้อมูลมาวิเคราะห์และตรวจสอบเพื่อแยกแยะและกำหนดรายละเอียดของปัญหาหรือโอกาส

2. การออกแบบ (Design) เป็นขั้นตอนของการพัฒนาและวิเคราะห์ทางเลือกในการปฏิบัติที่เป็นไปได้ รวมถึงการตรวจสอบและประเมินทางเลือกในการแก้ปัญหา ซึ่งอาจใช้ตัวแบบเพื่อสร้างทางเลือกต่าง ๆ ในการแก้ปัญหา หรือออกแบบหนทางแก้ปัญหาที่ดีที่สุด

3.การคัดเลือก (Choice) ผู้ตัดสินใจจะเลือกแนวทางเลือกที่เมาะสมกับปัญหาและสถานการณ์มากที่สุด โดยอาจใช้เครื่องมือมาช่วยวิเคราะห์ คำนวณค่าใช้จ่ายและผลตอบแทนของแต่ละแนวทางเพื่อให้เกิดความมั่นใจว่าได้เลือกแนวทางที่ดีที่สุด

4.การนำไปใช้ (Implementation) เป็นขั้นตอนที่นำผลการตัดสินใจไปปฏิบัติและคิดตามผลของการปฏิบัติเพื่อตรวจสอบว่าการดำเนินงานมีประสิทธิภาพหรือมีข้อขัดข้องประการใด จะต้องแก้ไข้หรือปรับปรุงให้สอดคล้องและเหมาะสมกับสถานการณ์อย่างไร

2.3.2 การตัดสินใจแบบโครงสร้างต้นไม้ (Decision Tree)

เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เป็นเทคนิคหนึ่งของ Classification ซึ่งเป็นวิธีการ แบ่งประเภทหรือแยกหมวดหมู่ข้อมูล โดย Classification นั้นเป็นเทคนิคหนึ่งของเหมืองข้อมูล (Data Mining)

2.3.2.1 เทคนิคต้นไม้ตัดสินใจ (Decision Tree) คือ แบบจำลองทางคณิตศาสตร์เพื่อการหาทาง เลือกที่ดีที่สุด โดยการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบของโครงสร้างต้นไม้ ซึ่ง มีการเรียนรู้ข้อมูลแบบมีผู้สอน (Supervised Learning) สามารถสร้างแบบจำลองการจัดหมวดหมู่ (Clustering) ได้จากกลุ่มตัวอย่างของข้อมูลที่กำหนดไว้ล่วงหน้า (Training set) ได้โดยอัตโนมัติ และสามารถพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ได้อีกด้วย โดยปกติมักประกอบด้วยกฎในรูปแบบ “ถ้า เงื่อนไข แล้ว ผลลัพธ์” เช่น (พยน, 2548)

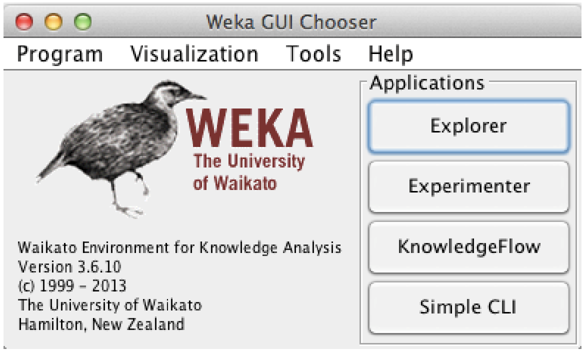
1) โหนด (Node) คือ คุณสมบัติต่าง ๆ เป็นจุดที่แยกข้อมูลว่าจะให้ไปในทิศทาง ใดซึ่งโหนดที่อยู่สูงสุดเรียกว่า โหนดราก (Root Node)

2) กิ่ง (Branch) คือ คุณสมบัติของคุณสมบัติในโหนดที่แตกออกมาโดยจำนวน ของกิ่งจะเท่ากับคุณสมบัติของโหนด

3) ใบ (Leaf) คือ กลุ่มของผลลัพธ์ในการแยกแยะข้อมูล โดยสามารถแสดงส่วนประกอบของต้นไม้ตัดสินใจ

2.3.2 ซอฟต์แวร์ Weka

Weka ย่อมาจาก Waikato Environment for Knowledge Analysis (ขจรศักดิ์ ศรีอ่อน , 2552) ซึ่ง เป็นซอฟต์แวร์สำเร็จภาพประกอบประเภทซอฟต์แวร์ให้เปล่า (Free Ware) ที่สามารถดาวน์โหลดได้ จากเว็บไซต์ ซึ่งอยู่ภายใต้การควบคุมของ GPL License ซึ่งโปรแกรม Weka ได้ถูกพัฒนามาจาก ภาษาจาวาทั้งหมด ซึ่งเขียนมาโดยเน้นกับงานทางด้านการเรียนรู้ด้วยเครื่อง (Machine Learning) และ การทำเหมืองข้อมูลโปรแกรมประกอบไปด้วยโมดูลย่อย ๆ สำหรับใช้ในการจัดการข้อมูลและ เป็นโปรแกรมที่สามารถใช้ Graphic User Interface (GUI) และใช้คำสั่งในการให้ซอฟต์แวร์ประมวลผล



ภาพที่ 2.15 ซอฟต์แวร์ Weka

ที่มา : http://preedeechanok.blogspot.com/2017/

2.3.3.1 โปรแกรมหลักของซอฟต์แวร์ Weka

2.3.3.1.1 Simple CLI (Command Line Interface) เป็นโปรแกรมรับคำสั่งการทำงานผ่านการพิมพ์

2.3.3.1.2 Explorer เป็นโปรแกรมที่ออกแบบในลักษณะ GUI

2.3.3.1.3 Experimenter เป็นโปรแกรมที่ออกแบบการทดลองและการทดสอบผล

2.3.3.1.4 KnowledgeFlow เป็นโปรแกรมออกแบบผังการไหลของความรู้

2.3.3.1.5 ArifViewer เป็นโปรแกรมที่ใช้สำหรับแก้ไขแฟ้มประเภท Aiff

2.3.3.1.6 Log เป็นโปรแกรมที่ใช้อ่านข้อความบันทึกเก็บระหว่างการทำงาน

2.3.3.2 เมนูหลักของ Explorer

2.3.3.2.1 Preprocess การเตรียมข้อมูล

2.3.3.2.2 Classify รวมโมดูลการทำเหมืองข้อมูลแบบจัดจำแนก ประเภท

2.3.3.2.3 Cluster รวมโมดูลการทำเหมืองข้อมูลแบบการเกาะกลุ่ม

2.3.3.2.4 Associate รวมโมดูลการทำเหมืองข้อมูลแบบกฎเชื่อมโยง

2.3.3.2.5 Select attributes รวมโมดูลสำหรับการวิเคราะห์ความเกี่ยวพันของลักษณะประจำ

2.3.3.2.6 Visualize นำเสนอข้อมูลด้วยภาพนามธรรมสองมิติ

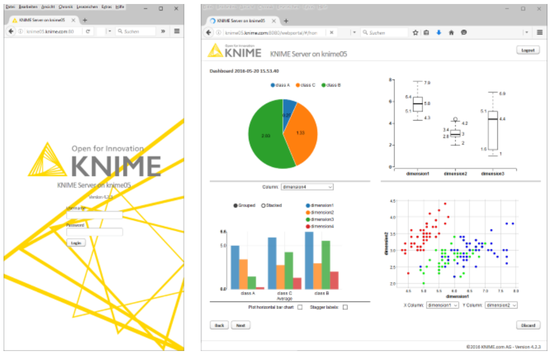
2.3.3.3 ประเภทของแฟ้มข้อมูลที่รับได้

2.3.3.3.1 แฟ้มข้อมูลที่รับต้องอยู่ในรูปแบบ ASCII อาจเป็น arft, csv, C45

2.3.3.3.2 ในกรณีแฟ้มข้อมูลอยู่ในเครือข่ายผู้ใช้สามารถเรียกใช้โดยอาศัย URL

2.3.3.3.3 หรืออาจใช้ข้อมูลที่อยู่ในฐานข้อมูลที่เชื่อมโยงผ่าน JDBC

2.3.3 ซอฟต์แวร์Konstanz Information Miner (KNIME)



ภาพที่ 2.16 ซอฟต์แวร์ KNIME

ที่มา : https://software.com.br/p/knime-server

KNIME เป็น Open Source ที่ได้รับการจัดอันดับจาก Gartner (ประกาศเมื่อเดือน Jan 2019) อยู่ในกลุ่ม Leader ด้าน Data Science & Machine Learning มาหลายปีติดกันแล้วเหตุผลสำคัญที่ทำให้ KNIME เป็น Software ที่ครบเครื่องและได้รับความนิยมจาก DataEngineer & Data Scientist ทั่วโลก มีหลายประเด็น อาทิเช่น

1) KNIME มีผู้ใช้และ community จากหลากหลายวงการ เช่น การตลาด, การผลิต, อุตสาหกรรมเภสัช เคมีภัณฑ์ ชีววิทยา และนักวิชาการต่าง ๆ

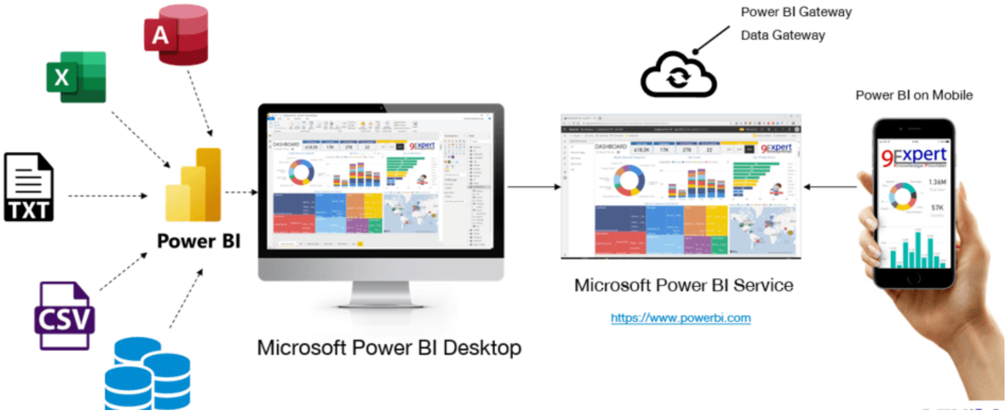
2) KNIME มีแหล่งเรียนรู้ที่เผยแพร่ไว้อย่างครบถ้วน คุณสามารถดู ตัวอย่างการประยุกต์ทำงานวิเคราะห์ข้อมูลได้หลากหลาย เช่น Customer Intelligence, SocialMedia, Finance, Manufacturing, Pharma/ Health Care, Retail, Industry & Government เป็นต้น

3) KNIME เป็น Open source ที่พร้อมใช้ มาแบบฟรี และไม่มีข้อจำกัดใด ๆ ซึ่งมันน่าสนใจและน่าใช้มาก

4) KNIME เป็น Data driven software ที่ทำได้หลายภาษา เช่น Textmining, Image processing, สามารถเชื่อมกับภาษา Python, R, Spark H2O, Keras/TensorFlowfor Deep Learning ได้อย่างลื่นไหล

การสำรวจการสร้างแบบจำลองการสร้างภาพการรายงานและการพัฒนาบางส่วนของแพลตฟอร์มเป็นโอเพ่นซอร์สเช่นเดียวกับส่วนขยายชุมชน Knime Server ซึ่งให้ความร่วมมือการทำงานอัตโนมัติการจัดการและความสามารถในการปรับใช้เป็นเชิงพาณิชย์เช่นเดียวกับส่วนขยายของคู่ค้า แพลตฟอร์ม Knime Analytics และเซิร์ฟเวอร์ Knime พร้อมใช้งานสำหรับการติดตั้งในครั้งแรกและสำหรับ AWS และ Azure clouds

2.3.4 ซอฟต์แวร์ power business intelligence (Power BI)



ภาพที่ 2.17 ซอฟต์แวร์ power business intelligence (Power BI)

ที่มา : https://www.9experttraining.com/articles/power-bi

Power BI เป็นชุดของเครื่องมือในการวิเคราะห์ข้อมูลธุรกิจ (Business Analytics Tool) และสร้างรายงานได้อย่างน่าสนใจ Power BI Dashboard ช่วยให้มุมมอง 360 องศา ให้กับผู้ใช้งานเพื่อประกอบการตัดสินใจ แบบรวมศูนย์ สามารถอัพเดต ได้อย่างทันที อีกทั้งยังสามารถดูได้จากทุก ๆ อุปกรณ์ ผู้ใช้สามารถทำ คลิกเพื่อดูข้อมูลในมุมที่ต้องการ เพื่อที่จะหาคำตอบ เพื่อตัดสินใจ การสร้างแดชบอร์ด ก็ยังสามารถเชื่อมต่อแหล่งข้อมูล (Data Source) ที่เป็นที่นิยมมากมาย และยังมีแดชบอร์ดที่สร้างมากึ่งสำเร็จรูปที่จะให้สร้างแดชบอร์ดได้อย่างรวดเร็ว และคุณยังสามารถเข้าถึงข้อมูลและรายงานของคุณได้จาก Apps Power BI Mobile ทั้งใน Windows, iOS, Android อีกด้วย ด้วยหลักการของ Design Once View Anywhere คือออกแบบครั้งเดียวแล้วดูได้จากทุก ๆ ที่ ทุก ๆ เวลา

สรุปความสามารถของ Power BI

1. สามารถนำเข้าข้อมูลจากแหล่งข้อมูล (Data Source) ที่หลากหลายได้

2. สามารถเตรียมข้อมูล โดยมีเครื่องมือที่ชื่อว่า Power Query เพื่อเตรียมข้อมูลให้พร้อมใช้งาน

3. สามารถทำ Data Model สามารถสร้าง Relationships, Data Hierachy, คำนวณด้วย DAX

4. สร้างรายงานด้วย Visualization มากมาย และยังสามารถติดตั้ง Visualization เพิ่มเติมได้จาก Appsource

5. สามารถสร้าง Dashboard และทำ Realtime Data Analytics ได้

6. สามารถเข้าถึงผ่าน Mobile Apps เพื่อดูรายงานของเราอัตโนมัติผ่าน Mobile, Tablet

7. สามารถแชร์ Report และ แชร์ Data Source ให้กับคนที่เกี่ยวข้อง

8. สามารถกำหนดให้มีการ Update ข้อมูลอัตโนมัติ (Automatic Refresh) ด้วย Power BI Gateway

2.4 วรรณกรรมที่เกี่ยวข้อง

2.4.1 พิศาลสิทธิ์ ธนวุฒิ, นิลวรรณ อยู่ภักดี (2558) ได้ทำวิทยานิพนธ์เรื่อง การคาดการณ์กำลังคนสำหรับการวางแผนอัตรากำลังเภสัชกรในประเทศไทย โดยวิธีการกำหนดอัตรากำลังคนมีหลายวิธี ขึ้นอยู่กับความเหมาะสมของการนำไปใช้ เมื่อคำนวณความต้องการได้ ไม่ควรตอบสนองเต็มจำนวนในทันที เนื่องจากอาจเกิดปัญหากำลังคนเกิน เพราะการลดกำลังการผลิตสามารถทำได้ช้า วิธีการกำหนดกำลังคนมีดังนี้ 1. วิธีอัตราส่วนต่อประชากร (population ratio method) เป็นการกำหนดตามความหนาแน่น (workforce density) เช่น แพทย์ 1 คน ต่อประชากร 10,000 คน วิธีนี้ อยู่บนสมมติฐานที่มีความคงที่และเหมือนกันทั้งหมด ทั้งบุคลากรทางการแพทย์และความจำเป็นของประชาชน ข้อดีคือ เป็นวิธีที่ง่าย ใช้ข้อมูลน้อย และนำไปใช้ได้เร็ว ข้อเสีย คือ ผลการประมาณการอาจไม่สะท้อนความเป็นจริง เนื่องจากไม่ได้คำนึงถึงปัจจัยอื่น ๆ เช่น การเพิ่มจำนวนประชากร อุปทานกำลังคนด้านสุขภาพทั้งภาครัฐและเอกชน เป็นต้น 2. วิธีเป้าหมายการบริการ (service targets method) เป็นวิธีที่องค์การอนามัยโลกยอมรับ แนวคิดเริ่มจากการตั้งเป้าหมายบริการที่สนใจจากตัวแปรที่เกี่ยวข้องและแนวโน้มในอดีต เช่น ความจำเป็นด้านสุขภาพ ความต้องการ

บริการสุขภาพ และระบบบริการสุขภาพของประเทศ โดยอนุมานว่าเป็นข้อมูลที่จะเกิดการใช้บริการ นำเป้าหมายบริการที่ได้มาเปลี่ยนเป็นปริมาณบุคลากรตามมาตรฐานตำแหน่งและบริการ ตัวอย่างเช่น หญิงตั้งครรภ์ต้องพบบุคลากรการแพทย์ทั้งหมด 4 ครั้ง และกำหนดเป้าหมายการบริการสำหรับพยาบาลต่อการบริการหญิงตั้งครรภ์ คือ 20 ครั้ง/วัน/คน จากตัวเลข 2 ค่านี้ ถ้าคาดการณ์อัตราการเกิดได้ ก็สามารถคำนวณและกำหนดจำนวนพยาบาลที่จำเป็นได้ วิธีนี้มีข้อดีกว่าวิธีอื่น คือ เข้าใจ ง่าย และแสดงความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ได้ การกำหนดอัตรากำลังคนวิธีนี้ สามารถนำประสิทธิภาพของแต่ละสถานบริการมาพิจารณา ประกอบ และยังสามารถประมาณต้นทุนการให้บริการจากการกำหนดอัตรากำลังคนได้ ข้อเสีย คือ หากตั้งสมมติฐาน ประสิทธิภาพและการกระจายของการบริการไม่ถูกต้องจะทำให้การคาดการณ์ผิดพลาด ไม่สะท้อนความเป็นจริง 3. วิธีภาระงานต่อวันร่วมกับค่ามาตรฐาน เป็นการใช้ค่ามาตรฐานที่กำหนด เช่น กฎหมายหรือข้อบังคับ มาคิดร่วมกับเวลาปฏิบัติงานในแต่ละงาน เช่น อาจารย์คณะเภสัชศาสตร์มีงานสอนในระดับปริญญาตรี 10 ชั่วโมง/สัปดาห์/ภาคการศึกษา โดยมีหลักเกณฑ์กำหนด 1 ชั่วโมงบรรยายเท่ากับ 3 หน่วยภาระ และ 1 ชั่วโมงปฏิบัติการเท่ากับ 1.5 หน่วยภาระ และมีค่าเกณฑ์มาตรฐาน 35 ชั่วโมงทำการ/สัปดาห์ ซึ่งจากงานวิจัยดังกล่าว เกี่ยวข้องกับ โครงงานของคณะผู้จัดทำในส่วนของการกำหนดอัตรากำลังคน

ซึ่งสามารถนำประสิทธิภาพของมหาวิทยาลัยมาพิจารณาหลักการทำงานของอาจารย์ เช่น อาจารย์ 1 ท่านต้องทำงานกี่ชั่วโมงต่อสัปดาห์ หรืออาจารย์ 1 ท่านสามารถสอนนักศึกษาได้ทั้งหมดกี่คนในแต่ละรายวิชา

2.4.2 ชนกานต์ กิ่งแก้ว (2558) ได้ศึกษาเรื่องการทบทวนการวิเคราะห์ข้อมูลเชิงทำนายสำหรับข้อมูลขนาดใหญ่ในธุรกิจ มีการใช้เครื่องมือในการวิเคราะห์ข้อมูล คือ Business Intelligence หรือ BI ซึ่งเป็นการนำข้อมูลที่มีอยู่มาจัดทำรายงานในรูปแบบต่าง ๆ ที่เหมาะสมกับมุมมอง ในการวิเคราะห์ให้ตรงตามความต้องการขององค์กรเพื่อประโยชน์ในการวางแผนกลยุทธด้านต่าง ๆ ซึ่งแสดงถึงระดับของเทคโนโลยี BI ต่อความซับซ้อน ของการวิเคราะห์ข้อมูล และคุณค่าที่ได้ต่อธุรกิจ ได้แก่ 1. เครื่องมือการสอบถาม การทำการรายงาน และ ค้นหา (Query, Reporting & Search Tools) เป็นรายงานที่สร้างขึ้นเฉพาะกิจ เมื่อผู้ต้องการใช้งานร้องขอ เช่น ต้องการจัดลำดับลูกค้าทั้งหมดที่ซื้อผลิตภัณฑ์ มากกว่า 5,000 ดอลลาร์ในช่วงเดือนมกราคม 2006 เป็นการสอบถามข้อมูล (Query) โดยใช้ภาษาสืบค้นข้อมูลที่อยู่ในรูปแบบฐานข้อมูลเชิงสัมพันธ์ (Relational Database) โดยใช้ภาษา SQL เพื่อนำข้อมูลออกมา แสดงในรูปแบบของตาราง 2. กระบวนการวิเคราะห์ผลแบบ

ออนไลน์ และเครื่องมือสร้างมโนทัศน์ (OLAP and Visualization tools) เป็น เทคโนโลยีที่ประกอบด้วยเครื่องมือที่ช่วยดึงและนำเสนอข้อมูลในหลายมิติ (Multidimensional) จากหลาย ๆ มุมมอง โดยที่ OLAP ได้รับการออกแบบมาสำหรับผู้ใช้ในระดับของผู้บริหารหรือหน่วยงานในองค์กรที่ต้องวิเคราะห์ข้อมูลเพื่อใช้ประกอบการตัดสินใจในระดับสูง 3. หน้าต่างแสดงข้อมูล (Dashboard and Score cards) เป็นการแสดงมโนทัศน์ (Visualization) ในด้านที่สำคัญ และจะมีการกำหนดตัวชี้วัด (KPI) ไว้ด้วย เพื่อแสดงถึงประสิทธิภาพขององค์กรจะมีการแบ่งแถบสี แสดงสถานะต่าง ๆ เพื่อสามารถติดตาม ประมวลผลได้และตัดสินใจได้ เช่น การวัดยอดขาย มีการกำหนดระดับเพื่อเอาไว้ชี้วัดมีการแบ่งตามมิติภูมิภาค เพื่อช่วยให้ผู้บริหารได้ทำการตัดสินใจได้ดีขึ้นจากข้อมูลที่มีอยู่ 4. การวิเคราะห์เชิงทำนาย (Predictive Analytics) เป็นการใช้โมเดล ขั้นสูงเพื่อสามารถวิเคราะห์ไปข้างหน้าได้จากข้อมูล ในอดีตที่มีขนาดใหญ่ (Big Data) ข้อมูลเหล่านี้จะถูกนำเข้ามาผ่านกระบวนการต่าง ๆ การทำให้ข้อมูลแสดงออกมาเป็นรายงานและการวิเคราะห์ (Report/Analysis) จะเป็นการหาคำตอบที่เกิดขึ้นในอดีต เช่น เกิดอะไรขึ้น และทำไมถึงเกิดเช่นนั้น การติดตามผล (Monitoring) เป็นการติดตามผลที่เกิดขึ้นในปัจจุบัน จากนั้นขั้นตอนการใช้การวิเคราะห์ข้อมูลเชิงทำนาย (Predictive Analytics) นั้นจะสามารถตอบคำถามที่ว่าอะไรจะเกิดขึ้นในอนาคต และอะไรที่น่าจะเกิดขึ้นมากที่สุด การวิเคราะห์จะได้สิ่งที่เกิดขึ้นในอนาคต ซึ่งจะก่อให้เกิด การตัดสินใจ (Action) เพื่อประโยชน์สูงสุดได้ ภาพรวมของการวิเคราะห์ข้อมูลเชิงทำนาย เริ่มจากการใช้ข้อมูลที่มีอยู่มาทำการวิเคราะห์ โดยข้อมูลสามารถแปลงไปเป็นการแสดงรายงาน จากเทคโนโลยีต่าง ๆ ที่กล่าวมาในข้อ 1-3 นั้น ล้วนแล้วแต่เป็นการนำเสนอข้อมูลในอดีตเท่านั้นเป็นการคาดการณ์ที่ได้จาก โมเดลขั้นสูง ซึ่งได้มาจากเทคนิคหลายเทคนิค ได้แก่ การทำเหมือนข้อมูล หลักการทางสถิติ และการเรียนรู้ของจากที่กล่าวมานั้นจะเห็นว่าเทคโนโลยีต่าง ๆ ได้มี วิวัฒนาการมาจากการเก็บรวบรวมข้อมูลต่าง ๆ ในรูปแบบ ดิจิตัลเป็นครั้งแรกในช่วงปี 1960 และพัฒนาเรื่อยมาจนถึงปัจจุบันที่ได้ทำการวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data) คือ ข้อมูลที่มีขนาดใหญ่ มีรูปแบบไม่ตายตัว ยากต่อการประมวลผล

ซึ่งจากบทความดังกล่าว เกี่ยวข้องกับ โครงงานของคณะผู้จัดทำในส่วนของการนำข้อมูลที่มีอยู่มาจัดทำรายงานในรูปแบบต่าง ๆ ให้เหมาะสมกับมุมมอง และการวิเคราะห์ข้อมูลให้ตรงตามความต้องการ เพื่อให้บรรลุเป้าหมายหรือวัตถุประสงค์ที่ได้ตั้งไว้

2.4.3 บุหลัน ทองกลีบ (2559) ได้ทำการวิจัยเรื่อง การวิเคราะห์อัตรากำลังของบุคลากรทางการพยาบาลตามกิจกรรมการพยาบาลและปริมาณเวลาในการดูแลผู้ป่วย โดยได้ทำการการ

เก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูลหาค่าความถี่ร้อยละค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานและคำนวณหาอัตรากำลังโดยใช้สูตรดังนี้

อัตรากำลังบุคลากรทางการพยาบาลที่ต้องการ (Productive FTE)

$$= \frac{\text{จำนวนผู้ป่วยเฉลี่ยต่อวัน} \times \text{ชั่วโมงการพยาบาลเฉลี่ยต่อวัน} \times 1.4 \times \text{FTE ทดแทนชั่วโมงที่ไม่ได้งาน}}{\text{จำนวนชั่วโมงการทำงานของบุคลากรทางการพยาบาล 1 คนใน 1 วัน}}$$

$$\textbf{จำนวนผู้ป่วยเฉลี่ยต่อวัน} = \frac{\text{จำนวนผู้ป่วยทุกคนใน 1 ปี}}{\text{365 วัน}}$$

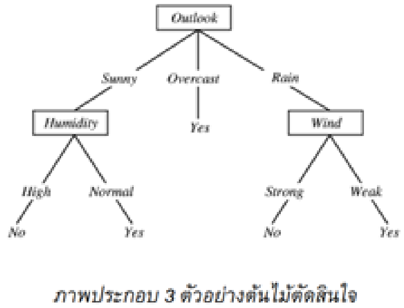
ชั่วโมงการพยาบาลเฉลี่ยต่อวัน

$$= \frac{\text{ผลรวมชั่วโมงความต้องการทางการพยาบาลของผู้ป่วยแต่ละประเภท}}{\text{จำนวนผู้ป่วยทุกคน}}$$

ค่าคงที่ 1.4 เป็นค่าคงที่อัตรากำลังที่ทดแทนการทำงานในวันหยุดแต่ละสัปดาห์ซึ่งกำหนดให้บุคลากรทางการพยาบาลแต่ละคนทำงานสัปดาห์ละ 5 วันแต่ความต้องการการดูแลของผู้ป่วยนั้นต้องการการดูแลสัปดาห์ละ 7 วันค่า 1.4 จึงเกิดจากสัดส่วนของจำนวนวันที่ผู้ป่วยต้องการใน 1 สัปดาห์คือ 7 วัน/5 วัน = 1.4 FTE (Full time equivalent employee) คือหลักการที่นักวางแผนทรัพยากรมนุษย์ใช้ในการคิดการทำงานของเจ้าหน้าที่หรือบุคลากรที่ทำงานบางเวลา (Part time staff) เทียบกับการทำงานของเจ้าหน้าที่หรือบุคลากรที่ทำงานเต็มเวลา(Full time staff) ทั้งนี้เพื่อใช้เป็นหน่วยนับอัตรากำลังตามชั่วโมงที่หน่วยงานกำหนดให้เจ้าหน้าที่ประจำทำงานซึ่งโดยทั่วไปภาครัฐมักกำหนดให้เจ้าหน้าที่ประจำทำงาน 8 ชั่วโมง/วัน สัปดาห์ละ 5 วันจึงเท่ากับ 40 ชั่วโมง/สัปดาห์หรือ2,080 ชั่วโมง/ปี (52 สัปดาห์/ปี x 40 ชั่วโมง/สัปดาห์)

ซึ่งจากงานวิจัยดังกล่าว เกี่ยวข้องกับ โครงงานของคณะผู้จัดทำในส่วนของการคำนวณเพื่อหาค่าเกณฑ์อาจารย์ต่อนักศึกษา เกณฑ์การแข่งขันระหว่างจำนวนนักศึกษาที่สมัครและจำนวนที่รับ

2.4.4 โชฑศ์รัตต ธรรมบุษดี, เอกบุรุษ นิรัติศัย (2558) ได้ทำการวิจัยเรื่อง แบบจำลองการตัดสินใจในการเลือกศึกษาต่อระดับบัณฑิตศึกษา จากปัจจัยด้านทัศนคติในการใช้สื่อสังคมออนไลน์ โดยภายในงานวิจัยจะนำเสนอแบบจำลองการตัดสินใจในรูปแบบต้นไม้ตัดสินใจ (Decision Tree) โดย Peter Cabena et al. (1998) อธิบายว่าการสร้างต้นไม้ตัดสินใจทำได้โดยใช้กระบวนการการให้เหตุผลเชิงอุปนัยจากระดับบน (Top–down induction) โดยกฎที่ถูกเลือกเป็นกฎแรกจะเป็นปัจจัยอันดับบนสุด และจะจำกัดจำนวนข้อมูลของแต่ละเงื่อนไขของปัจจัยนั้น ๆ และจากนั้นจะมีการวนซ้ำในข้อมูลที่อยู่ในแต่ละเงื่อนไขจนกว่าจะแบ่งแยกสมาชิกจนครบทุกข้อมูล ตัวอย่างของต้นไม้ตัดสินใจแสดงดังภาพประกอบที่ 3



ภาพที่ 2.18 ตัวอย่างต้นไม้ตัดสินใจ

ที่มา : https://mahara.org/artefact/file/download.php?file=162194&view=45421.

โดยภาพประกอบ 3 แสดงต้นไม้ตัดสินใจในการตัดสินใจเล่นกีฬากอล์ฟโดยดูปัจจัยจากสภาพอากาศ ความชื้น และความแรงของลมตัวอย่างการให้เหตุผลเช่น ถ้าสภาพอากาศแจ่มใส (Outlook มีค่าเป็น Sunny) ปัจจัยต้องพิจารณาปัจจัยความชื้นว่าปกติหรือสูง ถ้าความชื้นสูงก็จะไม่เล่นกอล์ฟแต่ถ้าความชื้นปกติก็จะตัดสินใจเล่นกอล์ฟ เป็นต้น จะเห็นว่าทุกการตัดสินใจจากต้นไม้ตัดสินใจจะสามารถนำมาสร้างเป็นกฎในการตัดสินใจได้

ซึ่งจากงานวิจัยดังกล่าว เกี่ยวข้องกับ โครงงานของคณะผู้จัดทำในส่วนของการนำเสนอเพื่อความสะดวกในแบบจำลองการตัดสินใจของงานโครงงานนี้ ใช้กระบวนการการให้เหตุผลและประกอบการตัดสินใจ

2.4.5 เบญจมาศ ปิยะ (2559) ได้ทำการวิจัยเรื่อง การศึกษาพฤติกรรมผู้ใช้งานตู้แช่เชิงพาณิชย์ด้วยเทคนิคต้นไม้ตัดสินใจ โดยใช้สูตรคำนวณเทคนิคต้นไม้ตัดสินใจเพื่อใช้ในการคำนวณหาค่ากลาง และนำเอาค่ากลางไปตั้งเป็นค่ามาตรฐานในการตั้งเวลาการเปิดปิดประตูตู้แช่ สูตรที่ 1 ถ้าเลือกคุณสมบัติ (Attribute) X เป็น Node โดยที่คุณสมบัติ X มีค่า (Value) ที่เป็นไปได้

ทั้งหมด n ค่า ซึ่ง node ที่สร้างจากคุณสมบัติ X แบ่งข้อมูลจำนวน S ข้อมูล ออกตามกิ่ง (Link) โดยมีจำนวนข้อมูลในแต่ละกิ่งเป็น {s1, s2, ...sn} ค่าเอนโทรปีหลังจากแบ่งข้อมูลตามคุณสมบัติ X สามารถคำนวณ

$$\text{Ex(S)} = \sum_{i=l}^{n} \left|\frac{Si}{s}\right| E(si)$$

สูตรที่ 2 ค่าเกนของคุณสมบัติ (Attribute) X สามารถคำนวณได้จากการ ลบค่าเอนโทรปีทั้งหมดของชุดข้อมูล S กับค่าเอนโทรปีที่ได้หลังจากแบ่งข้อมูลด้วยคุณสมบัติ X สามารถคำนวณ

$$\text{Gain(X)} = \text{E(S)} - \text{Ex(s)}$$

กระบวนการสร้างต้นไม้ตัดสินใจโดยทั่วไปแล้วจะมีขั้นตอน โดยสรุปดังต่อไปนี้ 1. เริ่มต้นสร้างโหนดขึ้นมาหนึ่งโหนดจากชุดข้อมูล 2. ถ้าข้อมูลทั้งหมดอยู่ในกลุ่มเดียวกันแล้ว ให้โหนดที่สร้างขึ้นเป็นโหนดใบและตั้งชื่อแยกตามกลุ่มของข้อมูลนั้น 3. ถ้าข้อมูลไม่มีคุณลักษณะใดที่เหมาะสมในการแบ่งกลุ่ม ให้โหนดที่สร้างขึ้นเป็นโหนดใบและตั้งชื่อตามกลุ่มที่มีข้อมูลสนับสนุนมากที่สุด 4. ถ้าข้อมูลภายในโหนดมีหลากหลายกลุ่มปะปนกันให้ทำการเลือกคุณลักษณะที่มีความเหมาะสมที่สุด โดยพิจารณาจากอินฟอร์เมชันเกน ซึ่งเป็นตัวชี้วัดความสามารถในการจำแนกกลุ่ม 5. เมื่อได้ตัวทดสอบการตัดสินใจ ให้สร้างกิ่งของต้นไม้ด้วยค่าต่าง ๆ ที่เป็นไปได้ของตัวทดสอบและแบ่งข้อมูลออกตามกิ่งต่าง ๆ ที่สร้างขึ้น 6. พิจารณาข้อมูลในแต่ละกิ่ง ถ้าพบว่าข้อมูลทั้งหมดอยู่ในกลุ่มเดียวกัน ให้ต่อกิ่งด้วยโหนดใบ และกำหนดค่าด้วยกลุ่มของข้อมูลนั้น แต่ถ้าพบว่าข้อมูลมีหลากหลายกลุ่มปะปนกัน ให้ทำการวนซ้ำการหาตัวทดสอบการตัดสินใจที่เหมาะสมต่อไป 7. ทำการวนซ้ำแบ่งข้อมูลและแตกกิ่งของต้นไม้ไปเรื่อย ๆ โดยการวนซ้ำจะสิ้นสุดก็ต่อเมื่อเงื่อนไขข้อใดข้อหนึ่งต่อไปนี้เป็นจริง

ซึ่งจากงานวิจัยดังกล่าว เกี่ยวข้องกับ โครงงานของผู้จัดทำในส่วนของสูตรในการคำนวณเพื่อสร้างต้นไม้ตัดสินใจ และขั้นตอนหรือกระบวนการในการสร้างต้นไม้ตัดสินใจ

2.5 บทสรุป

จากการที่ได้ศึกษาวรรณกรรมทั้งหมดที่กล่าวข้างต้น คณะผู้จัดทำได้ความรู้และประโยชน์ในเรื่องของวิธีการคาดการณ์กำลังคน เช่น วิธีอัตราส่วนบุคลากรต่อประชากร วิธีเป้าหมายบริการวิธีภาระงานต่อวันกับค่ามาตรฐาน ซึ่งเป็นส่วนสำคัญของการคาดการณ์กำลังคน การเลือกใช้ขึ้นอยู่กับข้อมูลที่มีอยู่หรือที่สามารถรวบรวมได้และวัตถุประสงค์ของการนำผลไปใช้ประโยชน์ ทั้งนี้เนื่องจากอนาคตเป็นสิ่งที่คาดการณ์ได้ยาก และสถานการณ์มีการเปลี่ยนแปลงตลอดเวลา ดังนั้น

เพื่อให้ได้ข้อมูลที่ดีจึงควรทำการศึกษากำลังคนอย่างต่อเนื่อง โดยในการวิเคราะห์นั้นต้องมีการเก็บรวบรวมข้อมูล หาค่าความถี่ ร้อยละ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานและคำนวณหาอัตรากำลังโดยใช้สูตรต่าง ๆ และการใช้ต้นไม้ตัดสินใจ (Decision Tree) นั้นมีส่วนช่วยให้การวิเคราะห์ข้อมูลที่มีจำนวนมาก ๆ ได้โดยจำลองความเป็นไปได้ของสิ่งที่ต้องการทราบและเงื่อนไขทั้งหมดที่ได้จากข้อมูลแต่ละตัวทำให้มองเห็นภาพรวมและวิธีการที่ต้องการได้ดียิ่งขึ้น และต้นไม้การตัดสินใจยังช่วยในการตัดสินใจทำสิ่งต่าง ๆ โดยใช้วิธีการที่ดีที่สุดได้ง่ายขึ้นกว่าการไม่ได้ใช้ต้นไม้การตัดสินใจ และในปัจจุบันการใช้เครื่องมือในการวิเคราะห์ข้อมูล คือ ความชาญฉลาดทางธุรกิจ (Business Intelligence) ซึ่งเป็นการนำข้อมูลที่มีอยู่มาจัดทำรายงานในรูปแบบต่าง ๆ ที่เหมาะสมกับมุมมอง ในการวิเคราะห์ให้ตรงตามความต้องการขององค์กรเพื่อประโยชน์ในการวางแผนกลยุทธด้านต่าง ๆ ซึ่งทำให้ทางคณะผู้จัดทำนั้นตัดสินใจได้ว่าควรที่จะทำการวิเคราะห์ปัญหาไปในทิศทางไหนได้บ้างและควรวิเคราะห์ถึงส่วนใดบ้าง เพื่อความถูกต้องและแม่นยำตามที่ได้ทำการคาดการณ์ในการวิเคราะห์ไว้