

บทที่ 2

แนวคิด ทฤษฎีเอกสารที่เกี่ยวข้อง

จากการศึกษาเอกสารงานการการพัฒนาแบบจำลองมาตรการการช่วยเหลือด้านการศึกษา สำหรับนักศึกษา ด้วยเทคนิคต้นไม้ตัดสินใจ ผู้ศึกษาพบว่าเอกสารแนวความคิดและโครงการต่าง ๆ ที่เกี่ยวข้องกับการศึกษาสรุปได้ ดังนี้

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytic)

2.1.2 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการจัดการข้อมูลขนาดใหญ่

2.2.2 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

2.2.3 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่ง HTML

2.2.5 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS

2.2.6 ทฤษฎีเกี่ยวกับชุดคำสั่ง PHP

2.2.7 ทฤษฎีเกี่ยวกับการ Visualization

2.2.8 ทฤษฎีเกี่ยวกับการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ

2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.3.1 โปรแกรม Rapid Miner

2.3.2 โปรแกรม Weka

2.3.3 โปรแกรม Atom

2.3.4 โปรแกรม Adobe XD

2.4 วรรณกรรมที่เกี่ยวข้อง

2.5 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytic)

Data analytics เป็นศาสตร์ของการวิเคราะห์ข้อมูลต่าง ๆ จาก Big data เพื่อช่วยในด้านธุรกิจ หรือตามวัตถุประสงค์อื่น ๆ ที่ต้องการ ซึ่งเริ่มต้นจากความสามารถของเราในการนำข้อมูลเหล่านั้น มาให้อยู่ในรูปแบบที่พร้อมจะประมวลผลได้ โดยเทคโนโลยีหรือชุดคำสั่งและแบบจำลองที่สร้างขึ้น เพื่อนำข้อมูลที่ได้ผ่านการวิเคราะห์แล้วมาใช้ หรือแปลความหมายโดยบุคคลที่ได้รับการฝึกอบรมการใช้เทคโนโลยีเหล่านั้น รูปแบบของการวิเคราะห์ข้อมูลสามารถแบ่งได้ดังนี้

1) การวิเคราะห์ข้อมูลแบบพื้นฐาน (Descriptive analytics) เป็นการวิเคราะห์เพื่อแสดงผลของรายการทางธุรกิจ เหตุการณ์ หรือกิจกรรมต่าง ๆ ที่ได้เกิดขึ้น หรืออาจกำลังเกิดขึ้นในลักษณะที่ง่ายต่อการเข้าใจ หรือต่อการตัดสินใจ ตัวอย่างเช่น รายงานการขาย รายงานผลการดำเนินงาน

2) การวิเคราะห์แบบเชิงวินิจฉัย (Diagnostic analytics) เป็นการอธิบายถึงสาเหตุของสิ่งที่เกิดขึ้นปัจจัยต่าง ๆ และความสัมพันธ์ของปัจจัยหรือตัวแปรต่าง ๆ ที่มีความสัมพันธ์ต่อกันของสิ่งที่เกิดขึ้น ตัวอย่างเช่น ความสัมพันธ์ระหว่างยอดขายต่อกิจกรรมทางการตลาดแต่ละประเภท ซึ่งเป็นก้าวใหม่ที่ช่วยเสริมให้ตัดสินใจไปในทางที่ถูกต้อง

3) การวิเคราะห์แบบพยากรณ์ (Predictive analytics) เป็นการวิเคราะห์เพื่อพยากรณ์สิ่งที่กำลังจะเกิดขึ้นหรือน่าจะเกิดขึ้น โดยใช้ข้อมูลที่ได้เกิดขึ้นแล้วกับแบบจำลองทางสถิติ หรือเทคโนโลยีปัญญาประดิษฐ์ต่าง ๆ (Artificial intelligence) ตัวอย่างเช่น การพยากรณ์ยอดขาย การพยากรณ์ผลประชามติ

4) การวิเคราะห์แบบให้คำแนะนำ (Prescriptive analytics) เป็นการวิเคราะห์ข้อมูลที่มีความซับซ้อนที่สุด เป็นทั้งการพยากรณ์สิ่งต่าง ๆ ที่จะเกิดขึ้น ข้อดี ข้อเสีย สาเหตุ และระยะเวลาของสิ่งที่เกิดขึ้น และการให้คำแนะนำ ทางเลือกต่าง ๆ ที่มีอยู่ และผลของแต่ละทางเลือก

2.1.2 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

การนำข้อมูลดิบมาเปลี่ยนเป็นกราฟ แผนภูมิ หรือแม้กระทั่งวิดีโอที่ช่วยอธิบายปริมาณตัวเลข และช่วยให้คุณได้ข้อมูลเชิงลึกจากข้อมูลเหล่านั้น การใช้ Data Visualization เปลี่ยนแนวคิดของเราเกี่ยวกับการสร้างคุณค่าให้กับข้อมูล ทำให้เราค้นพบรูปแบบใหม่ ๆ และมองเห็นแนวโน้มของข้อมูลมากขึ้น

หากต้องการประสบความสำเร็จในโลกธุรกิจที่มีการแข่งขันสูง ความสามารถในการตัดสินใจที่ถูกต้องในเวลาที่เหมาะสม โดยอาศัยข้อมูลเชิงลึกที่รัดกุมเป็นสิ่งสำคัญ ปัจจุบันองค์กรธุรกิจส่วนใหญ่ในประเทศไทย เริ่มใช้ประโยชน์จากการทำ Data Visualization เพราะว่าเครื่องมือนี้เป็นกุญแจสำคัญในการแสดงผลของข้อมูลให้เข้าใจง่าย ซึ่งจะช่วยให้ทุกแผนกในฝ่ายงานได้ใช้ประโยชน์ ในการหาแนวโน้มและความสัมพันธ์เชื่อมโยง อีกทั้งทำให้ทราบข้อมูลได้เร็วขึ้น ส่งผลให้การทำงานขององค์กรมีประสิทธิภาพเพิ่มขึ้น จนดำเนินการถึงการวางกลยุทธ์ภายในองค์กร และสามารถนำไปใช้กับการทำเว็บไซต์ได้ Data Visualization ช่วยให้การสื่อสารข้อมูล มีความรวดเร็วและเข้าใจง่ายมากขึ้น

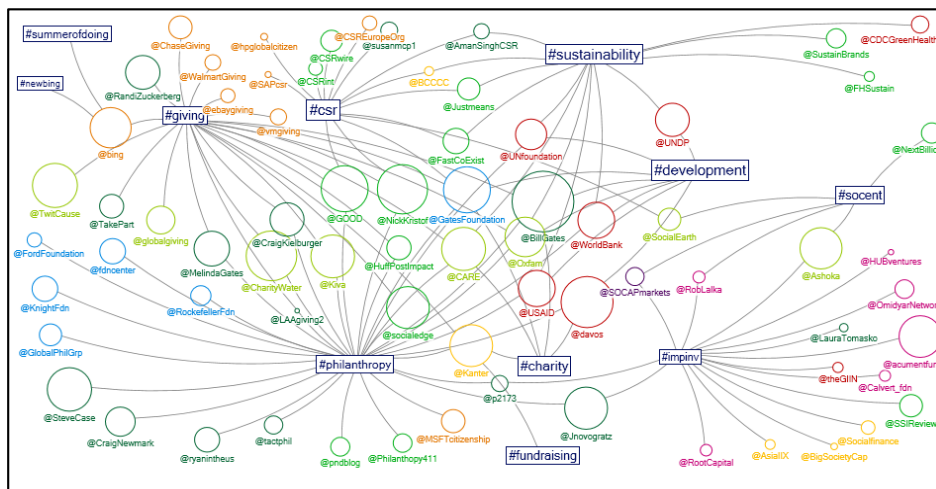


ภาพที่ 2.1 แสดง Data Visualization

พื้นฐานของการสร้าง Data Visualization คือ การ Mapping ส่วนข้อมูลกับส่วนของ Graphic เข้าด้วยกัน ซึ่งตอนนี้มีโปรแกรมสำเร็จรูปในการสร้าง Data Visualization หลากหลายโปรแกรม มีฟังก์ชันการใช้งานที่เข้าใจง่าย เช่น การสร้างฟิลเตอร์ การออกแบบเพื่อให้งานการวิเคราะห์ข้อมูล มีความยืดหยุ่นเป็นต้น ตัวอย่างรูปแบบ Data Visualization ที่แสดงผลให้อย่างสนใจ มีดังนี้

1) NETWORK/FLOW

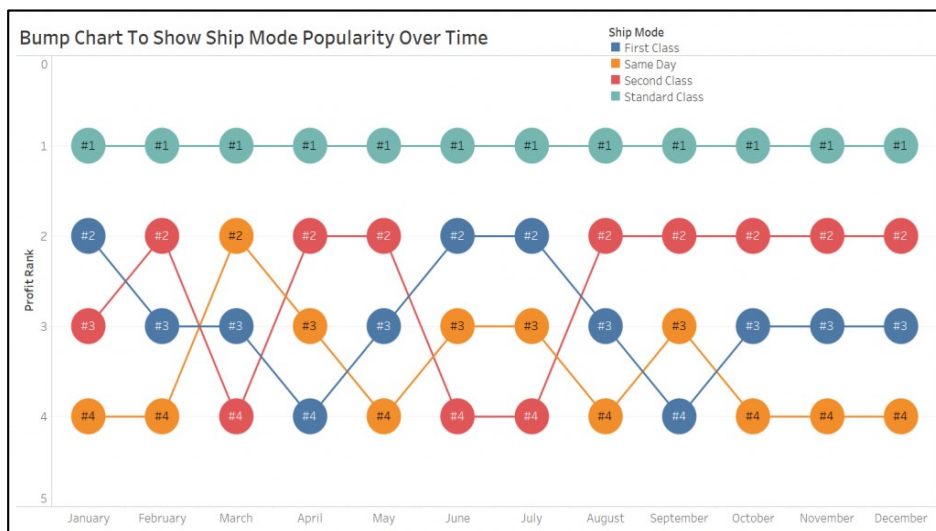
Network Graph ใช้แสดงความเชื่อมโยงของเครือข่ายหรือความสัมพันธ์ในกลุ่ม



ภาพที่ 2.2 แสดง Network Graph

2) RANKING

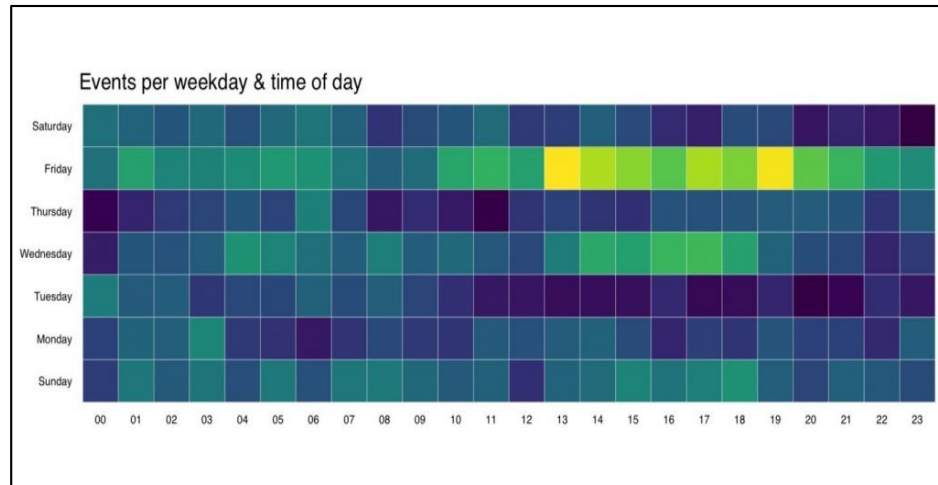
Bump Chart ใช้แสดงการเปลี่ยนแปลงลำดับของข้อมูลในหลายช่วงเวลา โดยใช้สีเพื่อช่วยให้อ่านข้อมูลได้ง่ายขึ้น



ภาพที่ 2.3 แสดง Bump Chart

3) RELATIONSHIP

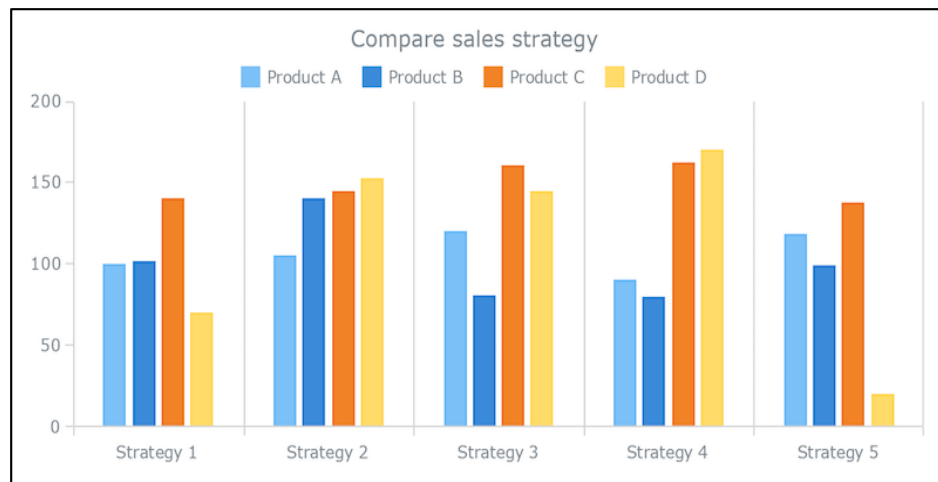
Heat map ใช้แสดงรูปแบบความสัมพันธ์ของข้อมูล โดยจะแสดงออกมาในรูปแบบของ “ สี ” ซึ่งแต่ละสีจะบ่งบอกถึงระดับความถี่ของพฤติกรรม แต่ใช้อ่านค่าความแตกต่างเล็กน้อยได้ยาก



ภาพที่ 2.4 แสดง Heat map

4) COMPARISON

Stacked Bar Chart ใช้เปรียบเทียบค่าผลรวมและสัดส่วนจากข้อมูลหลายกลุ่ม อาจอ่านสัดส่วนยากเมื่อมีข้อมูลหลายกลุ่มมากเกินไป

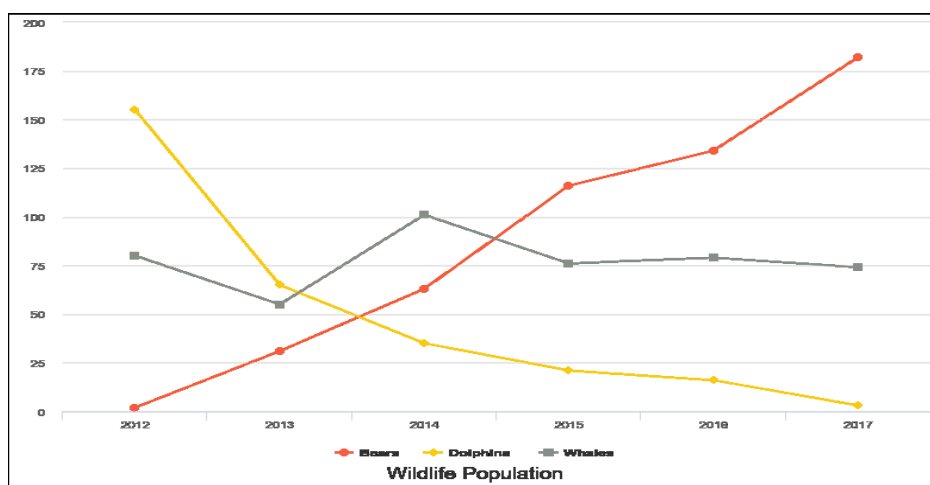


ภาพที่ 2.5 แสดง Stacked Bar Chart

5) TIME-SERIES

Line Chart ใช้เปรียบเทียบข้อมูลเพื่อดูแนวโน้ม (Trends) โดยอาจจะเทียบกับมิติของเวลา

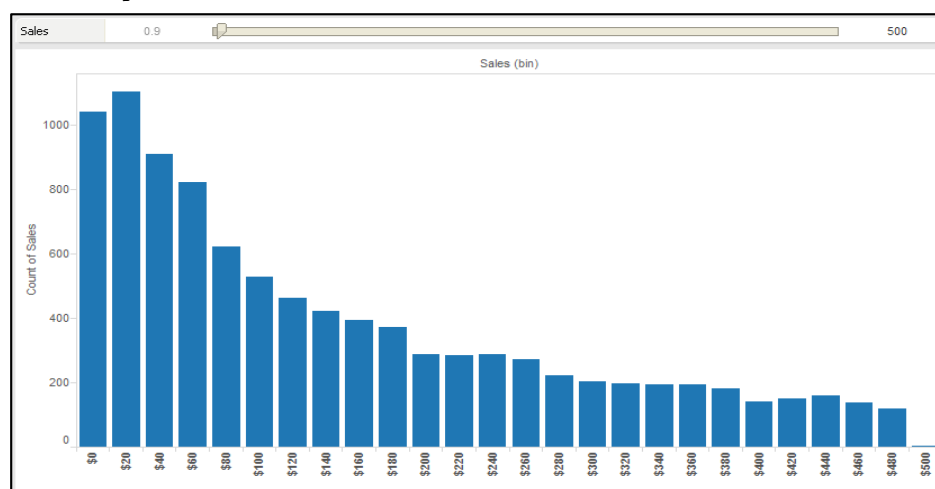
Time series



ภาพที่ 2.6 แสดง Stacked Bar Chart

6) DISTRIBUTION

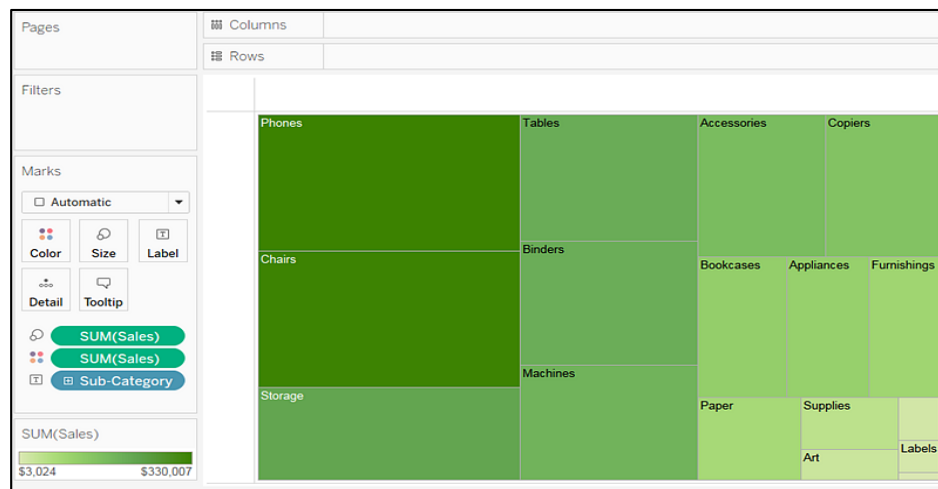
Histogram กราฟแท่งแบบเฉพาะที่แสดงความสัมพันธ์ระหว่างข้อมูลเป็นหมวดหมู่ที่เรียกว่า ชั้นข้อมูลกับความถี่ของข้อมูล เพื่อดูการกระจายของข้อมูล ลดระยะห่างระหว่างแท่งเพื่อให้เห็น “รูปทรง” ของข้อมูลชัดเจน



ภาพที่ 2.7 แสดง Histogram

7) COMPOSITION

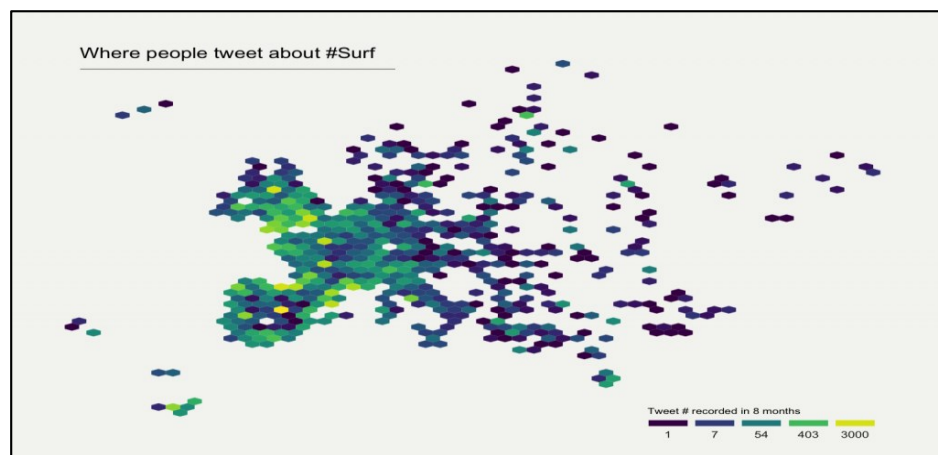
Tree map ใช้แสดงสัดส่วนของข้อมูลแบบมีลำดับชั้น (hierarchy) แสดงเป็นรูปสี่เหลี่ยม โดยจะแสดงทั้งในแง่ของสีเดียวกัน จัดเป็นกลุ่มเดียวกัน อาจอ่านสัดส่วนยากเมื่อมีกลุ่มขนาดเล็กจำนวนมาก



ภาพที่ 2.8 แสดง Tree map

8) SPATIAL

Hexbin Map เป็นการแสดงผลข้อมูลด้วยรูปทรงหกเหลี่ยมที่มีขนาดเท่ากัน โดยกำหนดสีแทนระดับความถี่ของจำนวนข้อมูล ส่วนมากนำไปใช้ลงบนตำแหน่งทางภูมิศาสตร์ เพื่อให้เห็นภาพได้ชัดเจน



ภาพที่ 2.9 แสดง Hexbin Map

จากที่ได้กล่าวไปจะเห็นได้ว่าข้อมูลแบบ Data Visualization มีพลังทางการตลาดอย่างมาก เพราะสามารถแสดงผลและเพิ่มความสนใจได้เป็นอย่างดี นอกจากนี้ยังทำให้ผู้ได้รับข้อมูลสามารถจินตนาการตามข้อมูลที่แสดงได้อีกด้วย หากองค์กรได้ลองนำการแสดงผลข้อมูลแบบ Data Visualization มาใช้ก็จะช่วยในการจัดการระดับสายงานแผนกอื่น ๆ ในบริษัท ส่งผลให้การวิเคราะห์ข้อมูลและตัดสินใจในสายงานต่าง ๆ มีประสิทธิภาพ ซึ่งดีต่อธุรกิจและทีม หรือนำมาใช้ในหน้าแรกของเว็บไซต์ เพื่อทำการโปรโมทโปรโมชันหรือข้อเสนอทางการตลาดของธุรกิจ แล้วจะพบว่า Data Visualization ช่วยให้เกิดการตอบสนองกับข้อความมากขึ้น

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการจัดการข้อมูลขนาดใหญ่ (Big Data)

2.2.1.1 ข้อมูลขนาดใหญ่ (Big Data)

ข้อมูลขนาดใหญ่ (Big Data) ซึ่งสามารถนำไปประยุกต์ใช้งานให้เป็นประโยชน์ต่อองค์กรต่าง ๆ สำหรับภาคเอกชนทั่วไปก็ต้องนำมาใช้ให้เกิดประโยชน์ต่อองค์กรมากที่สุด วัตถุประสงค์การใช้งานเพื่อองค์กรของตนเป็นหลัก แต่นอกจาก Big Data จะมีความสำคัญกับภาคเอกชนแล้วสำหรับภาครัฐนั้นก็มีความสำคัญอย่างยิ่งเช่นกัน เพราะรัฐจำเป็นต้องใช้ข้อมูลของ Big Data มากก่อให้เกิดประโยชน์ต่อประชาชนทุกคน เพื่อจะกระจาย ความช่วยเหลือไปถึงทุกที่ พร้อมทั้งสามารถนำข้อมูลที่มีการวิเคราะห์ไปใช้ประโยชน์ ในการบริหาร วางแผนกำหนดทิศทางในด้านต่าง ๆ อย่างมีประสิทธิภาพที่สูงขึ้น ข้อมูลที่มีขนาดใหญ่ ซึ่งอาจจะแบ่งย่อยออกเป็นข้อมูล 3 รูปแบบคือ

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

2) ข้อมูลกึ่งโครงสร้าง (semi-structured data) คือข้อมูลที่ถูกจัดเก็บอย่างมีรูปแบบในระดับหนึ่งและข้อมูลที่สามารถค้นหา (search) หรือแท็ก (tag) ได้ เช่น เว็บเพจที่มีการระบุชื่อเพจ คำสำคัญในเพจ และวันที่อัปเดตข้อมูล

3) ข้อมูลที่ไม่มีโครงสร้างแน่นอน (Unstructured data) มักจะหมายถึงข้อความยาวๆ ข้อความเสียง รูปภาพ และวิดีโอ

บางองค์กรอาจจะสามารถใช้ประโยชน์จากข้อมูลที่ไม่ได้ใหญ่มาก แต่มีความสำคัญ สามารถนำมาใช้วิเคราะห์ เพื่อให้เกิดความเข้าใจเชิงลึก (insight) ทำให้ได้มาซึ่งความรู้ความเข้าใจใหม่ๆ ที่ไม่สามารถเห็นได้อย่างตรงไปตรงมา หรือเป็นข้อมูลที่รวบรวมจัดเก็บไว้ตลอดกระบวนการเชิงธุรกิจ ที่สามารถนำมาวิเคราะห์ร่วมกันได้เห็นภาพชัดเจน เช่น การจัดระบบข้อมูลเกี่ยวกับการผลิต รวบรวมข้อมูลไว้ทุกปัจจัย ตั้งแต่ข้อมูลวัตถุดิบ ถึงข้อมูลของเครื่องจักรแต่ละกระบวนการ ไปจนถึงข้อมูลผลผลิต และการขนส่ง ทำให้สามารถวิเคราะห์ข้อมูลได้ตลอดทั้งกระบวนการทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

โดยทั่วไปลักษณะพื้นฐานของ Big Data จะมีอยู่ 3 ประการ คือ ปริมาณ (Volume), ความหลากหลาย (Variety) และ ความเร็ว (Velocity) แต่วันนี้เราจะมาพูดถึงคุณลักษณะที่เพิ่มขึ้นมาอีก 3 ประการ ที่ทำให้ข้อมูลขนาดใหญ่ (Big Data) นี้มีคุณภาพสูงขึ้นและเหมาะที่จะนำมาใช้ในการวิเคราะห์เพื่อให้ได้มาซึ่งความเข้าใจเชิงลึก Big data ที่มีคุณภาพสูงควรมีลักษณะพื้นฐานอยู่ 6 ประการหลักๆ (6 Vs) ดังนี้

1) ปริมาณ (Volume) หมายถึง ปริมาณของข้อมูลควรมีจำนวนมากพอ ทำให้เมื่อนำมาวิเคราะห์แล้วจะได้ Insights ที่ตรงกับความเป็นจริง เช่น การที่เรามีข้อมูลอายุ เพศ ของลูกค้าส่วนใหญ่ ทำให้เราสามารถหา Demographic Profile ทั่วไปของลูกค้าที่ถูกต้องได้ ถ้าเรามีข้อมูลลูกค้าแค่ส่วนน้อย ค่าที่ประมาณออกมาอาจจะไม่ตรงกับความเป็นจริง

2) ความหลากหลาย (Variety) หมายถึง รูปแบบของข้อมูลควรหลากหลายแตกต่างกันออกไป ทั้งแบบโครงสร้าง, กึ่งโครงสร้าง, ไม่มีโครงสร้าง ทำให้เราสามารถนำมาวิเคราะห์ประกอบกัน จนได้ insights ครบถ้วน

3) ความเร็ว (Velocity) หมายถึง คุณลักษณะข้อมูลที่ถูกสร้างขึ้นอย่างรวดเร็ว ต่อเนื่องและทันเหตุการณ์ ทำให้เราสามารถวิเคราะห์ข้อมูลแบบ real-time นำผลลัพธ์มาทำการตัดสินใจและตอบสนองได้อย่างทันท่วงที เช่น ข้อมูล GPS ที่ใช้ติดตามตำแหน่งของรถ อาจจะสามารถวิเคราะห์โอกาสที่ทำให้เกิดอุบัติเหตุ และออกแบบระบบป้องกันอุบัติเหตุได้

4) ความถูกต้อง (Veracity) หมายถึง มีความน่าเชื่อถือของแหล่งที่มาข้อมูล และความถูกต้องของชุดข้อมูล มีกระบวนการในการตรวจสอบและยืนยันความถูกต้องของข้อมูล ซึ่งมีความเกี่ยวเนื่องโดยตรงกับผลลัพธ์การวิเคราะห์ข้อมูล

5) คุณค่า (Value) หมายถึง ข้อมูลมีประโยชน์และมีความสัมพันธ์ในเชิงธุรกิจ ซึ่งต้องเข้าใจก่อนว่าไม่ใช่ทุกข้อมูลจะมีประโยชน์ในการเก็บและวิเคราะห์ ข้อมูลที่มีประโยชน์

จะต้องเกี่ยวข้องกับวัตถุประสงค์ทางธุรกิจ เช่นถ้าต้องการเพิ่มขีดความสามารถในการแข่งขันในตลาดของผลิตภัณฑ์ที่ขาย ข้อมูลที่มีประโยชน์ที่สุดน่าจะเป็นข้อมูลผลิตภัณฑ์ของคู่แข่ง

6) ความแปรผันได้ (Variability) หมายถึง ข้อมูลสามารถในการเปลี่ยนแปลงรูปแบบไปตามการใช้งาน หรือสามารถคิดวิเคราะห์ได้จากหลายแง่มุม และรูปแบบในการจัดเก็บข้อมูลก็อาจจะต่างกันออกไปในแต่ละแหล่งของข้อมูล

คุณลักษณะเหล่านี้ทำให้การจัดการ Big Data เป็นเรื่องยาก แต่ในขณะเดียวกันก็เป็นโอกาสให้องค์กรสร้างความได้เปรียบเหนือคู่แข่งด้วยการพัฒนาขีดความสามารถในการวิเคราะห์ข้อมูล

2.2.1.2 การจัดการข้อมูลขนาดใหญ่

การรวบรวมข้อมูลมาจัดเก็บ (Storage) เป็นการรวบรวมข้อมูลของจากหลากหลายทั้งที่มา และการใช้งานที่แตกต่างกันอย่างมากมาย ซึ่งกลไกและเทคโนโลยีแบบดั้งเดิม ETL (Extract Transform and Load) ไม่สามารถทำได้ ซึ่ง Big Data หรือ ข้อมูลขนาดใหญ่ต้องการเทคนิค วิธีการ และเทคโนโลยีใหม่ในการรวบรวมข้อมูลขนาด เทราไบต์ และเพตาไบต์ ในการรวบรวมข้อมูลนั้นต้องมีการประมวลผล จัดรูปแบบ ให้เหมาะสำหรับการใช้ในการวิเคราะห์หรือใช้งานสำหรับธุรกิจหรือวัตถุประสงค์นั้น ๆ

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

2) การวิเคราะห์ (Analyst) การวิเคราะห์ข้อมูลทำให้เกิดความกระจ่าง และชัดเจนในชุดข้อมูลที่มีอยู่ในการสำรวจข้อมูลยังทำให้ค้นพบสิ่งใหม่ แชร์สิ่งทีค้นพบใหม่ ๆ ต่อผู้อื่น สร้างรูปแบบจำลองข้อมูลด้วยการเรียนรู้ของเครื่องจักร และปัญญาประดิษฐ์ AI และนำข้อมูลเหล่านั้นไปใช้งาน

2.2.2 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

การทำเหมืองข้อมูลหรือ Data Mining เทคโนโลยีเกี่ยวกับฐานข้อมูลได้เริ่มต้นขึ้นตั้งแต่ปี 1960 โดยพัฒนามาจาก File Processing จากนั้นพัฒนามาในรูปแบบตาราง และเริ่มมีการใช้ภาษาในการเรียกข้อมูล (Query Language) ข้อมูลที่เก็บเริ่มมีจำนวนมากขึ้น ทำให้การจัดเก็บมีความซับซ้อน และมีประสิทธิภาพเพิ่มขึ้นของข้อมูลเป็นจำนวนมาก และสะสมไว้ในฐานข้อมูลนั้นมีขนาดใหญ่เกินกว่าที่กำลังคนจัดการได้ เป็นผลให้เกิดเครื่องมือที่ช่วยในการวิเคราะห์ข้อมูล และหาความเป็นไปได้ของข้อมูลที่เป็นประโยชน์

เนื่องด้วยปัจจุบันเป็นยุคที่ข้อมูลสารสนเทศมีความสำคัญ การเผยแพร่และสื่อสารข้อมูล ข่าวสาร ที่ตรงกับความต้องการของผู้ใช้จึงเป็นสิ่งจำเป็น การประยุกต์เทคโนโลยีสารสนเทศเพื่อช่วยในการสื่อสารข้อมูลจำนวนมากให้แก่ผู้ใช้ เช่น การให้บริการเว็บไซต์เพื่อเผยแพร่ข้อมูล ข่าวสาร และแลกเปลี่ยนความรู้ จึงเป็นเครื่องมือที่สำคัญในการสื่อสารข้อมูลถึงผู้ใช้งานจำนวนมาก ดังนั้น การศึกษาเกี่ยวกับพฤติกรรมของผู้ใช้บริการเว็บไซต์จะช่วยให้องค์กรสามารถนำข้อมูลมาใช้ในการวางแผนพัฒนาเว็บไซต์ ให้ตรงกับความต้องการใช้งาน หรือใช้ในการวางแผนกลยุทธ์ เพื่อสร้างความได้เปรียบทางการแข่งขัน

การทำเหมืองข้อมูล (Data Mining) คือ วิธีการค้นหาความรู้ความสัมพันธ์ และรูปแบบทั้งหมดที่ถูกซ่อนอยู่ในฐานข้อมูลที่มีปริมาณข้อมูลจำนวนมาก การทำเหมืองข้อมูลสามารถทำการสำรวจ และวิเคราะห์อย่างอัตโนมัติ จากข้อมูลที่มีอยู่ในรูปแบบความหมาย และรูปแบบของกฎ โดยที่ความสัมพันธ์แสดงให้เห็นถึงความรู้ต่าง ๆ ที่เป็นประโยชน์และน่าสนใจ ประโยชน์หลักของการทำเหมืองข้อมูลคือ การสืบค้นความรู้ในฐานข้อมูลขนาดใหญ่เพื่อให้ได้ความรู้มาใช้ในการตัดสินใจ

การทำเหมืองข้อมูลหากผนวกเข้ากับระบบการจัดการความรู้ที่ได้จากฐานข้อมูล จะทำให้ได้ระบบการค้นหาข้อมูลและความรู้ หรือเรียกว่า (Knowledge and Data Discovery Management System : KDDMS) เป็นการทำเหมืองข้อมูลที่ใช้เทคนิคการกลั่นกรอง วิเคราะห์ความรู้ที่ซ่อนอยู่ในฐานข้อมูลที่มีปริมาณมาก เพื่อได้ข้อมูล ที่ต้องการ และมีประโยชน์ สามารถนำข้อมูลที่ได้รับการจัดการได้อย่างมีประสิทธิภาพ โดยผลลัพธ์มีอยู่หลายแบบ เช่น รูปแบบ (Pattern) ต้นแบบ (Model) กฎ (Rule) การทำนายหรือการคาดการณ์ล่วงหน้า (Prediction) ทั้งผลลัพธ์ที่ได้มาขึ้นอยู่กับการทำเหมืองข้อมูล

2.2.3 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

การออกแบบเว็บไซต์นั้นมีความสำคัญอย่างมาก เนื่องจากจะทำให้ผู้ชมเกิดความประทับใจ ซึ่งชอบ จนสามารถดึงดูดความสนใจได้จนกระทั่งมีความต้องการที่จะกลับเข้าชมเว็บนั้น ๆ อีกครั้งในอนาคต สำหรับการออกแบบเว็บไซต์นั้น หมายถึง การออกแบบพัฒนาส่วนหน้าแสดงของเว็บให้เกิดความประทับใจแก่ผู้ชม และนำเสนอความเป็นเอกลักษณ์ของหน่วยงานได้ด้วยโปรแกรมฟังก์ชันการใช้งานภายในเว็บไซต์ที่ทำให้ผู้ใช้อยากเข้ามามีปฏิสัมพันธ์ต่อเว็บไซต์นั้น โดยต้องให้ความสำคัญกับปัจจัยต่าง ๆ ดังต่อไปนี้

2.2.3.1 ความสำคัญของการออกแบบเว็บไซต์

เว็บไซต์เป็นสื่อที่อยู่ในการควบคุมของผู้ใช้โดยสมบูรณ์ ดังนั้นการออกแบบเว็บจึงมีความสำคัญที่จะดึงดูดให้ผู้ใช้เข้าชม และทำกิจกรรมร่วมในเว็บไซต์ของเราให้ได้มากที่สุด ซึ่งสามารถสรุปปัญหาต่าง ๆ เกี่ยวกับผู้ใช้ซึ่งเป็นองค์ประกอบหนึ่งในการกำหนดความสำเร็จ หรือล้มเหลวของการสร้างเว็บไซต์ ดังต่อไปนี้

- 1) ผู้ใช้มักไม่อดทนกับอุปสรรค และปัญหาที่เกิดขึ้นจากการออกแบบการใช้งานผิดพลาด
- 2) ผู้ใช้มีทางเลือกมากขึ้นที่จะเข้าใช้บริการเว็บไซต์อื่น ๆ ถ้าหากการออกแบบเว็บไซต์นั้นเกิดความรู้สึกไม่ประทับใจ
- 3) ผู้ใช้หลายคนชื่นชอบเว็บไซต์ที่สวยงาม และมีความสะดวกในการใช้งานมากกว่าเว็บที่ดูสับสนวุ่นวาย มีข้อมูลจำนวนมากแต่ค้นหาข้อมูลที่ต้องการ ไม่พบ หรือใช้เวลามากในการแสดงหน้าเว็บ
- 4) หากผู้ใช้มีความประทับใจกับเว็บไซต์หนึ่งแล้ว จะทำให้อยากกลับมาเข้าใช้อีกครั้ง

2.2.3.2 องค์ประกอบของการออกแบบเว็บไซต์อย่างมีประสิทธิภาพ

การออกแบบเว็บไซต์ที่ดีนั้น จะต้องมียุทธศาสตร์ประกอบในการออกแบบให้เหมาะสมกับกลุ่มเป้าหมาย และจะต้องคำนึงถึงความสะดวกของผู้ใช้งานเป็นหลัก โดยที่ผู้ออกแบบนั้น ต้องให้ความสำคัญระหว่างความสวยงามและความสะดวก ซึ่งต้องนึกถึงองค์ประกอบดังต่อไปนี้

- 1) ความเรียบง่าย (Simplicity) คือ การสื่อเนื้อหาถึงผู้ใช้โดยจำกัดองค์ประกอบเสริมที่เกี่ยวข้องกับการนำเสนอให้เหลือแต่สิ่งที่จำเป็นเท่านั้น ทั้งการแสดงผลออกใน

ด้านเนื้อหาตัวอักษร หรือแม้แต่สีที่ใช้ ทำให้ผู้ใช้รู้สึกใช้งานง่ายและสะดวกไม่ก่อให้เกิดความรู้สึกรำคาญใด ๆ

2) ความสม่ำเสมอ (Consistency) คือ การสื่อให้ผู้ใช้รู้ว่าเป็นเนื้อหาเดียวตลอดทั้งเว็บไซต์ โดยไม่ก่อให้เกิดความสับสนแก่ผู้ใช้ ดังนั้น ผู้ออกแบบเว็บไซต์ควรออกแบบระบบ Navigation และโทนสีให้มีความสม่ำเสมอคล้ายคลึงกันตลอดทั้งเว็บไซต์

3) ความเป็นเอกลักษณ์ (Identity) คือ ต้องออกแบบเพื่อแสดงถึงความเป็นเอกลักษณ์ขององค์กร ทั้งชุดสี ตัวอักษร รูปภาพกราฟิก เป็นสื่อถึงความเป็นเอกลักษณ์แทบทั้งสิ้น

4) เนื้อหาที่ประประโยชน์ (Useful content) เนื้อที่แสดงออกของเว็บไซต์นั้น เป็นสิ่งสำคัญที่สุด ซึ่งต้องแสดงออกถึงความถูกต้องและสมบูรณ์ โดยไม่ซ้ำกับเว็บไซต์อื่น และเป็นเนื้อหา ที่ทันต่อเหตุการณ์ เนื่องจากเป็นสิ่งที่ดึงดูดผู้เข้ามาใช้งานเว็บไซต์

5) ระบบ Navigation ที่ใช้งานง่าย (User-friendly navigation) เป็นส่วนที่สำคัญของเว็บไซต์ เนื่องจากจะทำให้ผู้ใช้ไม่เกิดความสับสนเวลาใช้งานเว็บไซต์ ซึ่งผู้ออกแบบจำเป็นต้อง ออกแบบให้ผู้ใช้ใช้งานได้ง่ายและสะดวก เช่นถ้าใช้ภาพกราฟิกที่สื่อความหมายควรมีคำอธิบายที่ชัดเจนร่วมด้วย แสดงไว้ส่วนบนหรือส่วนล่างของหน้าเว็บไซต์

6) มีลักษณะที่น่าสนใจ (Visual appeal) ควรสื่อให้เว็บไซต์มีคุณภาพสัมพันธ์กับองค์ประกอบต่าง ๆ เช่น คุณภาพของกราฟิกที่แสดงบนเว็บไซต์นั้นต้องสมบูรณ์ไม่มีร่องรอยของความเสียหายเป็นจุดต่าง หรือมีขอบขั้นบันได สำหรับตัวอักษรต้องเป็นตัวอักษรที่อ่านง่าย และสบายตาพร้อมกับโทนสีที่สวยงาม

7) การใช้งานได้อย่างไม่จำกัด (Compatible) การออกแบบเว็บไซต์นั้นต้องทำให้คนส่วนใหญ่เข้าถึงให้ได้มากที่สุด โดยไม่มีการบังคับให้ติดตั้งโปรแกรมใด ๆ เพิ่มเติม สามารถแสดงผลได้ทุกระบบปฏิบัติการ และที่ความละเอียดของหน้าจอต่าง ๆ กัน

8) คุณภาพการออกแบบ (Design Stability) ต้องออกแบบให้เว็บไซต์มีเนื้อหาที่น่าเชื่อถือ เพื่อดึงดูดให้มีผู้เข้าชมเว็บไซต์เกิดความประทับใจ และมีความเชื่อถือในตัวเว็บไซต์

9) ระบบการใช้งานที่ถูกต้อง (Functional stability) ต้องออกแบบให้เว็บไซต์นั้นมีความแน่นอนทำหน้าที่ได้อย่างถูกต้อง และสามารถใช้งานได้จริง เช่น ถ้ามีแบบฟอร์มให้ผู้ใช้

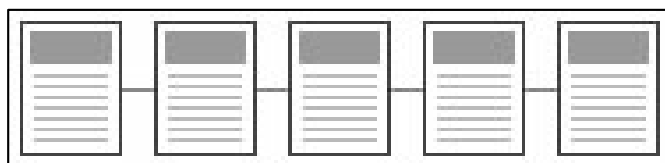
กรอกข้อมูล ก็ต้องทดสอบก่อนว่าสามารถใช้งานได้จริง หรือลิงค์ต่าง ๆ ที่แสดงอยู่บนหน้าเว็บไซต์ จะต้องเชื่อมโยงไปยังแหล่งข้อมูลนั้น ๆ ได้จริง

2.2.3.3 รูปแบบโครงสร้างของเว็บไซต์

การออกแบบโครงสร้างของเว็บไซต์ สามารถทำได้หลากหลายแบบซึ่งก็ขึ้นอยู่กับความชอบ และความถนัดของแต่ละบุคคล นอกจากนี้ยังขึ้นอยู่กับกลุ่มเป้าหมายที่ต้องการนำเสนอ เพราะจะต้องออกแบบให้เหมาะกับการใช้งานของกลุ่มเป้าหมายมากที่สุด โดยโครงสร้างของเว็บไซต์ส่วนใหญ่ก็จะประกอบไปด้วย 4 รูปแบบดังนี้

1) โครงสร้างเว็บไซต์แบบเรียงลำดับ

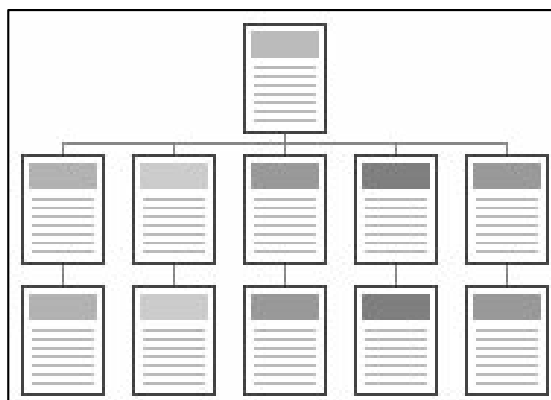
จะเป็นโครงสร้างแบบธรรมดาที่นิยมใช้งานกันมากที่สุด เนื่องจากมีความง่ายต่อการจัดระบบข้อมูล และสามารถนำเสนอเรื่องราวตามลำดับได้เป็นอย่างดี เหมาะกับเว็บไซต์ที่มีขนาดเล็ก มีเนื้อหาที่ไม่ซับซ้อน ส่วนใหญ่ก็จะเป็นพวกเว็บไซต์ที่ให้ความรู้ หรือเว็บไซต์องค์กรขนาดย่อม โดยลักษณะการลิงค์เนื้อหา ก็จะลิงค์ไปที่ละหน้า มีทิศทาง การเข้าสู่เนื้อหาต่าง ๆ ในแบบเส้นตรง ใช้ปุ่มเดินหน้า-ถอยหลัง ในการกำหนดทิศทาง จึงทำให้การใช้งานเป็นไปอย่างง่าย แต่โครงสร้างเว็บไซต์แบบเรียงลำดับก็มีข้อเสีย คือจะทำให้ผู้ใช้งานต้องเสียเวลาในการเข้าสู่เนื้อหา เพราะไม่สามารถกำหนดทิศทาง การเข้าสู่เนื้อหาด้วยตัวเองได้



ภาพที่ 2.10 แสดงโครงสร้างเว็บไซต์แบบเรียงลำดับ

2) โครงสร้างเว็บไซต์แบบลำดับชั้น

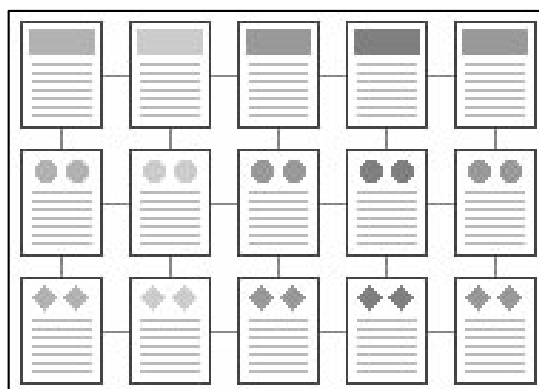
นิยมใช้กับเว็บที่มีความซับซ้อนของข้อมูล เพื่อให้สามารถเข้าถึงข้อมูลต่าง ๆ ได้ง่ายขึ้น โดยจะมีการแบ่งเนื้อหาออกเป็นส่วน ๆ และมีการนำเสนอรายละเอียดค่อย ๆ ที่ลดหลั่นกันมา ทำให้สามารถทำความเข้าใจกับโครงสร้างเนื้อหาได้ง่ายขึ้น โดยจะมีโฮมเพจเป็นจุดเริ่มต้น และจุดรวมจุดเดียวที่จะนำไปสู่การเชื่อมโยงเนื้อหาเป็นลำดับจากบนลงล่าง



ภาพที่ 2.11 แสดงโครงสร้างเว็บไซต์แบบลำดับชั้น

3) โครงสร้างเว็บไซต์แบบตาราง

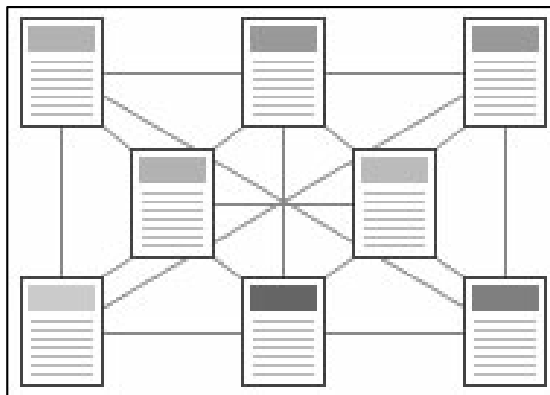
เป็นโครงสร้างการออกแบบเว็บไซต์ที่มีความซับซ้อน แต่ก็มีคามยืดหยุ่นในระดับหนึ่ง เพื่อให้ผู้ใช้งานสามารถเข้าสู่เนื้อหาต่าง ๆ ได้ง่ายขึ้น การออกแบบในลักษณะนี้จะมีการเชื่อมโยงเนื้อหาในแต่ละส่วนซึ่งกันและกัน ทำให้ผู้ใช้งานสามารถเปลี่ยนทิศทาง หรือกำหนดทิศทางในการเข้าสู่เนื้อหาด้วยตัวเองได้จึงไม่ทำให้เสียเวลา แถมยังทำให้เว็บไซต์มีความทันสมัยขึ้น



ภาพที่ 2.12 แสดงโครงสร้างเว็บไซต์แบบตาราง

4) โครงสร้างแบบใยแมงมุม

เป็นโครงสร้างที่ได้รับความนิยมเป็นอย่างมาก เพราะมีความยืดหยุ่นมากที่สุด โดยทุกหน้าเว็บจะมีการเชื่อมโยงถึงกันหมด ทำให้สามารถเข้าถึงหน้าเว็บเพจต่าง ๆ ที่ต้องการได้อย่างง่าย และมีความอิสระมากขึ้น นอกจากนี้ก็สามารถเชื่อมโยงไปสู่เว็บไซต์ภายนอกได้ดี



ภาพที่ 2.13 แสดงโครงสร้างเว็บไซต์แบบใยแมงมุม

2.2.3.4 ส่วนประกอบสำคัญของหน้าเว็บเพจบนหน้าเว็บเพจ จะมีส่วนประกอบสำคัญที่จำเป็นต้องมีอยู่ 3 ส่วน ได้แก่

1) ส่วนหัวของหน้า (Header) อยู่ตอนบนสุดของหน้าและเป็นส่วนที่สำคัญที่สุด โดยจะต้องทำให้สามารถดึงดูดผู้ชมให้รู้สึกอยากติดตามเนื้อหาในเว็บไซต์ต่อไป ซึ่งส่วนใหญ่มักจะมีการใส่ภาพกราฟิกให้ดูสวยงาม สิ่งสำคัญหลัก ๆ เลย ก็คือ โลโก้ ชื่อเว็บไซต์ และเมนูหลักที่สามารถลิงค์ไปยังเนื้อหาในหน้าเว็บเพจต่าง ๆ ได้

2) ส่วนของเนื้อหา (Body) อยู่บริเวณตอนกลางของหน้าเว็บ โดยจะแสดงข้อมูลเกี่ยวกับเนื้อหาบนเว็บแบบคร่าวๆ ซึ่งก็จะมีข้อความกราฟิก ตารางข้อมูลหรือวิดีโอประกอบอยู่ และหากมีเมนูแบบเฉพาะกลุ่มก็จะถูกจัดไว้ในหน้านี้เช่นกัน และที่สำคัญเนื้อหาในส่วนนี้ควรจะมีคามกระชับ เข้าใจง่าย มีการใช้รูปแบบตัวอักษรแบบเรียบง่ายและเป็นระเบียบ

3) ส่วนท้ายของหน้า (Footer) อยู่ล่างสุดของหน้าเว็บ ซึ่งจะมีหรือไม่มีก็ได้ ส่วนนี้จะแสดงถึงข้อมูลต่าง ๆ เพิ่มเติมเข้าไป เช่น ข้อความที่แสดงถึงการเป็นลิขสิทธิ์ ข้อมูลเจ้าของเว็บไซต์ วิธีการติดต่อและคำแนะนำต่าง ๆ เกี่ยวกับการใช้งานเว็บไซต์อย่างถูกต้อง เป็นต้น

2.2.3.5 วิธีการเลือกใช้สีสำหรับการออกแบบเว็บไซต์

การเลือกใช้สีในการออกแบบเว็บไซต์มีความสำคัญเป็นอย่างมาก เพราะสีสามารถกำหนดอารมณ์ ความรู้สึก และกระตุ้นการรับรู้ทางด้านจิตใจของมนุษย์ได้ดี ดังนั้น สีที่ใช้จึงต้องมีความสอดคล้องกับเนื้อหาและจุดประสงค์ของเว็บไซต์ว่าต้องการให้ผู้เข้าชมรู้สึกอย่างไร ต่อเนื้อหาที่ได้อ่าน โดยรูปแบบของสีที่สายตาของมนุษย์สามารถมองเห็นได้ก็แบ่งออกเป็น 3 กลุ่ม ดังต่อไปนี้

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 ดังนั้น หากเลือกใช้สีไม่เหมาะสมกับเนื้อหาหรือจุดประสงค์ของเว็บไซต์ ก็จะทำให้เว็บดูไม่น่าสนใจ ขาดความน่าเชื่อถือและทำให้ผู้ที่เคยเข้ามาใช้บริการไม่คิดจะกลับมาใช้บริการอีก

2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่ง HTML

HTML ย่อมาจาก Hyper Text Markup Language คือ ภาษาคอมพิวเตอร์ที่ใช้ในการแสดงผลของเอกสารบน Website หรือที่เราเรียกกันว่าเว็บเพจ ถูกพัฒนาและกำหนดมาตรฐานโดยองค์กร World Wide Web Consortium (W3C) และจากการพัฒนาทางด้าน Software ของ Microsoft ทำให้ภาษา HTML เป็นอีกภาษาหนึ่งที่ใช้เขียนโปรแกรมได้หรือที่เรียกว่า HTML Application

HTML เป็นภาษาประเภท Markup สำหรับการการสร้างเว็บเพจ โดยใช้ภาษา HTML สามารถทำได้โดยใช้โปรแกรม Text Editor ต่าง ๆ เช่น Notepad, Edit plus หรือจะอาศัยโปรแกรมที่เป็นเครื่องมือช่วยสร้างเว็บเพจ เช่น Microsoft FrontPage, Dream Weaver ซึ่งอำนวยความสะดวกในการสร้างหน้า HTML ส่วนการเรียกใช้งานหรือทดสอบการทำงานของเอกสาร HTML จะใช้

โปรแกรม Web Browser เช่น IE Microsoft Internet Explorer (IE), Mozilla Firefox, Safari, Opera, และ Netscape Navigator เป็นต้น

2.2.4.1 ไวยากรณ์ของ CSS

ภาษา HTML จะแบ่งออกเป็น 2 ส่วน คือ

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

2) ส่วนของบทความทั่ว ๆ ไป เป็นส่วนของข้อความที่เราต้องการแสดงผล ตัวอย่างการใช้งานภาษา HTML

```
<html>¶
<head>¶
    <title>หัวข้อเรื่องของหน้านี้</title>¶
</head>¶
<body>¶
¶
    เนื้อหาที่จะแสดงใน Web Browser¶
¶
</body>¶
</html>¶
```

ภาพที่ 2.14 ไวยากรณ์ของ HTML

การทำงานของ Code

1) <html> </html> ในการใช้งาน HTML เราจะต้องเริ่มด้วย <html> และปิดด้วย </html> เสมอ

2) <head> </head> เป็นส่วนที่ใช้ให้รายละเอียดเกี่ยวกับเว็บเพจ หน้านี้ ซึ่งจะไม่แสดงให้เห็นในส่วนของการแสดงผลของ Web Browser แต่จะมีผลกับส่วนอื่น ๆ เช่น

การหาของ Search Engine (Google, Yahoo) การใช้งานก็จะมีคำสั่งย่อยเพื่อบรรยายรายละเอียด เช่น <title> </title>, <meta> และอื่นอีกมากมาย

3) <title> </title> ในส่วนตัวอักษรที่อยู่ในคำสั่งนี้จะอยู่ใน title bar ของ web page

4. <body> </body> ตัวอักษรที่อยู่ในคำสั่งนี้จะแสดงส่วนแสดงผล ของ Web Browser

2.2.5 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS

CSS ย่อมาจาก Cascading Style Sheet มักเรียกโดยย่อว่า "สไตลชีต" คือ ภาษาที่ใช้เป็นส่วนของการจัดรูปแบบการแสดงผลเอกสาร HTML โดยที่ CSS กำหนดกฎเกณฑ์ในการระบุรูปแบบ หรือ Style ของเนื้อหาในเอกสาร อันได้แก่ สีของข้อความ สีพื้นหลัง ประเภทตัวอักษร และการจัดวางข้อความ ซึ่งการกำหนดรูปแบบ หรือ Style นี้ใช้หลักการของการแยกเนื้อหาเอกสาร HTML ออกจากคำสั่งที่ใช้ในการจัดรูปแบบการแสดงผล กำหนดให้รูปแบบของการแสดงผลเอกสาร ไม่ขึ้นอยู่กับเนื้อหาของเอกสาร เพื่อให้ง่ายต่อการจัดรูปแบบการแสดงผลล์พ์ของเอกสาร HTML โดยเฉพาะในกรณีที่มีการเปลี่ยนแปลงเนื้อหาเอกสารบ่อยครั้ง หรือต้องการควบคุมให้รูปแบบการแสดงผลเอกสาร HTML มีลักษณะของความสม่ำเสมอทั่วกันทุกหน้าเอกสารภายในเว็บไซต์เดียวกัน โดยกฎเกณฑ์ในการกำหนดรูปแบบ (Style) เอกสาร HTML ถูกเพิ่มเข้ามาครั้งแรกใน HTML 4.0 เมื่อปีพ.ศ. 2539 ในรูปแบบของ CSS level 1 Recommendations ที่กำหนดโดย องค์การ World Wide Web Consortium หรือ W3C

2.2.5.1 ไวยากรณ์ของ CSS

ไวยากรณ์ของ CSS นั้น มีโครงสร้างหลักของภาษาอยู่ 4 ส่วนด้วยกัน ซึ่งทั้งหมดมีความสำคัญต่อการกำหนดค่าและการใช้งาน CSS ซึ่งรูปแบบการกำหนดค่าจะประกอบไปด้วย property ตามด้วยเครื่องหมายโคลอน (:) เพื่อกำหนดค่าลงไปซึ่งอาจจะเป็นตัวเลข คีย์เวิร์ด หรือ URL และสุดท้ายของการกำหนดค่าคือ เซมิ-โคลอน (;) ที่ใช้ในการระบุว่าสิ้นสุดการกำหนดค่าให้กับproperty ดังตัวอย่างที่ 1

Selector [Properties: Value;]

ภาพที่ 2.15 ไวยากรณ์ของ CSS

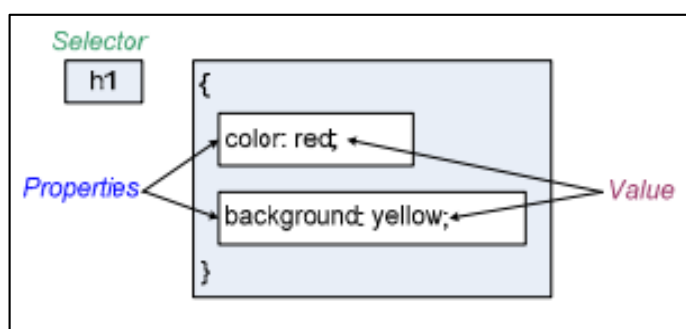
2.2.5.2 ส่วนประกอบของโครงสร้างของ CSS มีดังนี้

1) Selector สามารถเป็น HTML Tag ต่างๆ ก็ได้ เช่น <body>, <p> หรือเป็น Class name หรือ ID ที่ตั้งชื่อให้ก็ได้ซึ่งจะเรียนรู้ในหัวข้อถัดไป

2) เครื่องหมายปีกกาที่ครอบ Properties กับ Value ที่ต้องการกำหนดค่า

3) ชื่อของ Properties เป็นคุณสมบัติในการจัดรูปแบบการแสดงผล เช่น Color สำหรับกำหนดสี font-size สำหรับกำหนดขนาดตัวอักษร

4) ค่า Value ที่ต้องการจะกำหนดให้กับแต่ละ Properties เช่น color: black และ font-size:12px



ภาพที่ 2.16 แสดงซอร์ซโค้ดส่วนประกอบของโครงสร้างของ CSS

จากซอร์ซโค้ดข้างต้นนี้ สามารถพิจารณาตามรูปแบบโครงสร้างของ CSS ได้ดังนี้คือ Selector ของ CSS คือ h1, Properties ของ CSS คือ color และการใส่สีพื้นหลัง คือ Background, Value ของ CSS คือ red และ yellow

จะสังเกตเห็นได้ว่าจะมีเครื่องหมาย Colon (:) ที่เป็นตัวคั่นระหว่างชื่อ Properties กับ Value ที่เป็นการกำหนดค่าให้กับ Properties นั้น ๆ ซึ่งเมื่อสิ้นสุดแล้วแต่ละ Properties จะจบด้วย เครื่องหมาย Semi-colon (;) เสมอ ซึ่งเป็นข้อบังคับของการเขียน CSS นั่นเอง

2.2.6 ทฤษฎีเกี่ยวกับชุดคำสั่ง PHP

พีเอชพี (PHP) คือ ภาษาคอมพิวเตอร์ในลักษณะเซิร์ฟเวอร์-ไซด์ สคริปต์ โดยลิขสิทธิ์อยู่ในลักษณะโอเพนซอร์ส ภาษาพีเอชพีใช้สำหรับจัดทำเว็บไซต์ และแสดงผลออกมาในรูปแบบ HTML โดยมีรากฐานโครงสร้างคำสั่งมาจากภาษา ภาษาซี ภาษาจาวา และ ภาษาเพิร์ล ซึ่ง ภาษาพีเอชพี นั้นง่ายต่อการเรียนรู้ ซึ่งเป้าหมายหลักของภาษานี้คือให้นักพัฒนาเว็บไซต์สามารถเขียน เว็บเพจ ที่มีการตอบโต้ได้อย่างรวดเร็ว

PHP ย่อมาจากคำว่า “Personal Home Page Tool” (ปัจจุบันได้เพิ่มเติมคำย่อใหม่ โดยรวมกับตัวย่อเป็น PHP : PHP Hypertext Preprocessor) ซึ่งเป็นภาษาประเภท Script Language ที่ทำงานแบบ Server Side Script กระบวนการทำงานจะทำงานแบบโปรแกรมแปลคำสั่ง interpreter คือแปลภาษาทุกครั้งที่มีคนเรียกสคริปต์ ข้อดี คือ ไม่ต้องนำไปประมวลผลใหม่ (Compiler) เมื่อนำโปรแกรมไปใช้งาน หรือจะอัปเดตเวอร์ชันของโปรแกรม สามารถอัปเดตขึ้นไปทับไฟล์เดิมแล้วใช้งานได้ทันที ข้อเสียที่ต่างกันอย่างชัดเจนก็คือ กรณี Syntax ผิดจะรู้ก็ต่อเมื่อมีผู้ใช้งานเจอบั๊ก

ภาษา PHP จัดอยู่ในประเภท การเขียนโปรแกรมบนเว็บ (Web-based Programming) เพราะเราจะเก็บโค้ดคำสั่ง หรือสคริปต์ทั้งหมดที่เขียนขึ้นมาไว้บนเครื่องเซิร์ฟเวอร์ที่เดียว (Web Server) และให้ผู้ใช้งาน (Client) เรียกใช้งานโปรแกรมผ่านเว็บเบราว์เซอร์ต่าง ๆ เช่น Internet Explorer, Mozilla Firefox, Google Chrome, Opera, Safari ฯลฯ เพื่อนำข้อมูลมาแสดงผลที่หน้าจอของผู้ใช้แต่ละคนนั่นเอง

2.2.6.1 ไวยากรณ์ของ PHP

ภาษาพีเอชพี จะเป็นส่วนประกอบภายในเว็บเพจ โดยคำสั่งจะปรากฏระหว่าง `<?php ... ?>` เช่น

```
<?php
    echo "Hello, World!";
?>
```

ภาพที่ 2.17 ไวยากรณ์ของ PHP

2.2.7 ทฤษฎีเกี่ยวกับการ Visualization

Visualization เป็นส่วนประกอบสำคัญใน Cognitive System ซึ่งเป็นส่วนในการแสดงข้อมูลหรือผลลัพธ์ต่าง ๆ ในระหว่างคอมพิวเตอร์และผู้ใช้งานในรูปแบบของภาพโดยผู้ใช้งานสามารถเรียนรู้ และจดจำข้อมูลผ่านการมองเห็นได้มากกว่าการใช้ประสาทสัมผัสอื่น ๆ หรือจะกล่าวว่าการสร้างมโนภาพของสิ่งต่าง ๆ ที่เราสนใจขึ้นมาในใจ ซึ่งต่อมาได้กลายเป็นการนำภาพมาใช้กับการนำเสนอหรือนำมาเป็นกรอบความคิดซึ่งได้นำไปใช้ในการสนับสนุนการตัดสินใจ

Visualization เป็นมากกว่า วิธีการทาง Computer Visualization เป็นการนำเสนอข้อมูลแบบหนึ่งที่ทำให้การแสดงผลให้อยู่ในรูปแบบของ Visual Form ซึ่งอาจจะเป็นการแสดงในรูปแบบของรูปภาพ, กราฟ หรือแผนภาพ ซึ่งผลลัพธ์ของการทำ Visualization คือ การนำข้อมูลต่าง ๆ ที่ซ่อนอยู่ในตัวของข้อมูลเองออกมาให้ผู้ใช้งานสามารถสัมผัสได้ซึ่งกระบวนการต่าง ๆ ในการแสดงหรือนำข้อมูลต่าง ๆ ออกมา อาจจะถูกแอบซ่อนอยู่หรือผู้ใช้ไม่ทันสังเกต แต่อย่างไรก็ตาม Visualization เป็นสิ่งที่จำเป็นในการค้นหาข้อมูล หรือใช้ในการวิเคราะห์ข้อมูลโดยจุดมุ่งหมายของ Visualization ก็คือ การถ่ายทอดข้อมูลไปสู่ระบบการรับรู้โดยภาพของผู้ใช้ระบบ เพื่อช่วยในการลดช่องว่างระหว่างผู้ใช้และตัวข้อมูล และยังช่วยผู้ใช้งานสามารถเรียนรู้ ได้อย่างมีประสิทธิภาพ

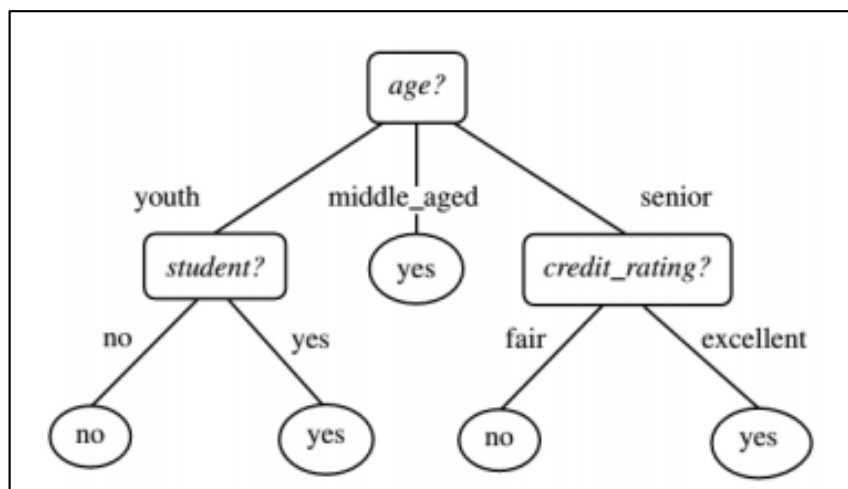
2.2.7.1 วิธีการของ Visualization

การใช้ Visualization เพื่อทำการค้นหาข้อมูลนั้นผู้ใช้งานจะทำขั้นตอนหลัก ๆ อยู่ 3 ขั้นตอน คือ Overview First, Zoom and Filter และ Detail on Demand โดยอันดับแรก ผู้ใช้ต้องการที่จะดูข้อมูลภาพรวมทั้งหมดซึ่งหลังจากดูภาพรวมทั้งหมดแล้วผู้ใช้งานก็จะทำการตัดสินใจเลือกรูปแบบหรือกลุ่มข้อมูลที่สนใจซึ่งก็จะมาถึงขั้นตอนในการวิเคราะห์ข้อมูลผู้ใช้งานก็จะทำการเจาะลึกถึงข้อมูลในรายละเอียด ซึ่ง Visualization Technology ก็จะอ้างอิงหรือพัฒนาจากขั้นตอนเหล่านี้ซึ่ง Visualization Technique มีประโยชน์มากในการแสดงภาพรวมหรือแสดงข้อมูลย่อยที่ผู้ใช้งานต้องการ โดยอาจจะใช้หลาย ๆ วิธีการรวมกัน เพื่อให้ได้ข้อมูลที่ผู้ใช้งานต้องการ ซึ่งช่วยลดช่องว่างของกิจกรรมที่ใช้ในการดึงข้อมูลต่าง ๆ ไปใช้ ซึ่งลักษณะของข้อมูลที่นำมาผ่านกระบวนการของ Visualization มีลักษณะต่าง ๆ มากมาย ดังนี้

ข้อมูล 1D ได้แก่ ข้อมูลที่เกี่ยวข้องกับช่วงเวลาใดเวลาหนึ่ง, ข้อมูล 2D ได้แก่ ข้อมูลที่เกี่ยวข้องกับแผนที่ภูมิศาสตร์, Multi-Dimensional Data ได้แก่ Relation Table, Text และ Hypertext ได้แก่ ข้อมูลหัวข้อข่าวต่าง ๆ และ Web Document, Hierarchies และ Graph ได้แก่ หมายเลขโทรศัพท์ และ Web Document, Algorithms และ Software ได้แก่ Debugging Operation ซึ่งแต่ละข้อมูลก็จะมีวิธีการที่ช่วยในการจัดการแสดงผลข้อมูลที่หลากหลาย

2.2.8 ทฤษฎีเกี่ยวกับการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ

การจำแนกข้อมูลด้วยต้นไม้ตัดสินใจจะเป็นกระบวนการสร้างต้นไม้ขึ้นเพื่อใช้ในการตัดสินใจจากข้อมูลที่มีหมวดหมู่ข้อมูลแนบอยู่ด้วย ต้นไม้ตัดสินใจจะประกอบไปด้วยโหนดต่าง ๆ ที่ไม่ใช่โหนดใบ (non-leaf node) ที่ซึ่งถูกใช้ในการแสดงถึงเงื่อนไขหรือแอทริบิวต์หนึ่ง ๆ ของข้อมูล โดยที่แต่ละกิ่งก้านของโหนดหนึ่ง ๆ จะหมายถึงค่าที่เป็นไปได้จากการทดสอบกับแอทริบิวต์นั้น ๆ และจะประกอบไปด้วยโหนดใบ (leaf node) ที่ซึ่งจะมีหมวดหมู่ข้อมูลจัดเก็บอยู่ โดยตัวอย่างต้นไม้ตัดสินใจจะถูกแสดงใน ภาพที่ 2.14 ด้านล่าง ที่จะแสดงการทำนายคุณลักษณะของลูกค้าที่จะทำการซื้อคอมพิวเตอร์จากร้านขายอุปกรณ์ไฟฟ้า โดยโหนดต่าง ๆ ที่ไม่ใช่โหนดใบจะถูกแทนด้วยสี่เหลี่ยม และโหนดใบจะถูกแทนด้วยวงรีตามลำดับ จากรูปจะเห็นว่าโหนดใบจะเป็นโหนดที่บ่งบอกถึงข้อมูลหมวดหมู่ของคำตอบที่เราต้องการ อาทิ เช่น “yes” หมายถึง ลูกค้าจะซื้อคอมพิวเตอร์ และ “no” หมายถึง ลูกค้าจะไม่ซื้อคอมพิวเตอร์โดยต้นไม้ที่ถูกสร้างขึ้น อาจเป็นต้นไม้ที่มีลักษณะเป็นไบนารีหรืออาจจะไม่เป็นไบนารีก็ได้ หลังจากทำการสร้างต้นไม้ตัดสินใจแล้ว เราจะสามารถใช้ต้นไม้ตัดสินใจในการจำแนกข้อมูลได้ โดยจะทำการจำแนกหมวดหมู่ของข้อมูลเรคคอร์ดหนึ่ง ๆ ที่ประกอบไปด้วยแอทริบิวต์ต่าง ๆ แต่เราจะไม่ทราบหมวดหมู่ข้อมูลในเรคคอร์ดนั้น ๆ ด้วยการเปรียบเทียบแอทริบิวต์ที่อยู่ในโหนดราก กับค่าของแอทริบิวต์ในเรคคอร์ดที่พิจารณา โดยจะทำการเปรียบเทียบจากโหนดรากไปจนถึงโหนดใบ เมื่อเราทราบถึงโหนดใบจะทำให้เราทราบถึงหมวดหมู่ข้อมูลของเรคคอร์ดที่ทำการพิจารณา



ภาพที่ 2.18 แสดงตัวอย่างต้นไม้ตัดสินใจของลูกค้าที่ทำการซื้อคอมพิวเตอร์

การจำแนกหมวดหมู่ข้อมูลด้วยต้นไม้ตัดสินใจได้รับความนิยมเป็นอย่างมาก และถูกประยุกต์ใช้ในหลาย ๆ งาน อาทิ เช่น การผลิตและการใช้ยา (Medicine) การผลิตสินค้าต่าง ๆ (Manufacturing and Production) การวิเคราะห์ทางการเงิน (Financial Analysis) ดาราศาสตร์ (Astronomy) อณูชีววิทยา (Molecular Biology) เป็นต้น สาเหตุที่ต้นไม้ตัดสินใจได้รับความนิยมอันเนื่องมาจากเหตุผลหลายประการด้วยกัน เช่น 1) ไม่ต้องการองค์ความรู้ใด ๆ หรือการกำหนดค่าพารามิเตอร์ใด ๆ เพื่อที่จะทำการสร้างต้นไม้ตัดสินใจ 2) สามารถจัดการกับข้อมูลที่มีหลาย ๆ มิติหรือข้อมูลที่มีหลาย ๆ แอตริบิวต์ได้ 3) ผลลัพธ์ที่ได้จะเป็นต้นไม้ตัดสินใจที่อยู่ในรูปแบบที่เข้าใจง่าย 4) ขั้นตอนการสร้างต้นไม้ตัดสินใจค่อนข้างง่ายและสามารถทำงานได้อย่างรวดเร็ว และ 5) มักจะให้ผลการจำแนกข้อมูลที่มีความถูกต้องค่อนข้างสูง ที่ซึ่งอาจขึ้นอยู่กับคุณลักษณะ ของข้อมูลที่ทำการพิจารณาด้วยเช่นกัน

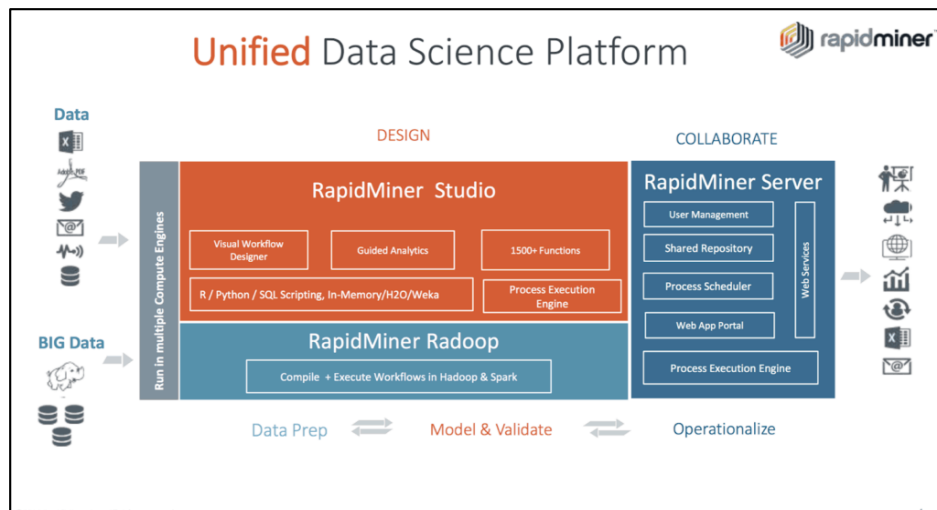
2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.3.1 โปรแกรม Rapid Miner

Rapid Miner เป็นวิทยาศาสตร์ข้อมูลแพลตฟอร์มซอฟต์แวร์ที่พัฒนาโดย บริษัท ที่มีชื่อเดียวกันที่ให้สภาพแวดล้อมแบบบูรณาการสำหรับการเตรียมข้อมูล, การเรียนรู้เครื่อง, การเรียนรู้ลึก, การทำเหมืองข้อความ, และการวิเคราะห์การทำนาย มันถูกใช้สำหรับธุรกิจและการใช้งานเชิงพาณิชย์ เช่นเดียวกับการวิจัยการศึกษาการฝึกอบรมการสร้างต้นแบบอย่างรวดเร็ว และการพัฒนาแอปพลิเคชัน และสนับสนุนทุกขั้นตอนของเครื่องกระบวนการการเรียนรู้รวม ถึงการเตรียมข้อมูลผลการสร้างภาพการตรวจสอบรูปแบบ และการเพิ่มประสิทธิภาพ

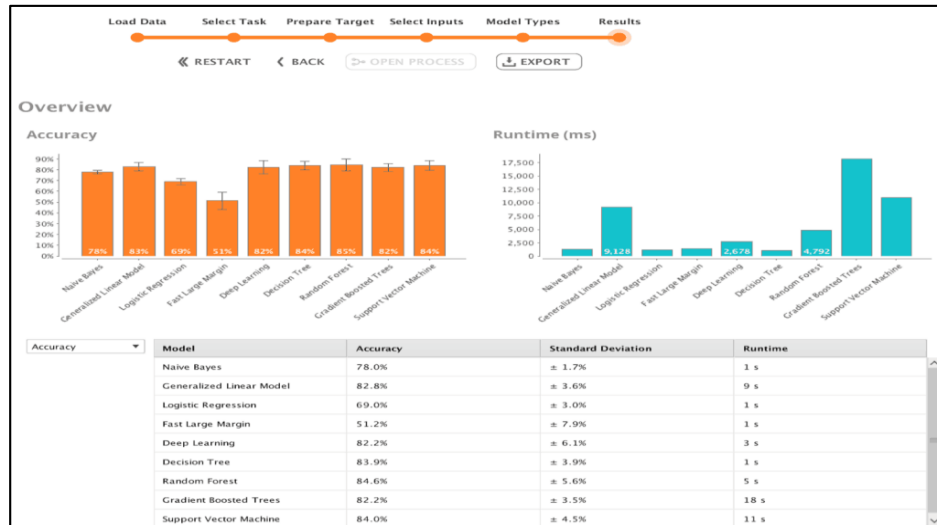
ตามการวิจัยของ Bloor Rapid Miner มอบคุณค่าการวิเคราะห์ขั้นสูง 99% ผ่านกรอบการทำงานของเทมเพลตที่ช่วยให้การจัดส่ง และลดข้อผิดพลาดโดยเกือบคำนึงถึง]ไม่จำเป็นต้องเขียนโค้ด Rapid Miner มีขั้นตอนการชุดข้อมูลและการเรียนรู้ของเครื่องรวมถึง: การไหลข้อมูลและการแปลง(ETL) การประมวลผลข้อมูลล่วงหน้าและการมองเห็นการวิเคราะห์เชิงพยากรณ์และการสร้างแบบจำลองทางสถิติการประเมินผลและการปรับใช้ Rapid Miner เขียนด้วยภาษาการเขียนโปรแกรม Java Rapid Miner มี GUI เพื่อออกแบบและดำเนินการเวิร์กโฟลว์การวิเคราะห์เวิร์กโฟลว์เหล่านั้นเรียกว่า "กระบวนการ" ใน Rapid Miner และประกอบด้วย "ตัวดำเนินการ" หลายรายการ ผู้ประกอบการแต่ละคนดำเนินงานเดี่ยวภายในกระบวนการและเอาต์พุตของผู้ประกอบการแต่ละรูปแบบการป้อนข้อมูลของหนึ่งต่อไป อีกวิธีหนึ่งเครื่องยนต์สามารถเรียกจากโปรแกรมอื่น ๆ หรือใช้เป็น API ฟังก์ชันส่วนบุคคลสามารถเรียกได้จากบรรทัดคำสั่ง Rapid Miner จัดเตรียมแผนการเรียนรู้แบบจำลองและอัลกอริทึมและสามารถขยายได้โดยใช้สคริปต์ R และ Python [6]

ฟังก์ชัน Rapid Miner สามารถขยายได้ด้วยปลั๊กอินเพิ่มเติมซึ่งมีให้ผ่านทาง Rapid Miner Marketplace Rapid Miner Marketplace เป็นแพลตฟอร์มสำหรับนักพัฒนาในการสร้างอัลกอริทึมการวิเคราะห์ข้อมูลและเผยแพร่สู่ชุมชน ซึ่ง Rapid Miner platform สามารถทำทั้งหมดนี้ได้ โดย Rapid Miner จะมี 3 โมดูลใหญ่ๆ ได้แก่ Rapid Miner Studio , Rapid Miner Server , Rapid Miner Radoop ตามรูปดังนี้



ภาพที่ 2.19 แสดงตัวอย่าง RapidMiner platform

โมดูลแรก : Rapid Miner Studio ใช้สำหรับการออกแบบการวิเคราะห์ข้อมูลผ่านทางหน้า GUI ซึ่งสามารถทำการจัดการข้อมูล และสร้างโมเดลแบบต่าง ๆ ได้โดยเวอร์ชันใหม่ ๆ ก็จะมีเรื่องของการทำงาน Auto Model ที่จะรันหลาย ๆ โมเดลในรูปแบบอัตโนมัติ ตัวอย่างรูปภาพดังนี้



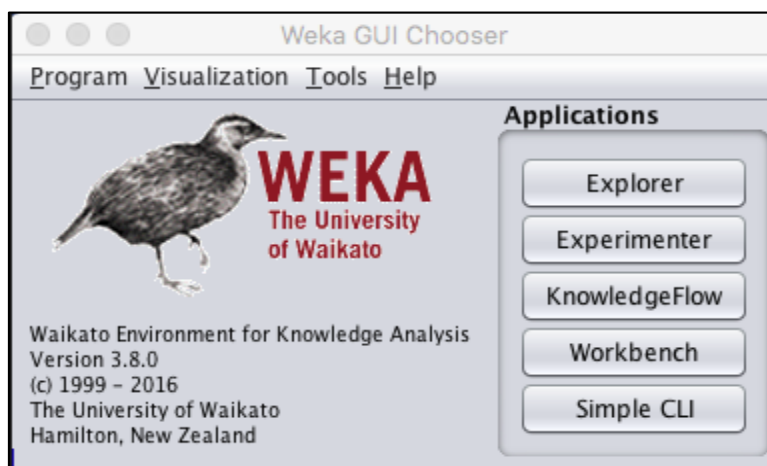
ภาพที่ 2.20 แสดงตัวอย่าง Auto Model ที่เจอรึสำหรับการทำ Machine Learning

โมดูลสอง : Rapid Miner Server เป็นโมดูลที่รองรับการทำงานของหลาย ๆ User ได้ ช่วยในเรื่องการตั้งเวลาให้ทำงาน (Scheduler) หรือสร้าง Web Service หรือ Web Application ได้ด้วย

โมดูลสาม : Rapid Miner Radoop เป็นโมดูลที่ใช้ในการจัดการข้อมูลที่มีขนาดใหญ่ ๆ แบบ Big Data โดยการทำงานจะไปประมวลผลบน Hadoop / Spark และไม่จำเป็นต้องเขียน Code

2.3.2 โปรแกรม Weka

โปรแกรม Weka (Waikato Environment for Knowledge Analysis) เริ่มพัฒนามาตั้งแต่ปี 1997 โดยมหาวิทยาลัย Waikato ประเทศนิวซีแลนด์ เป็นซอฟต์แวร์สำเร็จรูป อยู่ภายใต้การควบคุมของ GPL License โปรแกรม Weka ได้ถูกพัฒนามาจากภาษาจาวาทั้งหมด ซึ่งเขียนมาโดยเน้นกับงานทางด้านการเรียนรู้ด้วยเครื่อง (Machine Learning) และ การทำเหมืองข้อมูล (Data Mining) โปรแกรมจะประกอบไปด้วยโมดูลย่อย ๆ สำหรับใช้ในการจัดการข้อมูล และเป็นโปรแกรมที่สามารถใช้ Graphic User Interface (GUI) และใช้คำสั่งในการให้ซอฟต์แวร์ประมวลผล และสามารถรัน (Run) ได้หลายระบบปฏิบัติการ และสามารถพัฒนาต่อยอดโปรแกรมได้ เป็นเครื่องมือที่ใช้ทำงานในด้านการทำดาตาไมนนิ่งที่รวบรวมแนวคิดอัลกอริทึมมากมาย ซึ่งอัลกอริทึมสามารถเลือกใช้งานโดยตรงได้จาก 2 ทางคอกจากชุดเครื่องมือที่มีอัลกอริทึมมาให้ หรือเลือกใช้จากอัลกอริทึมที่ได้เขียนเป็นโปรแกรมลงไปเป็นชุดเครื่องมือเพิ่มเติม และชุดเครื่องมือมีฟังก์ชันสำหรับการทำงานร่วมกับข้อมูล ได้แก่ Pre-Processing, Classification, Regression, Clustering, Association rules, Selection และ Visualization ส่วนของการ Import ข้อมูล โดยข้อมูลนั้นจะต้องอยู่ในรูปแบบ ASCII อาจเป็น arff , csv , c45



ภาพที่ 2.21 แสดงตัวอย่างหน้าโปรแกรม Weka

1) กฎความสัมพันธ์ (Association Rule) แสดงความสัมพันธ์ของเหตุการณ์หรือวัตถุ ที่เกิดขึ้นพร้อมกัน ตัวอย่างของการประยุกต์ใช้กฎเชื่อมโยง เช่น การวิเคราะห์ข้อมูลการขายสินค้า โดยเก็บข้อมูลจากระบบ ณ จุดขาย (POS) หรือร้านค้าออนไลน์ แล้วพิจารณาสินค้าที่ผู้ซื้อมักจะซื้อพร้อมกัน เช่น ถ้าพบว่าคนที่ซื้อเทปวีดีโอ มักจะซื้อเทปกาวด้วย ร้านค้าก็อาจจะจัดร้านให้สินค้าสองอย่างอยู่ใกล้กัน เพื่อเพิ่มยอดขาย หรืออาจจะพบว่าหลังจากคนซื้อหนังสือ ก แล้ว มักจะซื้อหนังสือ ข ด้วย ก็สามารถนำความรู้นี้ไปแนะนำผู้ที่กำลังจะซื้อหนังสือ ก ได้

2) การจำแนกประเภทข้อมูล (Data Classification) หากกฎเพื่อระบุประเภทของวัตถุ จากคุณสมบัติของวัตถุ เช่น หาคความสัมพันธ์ระหว่างผลการตรวจร่างกายต่าง ๆ กับการเกิดโรค โดยใช้ข้อมูลผู้ป่วยและการวินิจฉัยของแพทย์ที่เก็บไว้ เพื่อนำมาช่วยวินิจฉัยโรคของผู้ป่วย หรือการวิจัยทางการแพทย์ ในทางธุรกิจจะใช้เพื่อดูคุณสมบัติของผู้ที่จะก่อหนี้ดีหรือหนี้เสีย เพื่อประกอบการพิจารณาการอนุมัติเงินกู้

3) การแบ่งกลุ่มข้อมูล (Data Clustering) แบ่งข้อมูลที่มีลักษณะคล้ายกันออกเป็นกลุ่ม แบ่งกลุ่มผู้ป่วยที่เป็นโรคเดียวกันตามลักษณะอาการ เพื่อนำไปใช้ประโยชน์ในการวิเคราะห์หาสาเหตุของโรค โดยพิจารณาจากผู้ป่วยที่มีอาการคล้ายคลึงกัน

4) จินตทัศน์ (Visualization) สร้างภาพคอมพิวเตอร์กราฟิกที่สามารถนำเสนอข้อมูลมากมายอย่างครบถ้วนแทนการใช้ข้อความนำเสนอข้อมูลที่มากมาย เราอาจพบข้อมูลที่ซ้อนเร้นเมื่อดูข้อมูลชุดนั้นด้วยจินตทัศน์

การทำเหมืองข้อมูล จำเป็นต้องอาศัยบุคลากรจากหลายฝ่าย และต้องอาศัยความรู้จำนวนมาก ถึงจะได้รับการประโยชน์อย่างแท้จริง เพราะสิ่งที่ได้จากขั้นตอนวิธีเป็นเพียงตัวเลข และข้อมูล ที่อาจจะนำไปใช้ประโยชน์ได้หรือใช้ประโยชน์อะไรไม่ได้เลยก็เป็นได้ ผู้ที่ศึกษาการทำเหมืองข้อมูลจึงควรมีความรู้รอบด้าน และต้องติดต่อกับทุก ๆ ฝ่าย เพื่อให้เข้าใจถึงขอบเขตของปัญหาโดยแท้จริงก่อน เพื่อให้การทำเหมืองข้อมูลเกิดประโยชน์อย่างแท้จริง

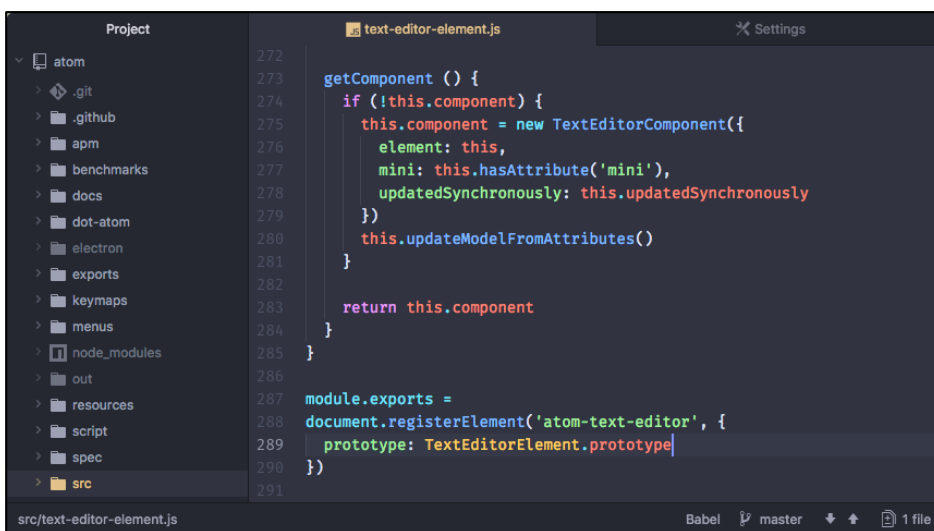
2.3.3 โปรแกรม Atom

Atom (โปรแกรม Text Editor แก้ไขข้อความ เขียนโค้ด ฟรี จาก GitHub) : สำหรับโปรแกรมนี้มีชื่อว่า โปรแกรม Atom เป็นโปรแกรมประเภทแก้ไขข้อความ (Text Editor) สามารถใช้ได้ทั้ง เขียนโค้ด รวมไปถึง แก้ไขส่วนต่าง ๆ ของโค้ดโปรแกรม หรือ แอปพลิเคชันต่าง ๆ ได้อย่างง่ายดาย ด้วยรูปแบบโปรแกรมที่เรียบง่าย พร้อมเครื่องมือ Tool สำหรับใช้งานที่หลากหลาย เหมาะสำหรับ

ผู้ที่เริ่มต้นเขียนโปรแกรมและมีอาชีพที่กำลังมองหาโปรแกรมที่ใช้งานง่าย ที่จะมาช่วยในการเขียนโค้ดง่ายและรวดเร็วมากขึ้น

โปรแกรม Atom ตัวนี้พัฒนามาจาก เว็บไซต์ชื่อดัง GitHub ที่ช่วยกันร่วมพัฒนาและเปิดให้ใช้งานได้ฟรี ๆ คุณสามารถใช้งานโปรแกรมแก้ไขเขียนโค้ด Atom เปิดใช้งาน Project ของคุณ ช่วยให้คุณดูงานเขียนโค้ดของคุณได้ง่ายมากยิ่งขึ้น มีเครื่องมือให้ใช้งาน ค้นหา คัดลอก และ แสดงตัวอย่าง หลังจากเขียนเสร็จแล้วก็สามารถ Preview ดูตัวอย่างของโค้ดที่เขียนได้เลขทันที พร้อมแสดงผลและแก้ไขส่วนข้อผิดพลาดต่าง ๆ ได้ง่ายมากยิ่งขึ้น

นอกจากนี้ โปรแกรม Atom ตัวนี้ก็ยังมีฟังก์ชัน Smart auto completion ช่วยแสดงและเติมส่วนต่าง ๆ ของโค้ดคุณ ทำให้คุณเขียนโค้ดได้รวดเร็วมากยิ่งขึ้น พร้อมแบ่งบรรทัดการเขียน เติมแต่งสีให้เห็นชัดเจน หากผู้ใช้งานคนไหนชื่นชอบหน้าต่างโปรแกรมแบบอื่น ๆ ก็สามารถหา Theme มาติดตั้งเองจากเว็บไซต์ GitHub ได้อีกด้วย โปรแกรมนี้ รองรับการใช้งานข้ามแพลตฟอร์ม (Cross Platform) ไม่ว่าจะเป็น บนระบบปฏิบัติการ Windows Linux และ MacOS



```

Project
├── atom
├── .git
├── .github
├── apm
├── benchmarks
├── docs
├── dot-atom
├── electron
├── exports
├── keymaps
├── menus
├── node_modules
├── out
├── resources
├── script
├── spec
└── src

src/text-editor-element.js
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291

text-editor-element.js
getComponent () {
  if (!this.component) {
    this.component = new TextEditorComponent({
      element: this,
      mini: this.hasAttribute('mini'),
      updatedSynchronously: this.updatedSynchronously
    })
    this.updateModelFromAttributes()
  }
  return this.component
}

module.exports =
document.registerElement('atom-text-editor', {
  prototype: TextEditorElement.prototype
})

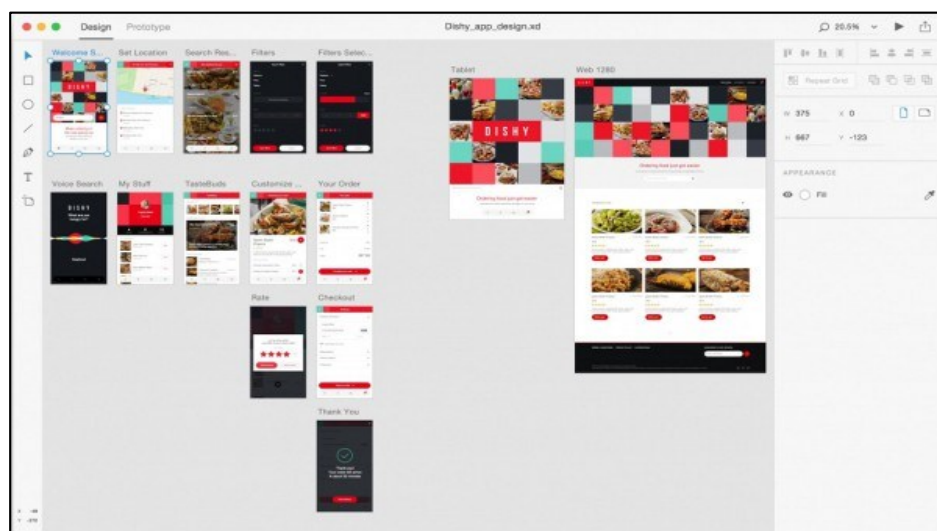
```

ภาพที่ 2.22 แสดงตัวอย่างหน้าโปรแกรม Atom

2.3.4 โปรแกรม Adobe XD

Adobe XD คือ แอปพลิเคชันบนคอมพิวเตอร์ ที่ช่วยให้นักออกแบบเริ่มทำงานตั้งแต่การสร้างไวร์เฟรม ไปจนถึงการสร้างตัวตนแบบเพื่อทดลองใช้เพื่อเก็บข้อมูลกลับมาปรับปรุงการใช้งานให้ดีขึ้น เพื่อที่ผู้จะใช้จะได้กลับมาใช้งานอีกมากขึ้น เวอร์ชันฟรีวิวแรกที่ทำให้ทดลองใช้ครั้งนี้ โหมดออกแบบ (Design) สามารถทำงานออกแบบได้อย่างง่ายดายตาย โหมดต้นแบบ (Prototype)

ก็สามารถกำหนด Hotspots และ Transitions ได้ตามต้องการ โหมดพรีวิว (Desktop Preview) สำหรับทดลองใช้งานเพื่อดูความเปลี่ยนแปลงจริงได้ นอกจากนี้ ยังสามารถแชร์พรีวิวให้ผู้ตัดสินใจและคนอื่นได้ทดลองใช้งานได้ ในบราวเซอร์บนคอมพิวเตอร์หรือบนอุปกรณ์พกพาอีกด้วย นอกจากนี้ นักออกแบบยังสามารถนำงานที่สร้างขึ้นในแอปพลิเคชันสำหรับการออกแบบอย่าง Illustrator CC และ Photoshop CC มาใช้งานร่วมกันได้ และเมื่อการออกแบบและทดลองใช้มาถึงขั้นสุดท้ายแล้ว ทุกอย่างก็สามารถเอ็กซ์พอร์ตไปให้นักพัฒนาทำงานต่อไปได้อย่างง่ายดาย



ภาพที่ 2.23 แสดงตัวอย่างหน้าโปรแกรม Adobe XD

2.4 วรรณกรรมที่เกี่ยวข้อง

อนุชิต พงศ์พรหม, พลกฤต แสงอาวูช และพูนฉัตร วิชัยดิษฐ์ (2562) การศึกษาการบริหารจัดการทุนการศึกษาของนักศึกษาระดับมหาวิทยาลัยของรัฐจังหวัดสุราษฎร์ธานีครั้งนี้มีจุดประสงค์เพื่อเปรียบเทียบ และศึกษาแนวทางในการบริหารจัดการทุนการศึกษาดังกล่าว โดยจำแนกตามปัจจัยส่วนบุคคล เนื่องจากยังมีนักศึกษาจำนวนมากที่ต้องการ และมีคุณสมบัติที่ตรงกับวัตถุประสงค์การให้ทุนการศึกษา แต่ยังไม่ได้รับการดูแลช่วยเหลือ ซึ่งกลุ่มตัวอย่างในงานวิจัยครั้งนี้เป็นกลุ่มตัวอย่างจากมหาวิทยาลัยราชภัฏสุราษฎร์ธานี และมหาวิทยาลัยสงขลานครินทร์วิทยาเขต สุราษฎร์ธานี เนื่องจากทั้ง 2 มหาวิทยาลัยเป็นองค์กรที่รับนโยบายการลดความเหลื่อมล้ำทางการศึกษา ซึ่งเป้าหมายหลักในการขับเคลื่อนมหาวิทยาลัย เพื่อผลิตบัณฑิตที่มีคุณลักษณะบัณฑิตที่พึงประสงค์สามารถเรียนรู้คิดเป็นทำเป็นและมีศักยภาพ

จากการศึกษาบทความดังกล่าว ทางผู้จัดทำเห็นว่า ทางผู้จัดทำจะนำแนวคิดและแนวทางในการบริหารจัดการทุนการศึกษา มาใช้เป็นแนวทางในการวิเคราะห์ และสร้างจำลองมาตรการช่วยเหลือด้านการศึกษา เพื่อเป็นส่วนช่วยในการตัดสินใจในการมอบทุนการศึกษาอย่างเป็นระบบ

สุมาลี ลิกเสน และมนตชัย เทียนทอง (2559)งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความถนัดของผู้เรียน ตามแนวทางพหุปัญญา ของ Howard Gardner ที่จำแนกความถนัดของผู้เรียนออกเป็น 3 กลุ่ม กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ คือ นักศึกษามหาวิทยาลัยราชภัฏนครปฐม ในภาคเรียนที่ 2/2556 จำนวน 431คน ทำการวิเคราะห์ข้อมูลด้วยโปรแกรม WEKA โดยใช้กฎการจำแนกเทคนิคต้นไม้ตัดสินใจ ภายใต้อัลกอริทึม j48 สร้างและทดสอบตัวแบบด้วยวิธี 10 fold Cross validation ผลการวิจัยพบว่า ตัวแปรต้นที่ส่งผลต่อความถนัดของผู้เรียน ตามแนวทางพหุปัญญา ประกอบด้วย สาขาวิชาเกรดเฉลี่ยวิชาที่ทำคะแนนได้สูงอยู่เสมอ ลักษณะการเรียนที่ชอบวิธีการเรียนที่มีความถนัดกิจกรรมที่ชอบและงานอดิเรกที่ชอบ ได้กฎการจำแนกความถนัดของผู้เรียนทั้งสิ้นจำนวน 54 กฎดังนี้ ความถนัดของผู้เรียนกลุ่มการวิเคราะห์ (Analytic) จำนวน 19 กฎ ความถนัดของผู้เรียนกลุ่มพินิจพิจารณา (Introspective) จำนวน 14 กฎ และความถนัดของผู้เรียนกลุ่มปฏิสัมพันธ์ (Interactive) จำนวน 21 กฎ ซึ่งมีค่าความถูกต้องคิดเป็นร้อยละ 85.90 ค่าความแม่นยำคิดเป็นร้อยละ 86.03 ค่าความระลึกคิดเป็นร้อยละ 85.56 และค่าถ่วงดุลคิดเป็นร้อยละ 85.89 โดยผลที่ได้จากงานวิจัยนี้ สามารถนำไปใช้ในการพัฒนาระบบ การวิเคราะห์ความถนัดของผู้เรียนตามแนวทางพหุปัญญาของ Howard Gardner โดยใช้กฎการจำแนกเทคนิคต้นไม้ตัดสินใจที่มีความถูกต้องและแม่นยำในการทำนายความถนัดของผู้เรียนได้สำหรับแบ่งกลุ่มเพื่อจัดทำระบบโครงการออนไลน์ต่อไป

จากการศึกษาบทความดังกล่าว ทางผู้จัดทำเห็นว่า ทางผู้จัดทำจะนำกฎการจำแนกด้วยเทคนิคต้นไม้ตัดสินใจมาใช้ในการสร้างแบบจำลองเพื่อเป็นมาตรการช่วยเหลือด้านการศึกษา สำหรับนักศึกษา โดยทำการวิเคราะห์ข้อมูลด้วยโปรแกรม Weka

จริยา ศรีวิจารณ์, อรัญญา ต้อยคำภีร์ และณัฐสุดา เต็มพันธ์ (2556) การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาประสบการณ์ด้านจิตใจของนักศึกษาที่ยากจนจากชนบทและได้รับทุนการศึกษา ผู้ให้ข้อมูลเป็นนักศึกษาปริญญาตรี จำนวน 10 ราย คัดเลือกแบบเจาะจงตามเกณฑ์

ที่กำหนด เก็บข้อมูลโดยการสัมภาษณ์รายบุคคล วิเคราะห์ข้อมูลโดยวิธีการวิจัยเชิงคุณภาพแบบ
 ทีมผู้วิจัยเห็นชอบร่วมกันผลการศึกษาพบ 5 ประเด็นหลัก คือ (1) การเผชิญความยากลำบากใน
 ชีวิต ได้แก่ ความเป็นอยู่ที่ขาดแคลน การไม่ยอมจำนนต่อความยากจน การมีความกดดันและรู้สึก
 ต้อยค่าจากความยากจน (2) การมีความรู้สึกสุข ระคนทุกข์เมื่อได้ทุนการศึกษา ได้แก่ ความรู้สึก
 ภาคภูมิใจร่วมกับความรู้สึกซาบซึ้งใจที่ได้รับทุนและความรู้สึกไม่มั่นใจบางขณะ (3) การปรับตัวใน
 สังคมใหม่ ได้แก่ การใช้ชีวิตในมหาวิทยาลัย การดำรงอยู่ภายใต้สิ่งแวดล้อมวัฒนธรรม และการใช้
 ชีวิตต่างถิ่นเพียงลำพัง (4) การมีทัศนคติดีและกตัญญูสำคัญในการปรับตัว ได้แก่ การมีจิตใจที่
 แข็งแกร่ง การมีสัมพันธภาพและการมีแหล่งเกื้อหนุนที่ดี และ (5) การมีจุดมุ่งหมายในชีวิต ได้แก่
 การทำสิ่งต่าง ๆ เพื่อครอบครัว เพื่อความมั่นคงของตัวเอง และเพื่อช่วยเหลือผู้อื่น จะเห็นว่าปัจจัย
 หนึ่งที่สำคัญ ในการที่ทำให้นักศึกษาต่อสู้กับความลำบากในชีวิตจากความยากจน คือ ทัศนคติด้าน
 บวกและการมีจุดมุ่งหมาย ในชีวิตของพวกเขา

จากการศึกษาบทความดังกล่าว ทางผู้จัดทำเห็นว่า ทางผู้จัดทำจะนำแนวคิดและแนวทาง
 ในการบริหารจัดการทุนการศึกษา มาใช้เป็นแนวทางในการวิเคราะห์ และสร้างจำลองมาตรการ
 การช่วยเหลือด้านการศึกษา เพื่อเป็นส่วนช่วยในการตัดสินใจในการมอบทุนการศึกษาอย่างเป็น
 ระบบ

อภิยศ เจริญวิวัฒน์ (2562) สำนักหอสมุด มหาวิทยาลัยเกษตรศาสตร์ มีแนวคิดในการ
 ประยุกต์ใช้ข้อมูลทางสถิติเพื่อเพิ่มประสิทธิภาพในการให้บริการแก่ผู้ใช้ โดยทำการศึกษากับตัวเลข
 ของการให้บริการในด้านต่าง ๆ ที่กระจุกกระจายอยู่เป็นจำนวนมาก มีปริมาณข้อมูลสูง และยังไม่
 เคยผ่านการวิเคราะห์ทางด้านสถิติมาก่อน โดยการวิเคราะห์ข้อมูลนี้ จะวิเคราะห์ข้อมูลที่มี
 หลากหลาย ทั้งจากการเข้าใช้งานอาคารของสำนักหอสมุด และการใช้บริการพื้นฐานและบริการ
 พิเศษต่าง ๆ โดยใช้วิธีวิเคราะห์ข้อมูลแสดงความสัมพันธ์ และนำมาแสดงผลในรูปแบบของ Data
 visualization โดยใช้ภาพ ตัวเลข แผนภูมิ กราฟ และอื่น ๆ ซึ่งช่วยเพิ่มความพึงพอใจแก่ผู้บริหารให้
 สามารถรับรู้ข้อมูลที่มากขึ้น และสามารถนำข้อมูลที่ได้ไปใช้ในการสนับสนุนการตัดสินใจ

จากการศึกษาบทความดังกล่าว ทางผู้จัดทำเห็นว่า ทางผู้จัดทำจะนำวิธีวิเคราะห์ข้อมูลแสดง
 ความสัมพันธ์ โดยการทำ Data visualization เพื่อเผยแพร่ผลวิเคราะห์ ขั้นตอนการนำผลลัพธ์ที่ได้
 ไปใช้งานเป็นการทั่วไป อาจจัดทำเป็นรูปแบบของรายงาน (Report) หรือแผนภาพ (Dashboard) ที่
 พร้อมให้ฝ่ายต่าง ๆ นำไปใช้ประโยชน์ในการวางแผน กำหนดกลยุทธ์ และดำเนินการต่าง ๆ

การวิจัยครั้งนี้มีวัตถุประสงค์ 1) เพื่อเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจ รุ่นที่ 3 ที่ใช้เทคนิคการแบ่งข้อมูลที่แตกต่างกัน และ 2) เพื่อทดสอบประสิทธิภาพของเทคนิคการแบ่งข้อมูลชนิดแบบไม่มีผู้สอน ชุดข้อมูลที่ใช้ในการทดลอง คือ ชุดข้อมูล KDD Cup 1999 ประสิทธิภาพ ของการจำแนกข้อมูลวัดจากอัตราการจำแนกข้อมูลถูกต้อง และอัตราการจำแนกข้อมูลผิดพลาดโดยใช้โปรแกรมเวก้า และแมทแลบในการประมวลผล จากการเฉลี่ยผลการทดลองจำนวน 10 รอบ โดยใช้เทคนิคในการแบ่งข้อมูลของชุดข้อมูลฝึกสอนให้อยู่ในรูปแบบไม้ต่อเนื่อง และจำแนกข้อมูลด้วยต้นไม้ตัดสินใจรุ่นที่ 3 พบว่าการแบ่งข้อมูลด้วยขนาดความถี่เท่ากันจำนวน 10 ชั้น มีอัตราการจำแนกข้อมูลถูกต้องสูงที่สุด เท่ากับ 99.79% รองลงมาคือ การแบ่งข้อมูลด้วยการจัดกลุ่มเคมีนส์ จำนวน 40 กลุ่ม เท่ากับ 99.75% และน้อยที่สุด คือ การแบ่งข้อมูลด้วยขนาดความถี่เท่ากัน จำนวน 20 ชั้น เท่ากับ 99.57% เมื่อเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลด้วยต้นไม้ตัดสินใจ รุ่นที่ 3 โดยนำกฎที่สร้างขึ้นจากชุดข้อมูลฝึกสอนไปใช้กับชุดข้อมูลทดสอบพบว่าประสิทธิภาพของการใช้เทคนิคแบ่งข้อมูลด้วยขนาดความถี่เท่ากันและจำแนกข้อมูลด้วยต้นไม้ตัดสินใจรุ่นที่ 3 มีประสิทธิภาพสูงที่สุด โดยมีอัตราการจำแนกข้อมูลถูกต้อง เท่ากับ 92.30% มีอัตราการจำแนกข้อมูลผิดพลาด เท่ากับ 4.89% และสามารถจำแนกข้อมูลที่ไม่สามารถจัดเข้ากลุ่มได้ เท่ากับ 2.81% ผลการทดลองยังแสดงให้เห็นว่าเทคนิคการแบ่งข้อมูลชนิดแบบไม่มีผู้สอน แต่ละชนิดส่งผลต่อประสิทธิภาพการจำแนกข้อมูลของต้นไม้ตัดสินใจ รุ่นที่ 3 และพบว่าการใช้เทคนิคการแบ่งข้อมูลร่วมกับเทคนิคต้นไม้ตัดสินใจ รุ่นที่ 3 สามารถช่วยกรองข้อมูลที่ไม่อยู่ในกฎที่สร้างขึ้นได้ ซึ่งเป็นสิ่งที่แตกต่างกับเทคนิคต้นไม้ตัดสินใจ รุ่นที่ C4.5 ฉะนั้นวิธีที่นำเสนอหากเลือกนำไปใช้ได้อย่างเหมาะสมจะเกิดประโยชน์อย่างมาก

จากการศึกษาบทความดังกล่าว ทางผู้จัดทำเห็นว่าจะนำวิจัยครั้งนี้มาทดสอบ เปรียบเทียบประสิทธิภาพของข้อมูลนักศึกษาด้วยเทคนิคต้นไม้ตัดสินใจ โดยการใช้เทคนิคการแบ่งข้อมูลเป็น 2 ส่วนส่วนแรก คือ การนำข้อมูลมาใช้จริง (Training Data) 70% และส่วนที่สอง คือ การนำข้อมูลมาทดสอบ (Testing Data) 30% เพื่อหาอัลกอริทึมที่เหมาะสมและให้ค่าความแม่นยำมากที่สุด ซึ่งวัดค่าความยำโดยใช้ Confusion Matrix ในโปรแกรม Weka

ดังนั้น วรรณกรรมที่เกี่ยวข้องทั้งหลายที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น ผู้จัดทำโครงการได้เล็งเห็นความสำคัญของการนำเทคโนโลยีมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลนักศึกษาเบื้องต้น เพื่อใช้สำหรับเป็นมาตรการช่วยเหลือในด้านของการศึกษา และเผยแพร่ข้อมูลบนเว็บไซต์บน Web

Browser เนื่องจากเทคโนโลยีปัจจุบันเข้ามาช่วยในการจัดการข้อมูล หรือดำเนินงานต่าง ๆ ให้ความสะดวกสบายเป็นอย่างมาก อีกทั้งผู้ใช้งานยังเข้าถึงข้อมูลได้ง่าย และมีประสิทธิภาพ

2.5 บทสรุป

จากแนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้องที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น ผู้วิเคราะห์ข้อมูลได้เลือกใช้ขั้นตอนกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบ Classification ด้วยการสร้างโมเดล Decision Tree เพื่อจัดกลุ่มข้อมูลที่มีความสัมพันธ์กัน โดยใช้โปรแกรมที่ใช้ทำเหมืองข้อมูล คือ โปรแกรม SQL Server 2014 Management Studio & SQL Server Data Tools for Visual Studio 2013, โปรแกรม Weka 3.8.4 และโปรแกรม Rapid Miner Studio 9.5.1 ในการสร้างโมเดล Decision Tree เพื่อทำการเปรียบเทียบและเลือกผลลัพธ์ที่ดีที่สุดมานำเสนอ จากนั้นนำข้อมูลสารสนเทศมาทำการแสดงผลแบบ Visualization ในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public เผยแพร่บน Web Browser ที่เป็นที่ยอมรับในยุคอินเทอร์เน็ตคือการเผยแพร่ทางสื่อออนไลน์ โดยใช้ภาษา HTML และ CSS ในการเขียนเว็บไซต์ขึ้นมา