บทที่ 2

กรอบแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาวิจัยเรื่อง การนำทฤษฎีเหมืองข้อมูลมาประยุกต์ใช้ในบริการยืมคืนของห้องสมุด กรณีศึกษา สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ ผู้วิจัยได้ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องเพื่อใช้เป็นแนวทางการวิจัย ดังนี้

2.1 การจัดการห้องสมุด และทฤษฎีที่เกี่ยวข้อง

2.1.1 การจัดการห้องสมุด (Library Management)

การจัดการห้องสมุด (Library Management) คือชุดของหน้าที่ต่างๆ (A Set of Functions) ที่กำหนดทิศทางในการใช้ทรัพยากรทั้งหลายอย่างมีประสิทธิภาพ (Efficient) หมายถึงการใช้ทรัพยากรของห้องสมุด ได้อย่างเฉลียวฉลาดและคุ้มค่า (Cost-Effective) และประสิทธิผล (Effective) หมายถึงการตัดสินใจได้อย่างถูกต้อง (Right Decision) อีกทั้งมีการปฏิบัติการสำเร็จตามแผนที่ได้กำหนดด้วย ทั้งนี้เพื่อให้บรรลุถึงเป้าหมายของห้องสมุด (Gibbons, 2001) โดยแบ่งส่วนประกอบทางการจัดการ ดังนี้ (Stueart, 1987)

2.1.1.1 การวางแผน (Planning) เป็นงานหน้าที่แรกทางการจัดการที่มีความสำคัญของผู้บริหารห้องสมุดที่ต้อง ศึกษา วิเคราะห์ และตัดสินใจกำหนดเป้าหมายที่ต้องการหาแนวทางและสร้างวิธีการปฏิบัติให้งานสำเร็จตามเป้าหมายนั้นๆ ได้อย่างมีประสิทธิภาพภายใต้ระยะเวลาและข้อจำกัดของแต่ละงาน

2.1.1.2 การจัดองค์การ (Organizing) เป็นการจัดโครงสร้าง การจัดระบบงาน การรวมกลุ่มกิจกรรม การกำหนดขอบเขตอำนาจหน้าที่ความรับผิดชอบ และการมอบหมายให้แก่ผู้ใต้บังคับบัญชาแต่ละคน เพื่อที่จะปฏิบัติงานให้สอดคล้องกันและบรรลุเป้าหมายที่ผู้จัดการกำหนด

2.1.1.3 การจัดบุคคลเข้าทำงาน (Staffing) เป็นการกำหนดความต้องการ สรรหา คัดเลือก บรรจุ ฝึกอบรม พัฒนา ธำรงรักษา และจูงใจบุคลากรภายใต้ภายใต้การบังคับบัญชาให้ปฏิบัติงานได้ตามที่ต้องการ

2.1.1.4 การนำ (Directing) คือการที่ผู้บริหารห้องสมุดใช้ความสามารถและอิทธิพลในการชักจูงให้ผู้ตาม (Follower) แสดงพฤติกรรมหรือปฏิบัติงานอย่างเต็มความสามารถและด้วยความเต็มใจ เพื่อให้บรรลุเป้าหมายที่ผู้บริหารห้องสมุดกำหนด

2.1.1.5 การควบคุม (Controlling) เป็นกระบวนการในการติดตาม ตรวจสอบ เปรียบเทียบผลการปฏิบัติงานจริงกับเป้าหมายที่วางไว้ตามเกณฑ์

2.1.2 กระบวนการค้นหาความรู้ในฐานข้อมูล (Knowledge Discovery in Database)

กระบวนการค้นหาความรู้ในฐานข้อมูล (Knowledge Discovery in Database หรือ KDD) คือกระบวนการในการค้นหาลักษณะแฝงของข้อมูลที่อยู่ในกลุ่มข้อมูลจำนวนมาก เพื่อช่วยในการบริหารงาน (Fayyad & Stolorz, 1997) โดยกระบวนการค้นหาความรู้ในฐานข้อมูลแบ่งออกเป็น 4 ขั้นตอน ดังนี้

1. Data Cleaning เป็นขั้นตอนการทำความสะอาดข้อมูล โดยอาจมีการเติมค่าที่สูญหายหรือจัดการกับสิ่งแปลกปลอม (Noise) ที่อยู่ในข้อมูล ตลอดจนกำจัดข้อมูลนอกขอ บเขต (Outlier) และความไม่คงที่ของข้อมูล หากไม่มีการทำความสะอาดข้อมูลที่ดี ผลที่ได้จากการค้นหาความรู้ในฐานข้อมูลจะไม่น่าเชื่อถือ อีกทั้งยังอาจทำให้การทำเหมืองข้อมูลมีความยุ่งยาก
2. Data Integration เป็นการรวบรวมข้อมูลจากหลาย ๆ แหล่งหรือหลาย ๆ ฐานข้อมูล
3. Data Transformation เป็นขั้นตอนการเปลี่ยนแปลงรูปแบบข้อมูล เพื่อให้มีข้อมูลที่จะนำมาใช้ในการวิเคราะห์มากขึ้น
4. Data Reduction เนื่องจากข้อมูลที่นำมาวิเคราะห์อาจมีขนาดใหญ่มาก ซึ่งหากนำข้อมูลทั้งหมดไปทำเหมืองข้อมูลจะยุ่งยาก ซับซ้อน และเสียเวลานาน ดังนั้นจึงจำเป็นต้อง มีการลดขนาดของข้อมูลลง โดยการลดจำนวนตัวแปรที่ไม่เกี่ยวข้องออกไป เพื่อให้สามารถนำข้อมูลไปวิเคราะห์ได้อย่างมีประสิทธิภาพมากขึ้น

2.1.3 เทคนิคการทำเหมืองข้อมูล (Data Mining Technique)

การทำเหมืองข้อมูลเกิดจากการวิวัฒนาการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการเก็บข้อมูลอย่างง่ายมาสู่การจัดเก็บในรูปแบบของฐานข้อมูลที่ซับซ้อนขึ้น อีกทั้งสามารถดึง

สารนิเทศของข้อมูลมาใช้งาน จนถึงการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ภายในข้อมูล ดังแสดงในภาพที่ 2.1

ภาพที่ 2.1
วิวัฒนาการจัดเก็บและตีความหมายข้อมูล

ที่มา : (Han, 2007)

วิวัฒนาการจัดเก็บและตีความหมายข้อมูลทั้ง 6 ยุค มีรายละเอียดดังต่อไปนี้

1. ยุคการเก็บข้อมูล (Data Collection) และการสร้างฐานข้อมูล (Database Creation)

   โดยเกิดในช่วงก่อน ค.ศ. 1960 ซึ่งเป็นการนำข้อมูลมาจัดเก็บอย่างเหมาะสมในอุปกรณ์ที่น่าเชื่อถือ และป้องกันการสูญหาย ได้เป็นอย่างดี

2. ยุคการจัดการฐานข้อมูล (Database Management System)

   โดยเกิดในช่วง ค.ศ. 1970 ถึง ค.ศ. 1980 (ตอนต้น) ซึ่งเป็นการนำข้อมูลที่จัดเก็บมาสร้างความสัมพันธ์ต่อกันในข้อมูลเพื่อประโยชน์ในการนำไปวิเคราะห์ และการตัดสินใจอย่างมีคุณภาพ

3. ยุคระบบจัดการฐานข้อมูลขั้นสูง (Advanced Database Systems)

   โดยเกิดในช่วง ค.ศ. 1980 (ตอนกลาง) ถึงปัจจุบัน

4. ยุคการวิเคราะห์ข้อมูลขั้นสูง ยุคทำคลังข้อมูล และยุค ทำเหมืองข้อมูล (Advanced Data Analysis: Data Warehousing and Data Mining)

   โดยเกิดในช่วง ค.ศ. 1980 (ตอนปลาย) ถึงปัจจุบัน ซึ่งเป็นการรวบรวมข้อมูลมาจัดเก็บลงไปในฐานข้อมูลขนาดใหญ่โดยครอบคลุม ทุกแง่มุมขององค์กร เพื่อนำข้อมูลที่ได้มาประมวลผล โดยการสร้างแบบจำลองเมืองข้อมูล และความสัมพันธ์ทางสถิติเพื่อช่วยสนับสนุนการตัดสินใจ

5. ยุคฐานข้อมูลเชิงเว็บ (Web-Based Databases)

   โดยเกิดในช่วง ค.ศ. 1990 ถึงปัจจุบัน ซึ่งมาช่วยเสริมการทำงานภายในเว็บไซต์ อีกทั้งยังเป็นการทำงานกับข้อมูลโดยตรง โดยสามารถให้ร ายละเอียดของข้อมูลและการเปลี่ยนแปลงข้อมูลบนเว็บไซต์ได้อย่างชัดเจน

6. ยุคใหม่ของการบูรณาการข้อมูลและระบบสารสนเทศ (New Generation of Integrated Data and Information Systems)

   โดยเกิดในช่วงอนาคต ซึ่งเป็นการจัดเก็บข้อมูลที่มีลักษณะเดียวกันให้รวมเข้าอยู่ด้วยกัน โดยมีการลดความซ้ำซ้อนของข้อมูลบางส่วนออกไป

มีผู้ให้คำนิยามการทำเหมืองข้อมูลไว้หลายความหมาย ดังนี้

การทำเหมืองข้อมูล คือ การนำเทคนิคจากศาสตร์ด้านการเรียนรู้เครื่องกล (Machine Learning) การรู้จำรูปแบบ ข้อมูล (Pattern Recognition) หลักสถิติ (Statistics) และฐานข้อมูล (Databases) มาใช้ในการดึงข้อมูลข่าวสารจากฐานข้อมูลที่มีอยู่ (Peter Cabena, 1998)

การทำเหมืองข้อมูล คือ การวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์และสรุปผลจากข้อมูลจำนวนมากเพื่อทำความเข้าใจและให้เกิดประโยชน์ต่อเจ้าของธุรกิจ (David Hand, 2001)

การทำเหมืองข้อมูล คือ กระบวนการ คัดเลือกและสำรวจข้อมูล ตลอดจน เป็นการสร้างแบบจำลองของข้อมูลเพื่อค้นหารูปแบบและค้นหาความสัมพันธ์จากข้อมูลจำนวนมากเพื่อให้ได้ผลลัพธ์ที่เป็นประโยชน์ (Giudici, 2003)

การทำเหมืองข้อมูล คือ กระบวนการ ทำงานโดยอัตโนมัติที่อาศัยความสามารถของคอมพิวเตอร์ในการเรียนรู้ วิเคราะห์และขุดค้นข้อมูลจากฐานข้อมูล เพื่อค้นหาแนวโน้มและรูปแบบของข้อมูลซึ่งสามารถนำมาใช้ในการทำนายแนวโน้มพฤติกรรมต่างๆ ของลูกค้า ตลอดจนสกัดแก่นความรู้ที่มีอยู่ในสารสนเทศนั้นได้ (Geatz, 2003)

เหมืองข้อมูลเป็นเทคโนโลยีที่ช่วยให้การวิเคราะห์ข้อมูลทำได้โดยอัตโนมัติและมีประสิทธิภาพสูง ซึ่งส่งผลให้มีการนำเหมืองข้อมูลไปใช้อย่างแพร่หลายในทุกวงการ โดยเฉพาะในกรณีที่ข้อมูลมีขนาดใหญ่มาก โดยระบบคอมพิวเตอร์จะทำหน้าที่ค้นหาแนวโน้มและลักษณะที่น่าสนใจต่างๆ ที่ปรากฏในข้อมูล ตลอดจนวิเคราะห์ความสัมพันธ์ระหว่างข้อมูลนั้น ซึ่งจะช่วยให้สามารถพยากรณ์แนวโน้ มของข้อมูลใหม่ที่จะเกิดขึ้นในอนาคตได้ รวมถึงเข้าใจความสัมพันธ์ที่เชื่อมโยงข้อมูลแต่ละกลุ่มเข้าด้วยกัน (Nicholson, 2003)

โดยมีวิธีสำหรับการทำเหมืองข้อมูล (Algorithms for Data Mining) ดังนี้

1. Classification (การแบ่งประเภท)

Classification ก็คือการแบ่งประเภ ท โดยจำเป็นต้องมีตัวแปรเป้าหมายในการแบ่งประเภท เช่น ถ้าแบ่งประเภทของนักเรียนตามแผนการเรียนต่อแล้ว อาจแบ่งได้เป็น 2 ประเภท คือ นักเรียนที่เรียนต่อ และนักเรียนที่ไม่เรียนต่อ สำหรับขั้นตอนวิธีที่จัดอยู่ในประเภทของ Classification ได้แก่ Naïve Bayes, ต้นไม้แห่งการตัดสินใจ (Decision Tree) และเครือข่าย

ประสาทเทียม (Neural Network) โดยที่ขั้นตอนวิธี Naïve Bayes นั้นจะช่วยให้สามารถสร้างแบบจำลองในการพยากรณ์ข้อมูลได้อย่างรวดเร็วและยังสามารถนำไปใช้สร้างตัวแบบอื่น ๆ ที่ทำให้เข้าใจข้อมูลได้มากยิ่งขึ้น แต่การวิเคราะห์ด้วยขั้นตอนวิธี Naïve Bayes นั้น สามารถพิจารณาได้ทีละตัวแปรเท่านั้น ไม่สามารถนำตัวแปรมากกว่า 1 ตัวแปร มาพิจารณาร่วมกัน เช่น ไม่สามารถบอกได้ว่าถ้าเป็นบริษัทขนาดเล็กที่มีกำไรมากกว่า 500,000 ดอลล่าห์ แล้วจะเป็นบริษัทที่มีความเสี่ยงสูงหรือไม่ อีกทั้งไม่สามารถบอกได้ว่าถ้าเป็นบริษัทขนาดใหญ่ที่ไม่มีกำไรแล้วจะเป็นบริษัทที่มีความเสี่ยงต่ำหรือไม่ (Tang, 2005)

2. Clustering (การจัดกลุ่ม)

บางครั้งเรียก Clustering ในอีกชื่อหนึ่งว่า Segmentation โดยเป็นชุดข้อมูลที่นำข้อมูลที่มีความคล้ายคลึงมารวมอยู่ด้วยกัน และชุดข้อมูลนี้ก็มีความแตกต่างจากข้อมูลที่อยู่ในกลุ่มอื่น Clustering นั้นต่างจาก Classification ตรงที่ว่าการแบ่งกลุ่มของ Clustering นั้นไม่มีตัวแปรเป้าหมายที่ใช้ในการแบ่งกลุ่ม โดยการทำ Clustering นั้นจะพิจารณาตัวแปรทุกตัวที่มีอยู่ จากนั้นจึงนำข้อมูลที่มีค่าในตัวแปรต่าง ๆ ที่ใกล้เคียงกันรวมกลุ่มกัน ซึ่งตัวแปรต่าง ๆ ในที่นี้ เช่น ข้อมูลของลูกค้าชุดหนึ่งประกอบด้วยสองตัวแปร คืออายุ และรายได้ ดังนั้นเมื่อทำ Clustering แล้วก็จะแบ่งกลุ่มด้วยตัวแปรโ ดยที่ผู้ใช้ไม่ต้องระบุตัวแปรที่ใช้ในการแบ่งกลุ่ม ในขณะที่การทำ Classification นั้นจะต้องมีการระบุตัวแปรเป้าหมายที่ใช้ในการแบ่งประเภทเช่น ในอายุเป็นเป้าหมายก็จะแบ่งประเภทลูกค้าตามอายุเป็นต้น

การทำ Clustering นี้จะทำให้ทราบพฤติกรรมของแต่ละกลุ่ม ถ้าหากใช้ในกา ร แบ่งกลุ่มลูกค้าแล้วก็สามารถทราบพฤติกรรมของลูกค้าว่าแต่ละกลุ่มมีลักษณะอย่างไร และปรับปรุงบริการให้สอดคล้องกับความต้องการของลูกค้าแต่ละกลุ่มได้ (Tang, 2005)

3. Association (การสัมพันธ์กัน)

Association เป็นงานที่นิยมทำกันมาก บางครั้งเรียกว่า Market Basket Analysis หรือ Affinity Analysis เป็นการศึกษาว่าตัวแปร หรือข้อมูลใดมีความสัมพันธ์กัน โดยวิเคราะห์จากข้อมูลการขายสินค้าว่ามีสินค้าชนิดใดที่มักถูกซื้อไปด้วยกัน นอกจากนี้การวิเคราะห์ด้วย Association ก็เพื่อหากฎความสัมพันธ์ (Association Rule) โดยกฎความสัมพันธ์จะอยู่ในรูป “ถ้าเหตุการณ์นี้เกิดแล้ว เหตุการณ์นี้จะเกิดตาม ” ซึ่งสามารถเขียนเป็นรูปแบบ A=>B พร้อมด้วยค่าความน่าจะเป็น เช่น

Product = “ผ้าอ้อม” => Product = “เบียร์” ด้วยความน่าจะเป็นร้อยละ 40 หมายความว่า ถ้าเมื่อใดที่ลูกค้าซื้อผ้าอ้อมแล้วมีโอกาสร้อยละ 40 ที่ลูกค้าจะซื้อเบียร์ด้วย

Product = “น้ำอัดลม” => Product = “มันฝรั่ง” => Product = “น้ำผลไม้” ด้วยความน่าจะเป็นร้อยละ 80 หมายความว่า ถ้าเมื่อใดที่ลูกค้าซื้อน้ำอัดลมกับมันฝรั่ง แล้วมีโอกาสร้อยละ 80 ที่ลูกค้าจะซื้อน้ำผลไม้ด้วย

จากนั้นจะนำธุรกรรมการซื้อสินค้านำมาพิจารณารูปแบบ และพฤติกรรมการซื้อของลูกค้าได้ เช่น พิจารณาธุรกรรมการซื้อสินค้าของลูกค้ากลุ่มที่ซื้อเบียร์พร้อมกับผ้าอ้อมก็พบว่าผู้ซื้อนั้นเป็นชายหนุ่มที่เพิ่งจะมีลูกอ่อน จึงทำให้เข้าใจพฤติกรรมการซื้อนี้และจัดวางสินค้าให้เข้ากับพฤติกรรมลูกค้าได้ โดยขยับชั้นวางผ้าอ้อมมาใกล้กับชั้นวางเบียร์เพื่อสะดวกในการหยิบ หรือกรณีที่ลูกค้าซื้อน้ำอัดลมและมันฝรั่งแล้วมีโอกาสมากที่จะซื้อน้ำผลไม้ด้วย ก็นำมาปรับรูปแบบการวางสินค้าให้น้ำอัดลม มันฝรั่งและน้ำผลไม้อยู่ด้วยกัน หรือจากข้อมูลดังกล่าวสามารถนำมาวางแผนการเก็บสินค้าในค ลังได้ว่าสินค้าทั้งสามชนิดนี้จะต้องมีจำนวนสำรองอยู่ในคลังพอ ๆ กัน เนื่องจากมักถูกขายไปพร้อม ๆ กัน ทำให้มีสินค้าเพียงพอต่อความต้องการของลูกค้า (Tang, 2005)

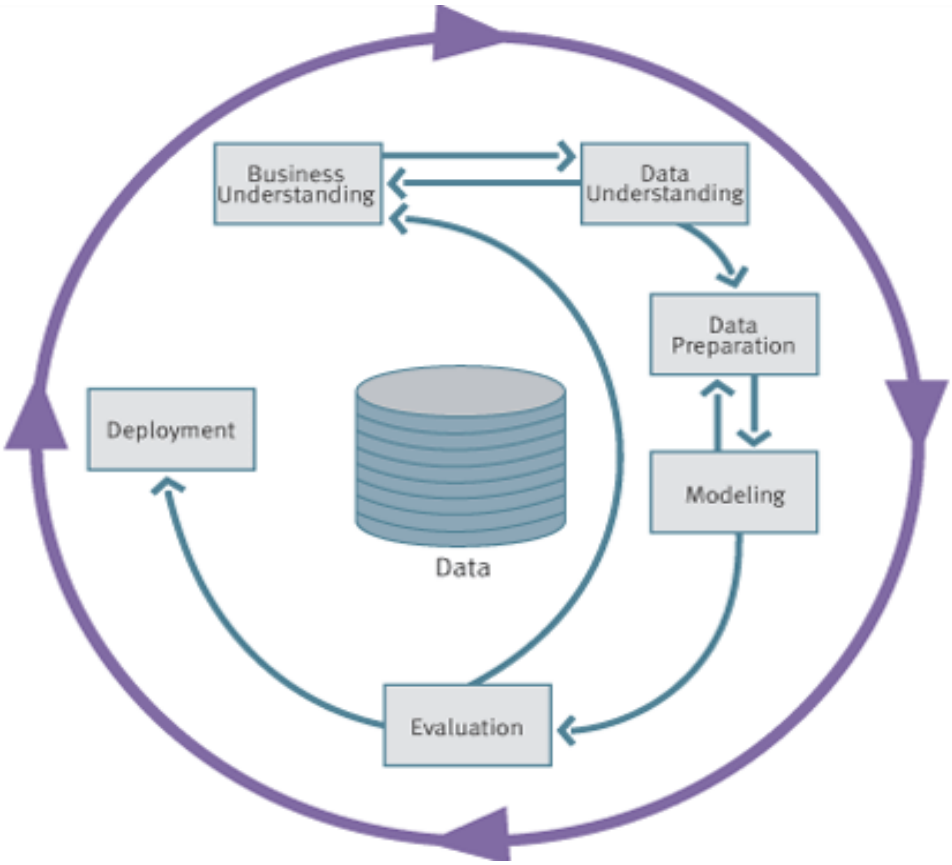
2.1.4 CRoss-Industry Standard Process for Data Mining (CRISP-DM)

จากความต้องการวิธีการในการค้นหาความรู้ที่มีประสิทธิภาพก่อให้เ กิดขั้นตอนวิธีในการทำเหมืองข้อมูล (Data Mining Algorithm) และเครื่องมือช่วยในการทำเหมืองข้อมูล (Data Mining Tool) อย่างไรก็ตามเนื่องจากความซับซ้อนในกระบวนการทำเหมืองข้อมูล การทำเหมืองข้อมูลจำเป็นต้องอาศัยกระบวนการมาตรฐานในการทำเหมืองข้อมูล (Mining Methodology) ซึ่งหนึ่งในนั้นก็คือ CRISP-DM (Pete Chapman)

CRISP-DM (CRoss-Industry Standard Process for Data Mining) มาตรฐานในการทำเหมืองข้อมูล ซึ่งเกิดจากความร่วมมือระหว่าง บริษัท DaimlerChrysler บริษัท SPSS และบริษัท NCR โดยเป็นแบบจำลองในการทำเหมืองข้อมูลที่น่าเชื่อถือ อีกทั้งสามารถเข้าใจได้อย่างรวดเร็ว โดยคนที่มีพื้นฐานความรู้ด้านเหมืองข้อมูลน้อยก็สามารถใช้งานได้

ภาพที่ 2.2

แบบจำลอง CRISP



ที่มา : (Pete Chapman)

กระบวนการทำเหมืองข้อมูล ตามขั้นตอนของ CRISP-DM มี 6 ขั้นตอนด้วยกัน คือ

<u>ขั้นตอนที่ 1</u> ช่วงทำความเข้าใจธุรกิจ (Business Research/Understanding Phase) โดยเป็นการกำหนดจุดประสงค์และความต้องการของโครงการ จากนั้นจะแปลงจุดประสงค์ให้อยู่ในรูปแบบของปัญหาเหมืองข้อมูล

ขั้นตอนที่ 2 ช่วงทำความเข้าใจข้อมูล (Data Understanding Phase) โดยเป็นการทำความเข้าใจกับข้อมูลที่จัดเก็บ รวมรวบข้อมูล ศึกษาและทำความคุ้นเคยกับข้อมูล ตลอดจนประเมินคุณภาพของข้อมูลที่ได้มา

ขั้นตอนที่ 3 ช่วงเตรียมข้อมูล (Data Preparation Phase) โดยเป็นการเตรียมข้อมูล โดยเตรียมข้อมูลดิบที่มีให้เป็นข้อมูลที่จะต้องใช้ในขั้นตอนที่เหลือ (ขั้นตอนนี้ใช้เวลานาน) ตลอดจนเลือกตัวแปรที่ต้องการวิเคราะห์ให้เหมาะสม อีกทั้งแปลงรูปแบบของตัวแปรให้อยู่ในรูปแบบเดียวกัน เพื่อให้ข้อมูลพร้อมสำหรับการนำไปสร้างแบบจำลอง

ขั้นตอนที่ 4 ช่วงสร้างแบบจำลอง (Modeling Phase) โดยเป็นการคัด เลือกแบบจำลองที่เหมาะสม ปรับปรุงตัวแปร ลักษณะแบบจำลองเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด อาจจะใช้เทคนิคหลายๆ เทคนิคเข้ามาช่วยวิเคราะห์ได้ ถ้าจำเป็น ก็สามารถย้อนกลับไปช่วงการเตรียมข้อมูล เพื่อเตรียมข้อมูลให้เหมาะสมกับการสร้างแบบจำลองใหม่ได้

ขั้นตอนที่ 5 ช่วงประเมินผล (Evaluation Phase) โดยเป็นการประเมินแบบจำลองที่ใช้ในการวิเคราะห์ทั้งหมด ประเมินว่าแบบจำลองไหนตอบโจทย์ในขั้นตอนแรกได้ดีที่สุด ตรวจสอบความถูกต้องและสภาพแวดล้อมต่างๆ ตัดสินใจในการนำผลลัพธ์ไปใช้

ขั้นตอนที่ 6 ช่วงนำไปใช้งาน (Deployment Phase) โดยเป็นการนำไปใช้งานจริง รวมถึงนำเสนอตัวอย่างจากการนำไปใช้งานจริง

โดยขั้นตอนย่อยในแต่ละขั้นตอนของ CRISP-DM แสดงได้ดังตารางที่ 2.1

ตารางที่ 2.1
ตารางแสดงขั้นตอนของ CRISP-DM

ขั้นตอน (Process)	วิธีการ (Method)	ผลลัพธ์ (Output)
1. ทำความเข้าใจในธุรกิจ (Business Research /Understanding)	กำหนดวัตถุประสงค์ของโครงการ (Determine Business Objective) โดยควรตอบสนองต่อความต้องการของลูกค้า ตลอดจนต้องสอดคล้องกับกฎเกณฑ์ในการดำเนินของธุรกิจ	● ข้อมูลสถานะขององค์กร (Background) ● วัตถุประสงค์ ในการทำเหมืองข้อมูล อีกทั้งต้องตอบสนองต่อการดำเนินของธุรกิจ (Business Objectives) ● ข้อกำหนด ที่ทำให้ธุรกิจประสบความสำเร็จ (Business Success Criteria)
	ประเมินสถานการณ์ของธุรกิจ (Assess Situation) โดยเป็นการลงรายละเอียดเกี่ยวกับทรัพยากร กฎเกณฑ์ ข้อสมมุติฐาน ตลอดจนปัจจัยอื่น ๆ ที่ควรพิจารณาในการดำเนินธุรกิจ เช่น รายละเอียดของทรัพยากรที่มีอยู่ภายในโครงการ รวมถึง รายละเอียด ส่วนบุคคล (ผู้เชี่ยวชาญทางธุรกิจ ผู้เชี่ยวชาญข้อมูล และผู้สนับสนุนทางเทคนิค)	● ข้อมูลการดำเนินงานและเครื่องมือที่ใช้ในการทำเหมืองข้อมูล ● ข้อสงสัยหรือ สมมุติฐาน และข้อยกเว้นหรือ ข้อควรหลีกเลี่ยงในการทำเหมืองข้อมูล นอกจากนี้ยังจำเป็นต้องระบุความเสี่ยงที่ทำให้โครงการล่าช้าหรือล้มเหลว ● ผลวิเคราะห์ค่าใช้จ่าย ● ผลกำไรของโครงการ

ตารางที่ 2.1 (ต่อ)
ตารางแสดงขั้นตอนย่อยของ CRISP-DM

ขั้นตอน (Process)	วิธีการ (Method)	ผลลัพธ์ (Output)
1. ทำความเข้าใจในธุรกิจ (Business Research /Understanding)	กำหนดเป้าหมายในการทำเหมืองข้อมูล (Determine Data Mining Goals) ซึ่งจำเป็นที่ต้องสอดคล้องกับเป้าหมายทางธุรกิจขององค์กร	● รายละเอียดผลลัพธ์ของการทำเหมืองข้อมูล ที่ช่วยให้เป้าหมายธุรกิจประสบผลสำเร็จ ● กฎเกณฑ์ ในการทำ งาน เพื่อให้ประสบผลสำเร็จในการทำเหมืองข้อมูล
	สร้างแผน การ ดำเนิน งานภายในโครงการ (Produce Project Plan) โดย บรรยายแผนสำหรับการทำเหมืองข้อมูล อีกทั้ง ระบุขั้นตอนในการดำเนินงาน	● กระบวนการ ในการดำเนินงาน ของโครงการ ซึ่งรวมถึงระยะ เวลา และทรัพยากรที่ใช้
2. การทำความเข้าใจข้อมูล (Data Understanding)	รวบรวมข้อมูลของทรัพยากร ในเบื้องต้นที่ต้องใช้ภายในโครงการ (Collect Initial Data) โดยขั้นตอนนี้เป็นขั้นตอนสำคัญที่นำไปสู่ ขั้นตอนการจัดเตรียมข้อมูล	● ชุดข้อมูลที่ต้องใช้ภายในโครงการ และขั้นตอน วิธีในการจัดเก็บชุดข้อมูล
	บรรยาย คุณลักษณะของ ข้อมูล (Describe) โดย มีการ ตรวจสอบคุณลักษณะของข้อมูลและจัดทำรายงานที่เกี่ยวข้อง	● รูปแบบและขนาดของข้อมูลที่ใช้ในโครงการ

ตารางที่ 2.1 (ต่อ)
ตารางแสดงขั้นตอนย่อยของ CRISP-DM

ขั้นตอน (Process)	วิธีการ (Method)	ผลลัพธ์ (Output)
2. การทำความเข้าใจข้อมูล (Data Understanding)	สำรวจข้อมูลที่มีอยู่ (Explore Data) โดยวิเคราะห์ปัญหาของการทำเหมืองข้อมูลว่าสามารถแก้ ปัญหาโดยอาศัยการเรียกดูข้อมูล (Query) การนำเสนอ รูป แบบเสมือน จริง (Visualization) และ การนำเสนอรายงาน (Report)	● รายงานการสำรวจข้อมูลที่ระบุการสมมุติฐานเบื้องต้นที่ส่งผลกระทบต่องานที่เหลือของโครงการ
	การ ตรวจสอบคุณภาพของข้อมูล (Verify Data) เช่น ข้อมูลสมบูรณ์หรือไม่ มีข้อมูลที่ขาดหายไปหรือไม่	● รายงานแสดงผลการตรวจสอบความถูกต้องของข้อมูล โดยที่หากพบปัญหาต้องระบุวิธีการแก้ไขปัญหา ซึ่งวิธีการแก้ไขปัญหาขึ้นอยู่กับตัวข้อมูลและองค์ความรู้ทางธุรกิจ
3. ช่วงการ จัด เตรียมข้อมูล (Data Preparation Phase)	คัดเลือกข้อมูล (Select Data) โดยตัดสินใจว่าควรวิเคราะห์ข้อมูลใดตลอดจนคำนึงถึงกฎเกณฑ์ภายใต้วัตถุประสงค์ในการทำเหมืองข้อมูล	● รายละเอียดของข้อมูลที่ควรหรือไม่ควรวิเคราะห์พร้อมทั้งระบุเหตุผลในการคัดเลือก
	การทำความสะอาด ข้อมูล (Clean Data) โดยเป็นการปรับปรุงคุณภาพของข้อมูลให้ดีขึ้น ให้ตรงกับรูปแบบข้อมูลที่ต้องการ	● รายงานแสดงการทำความสะอาดข้อมูล

ตารางที่ 2.1 (ต่อ)
ตารางแสดงขั้นตอนย่อยของ CRISP-DM

ขั้นตอน (Process)	วิธีการ (Method)	ผลลัพธ์ (Output)
3. ช่วงการ จัด เตรียมข้อมูล (Data Preparation Phase)	สร้างโครงสร้างข้อมูล (Construct Data) โดยเป็นกระบวนการในการจัดเตรียมข้อมูล เช่น การสร้างข้อมูลบางส่วนหรือสร้างข้อมูลใหม่ทั้งหมด ตลอดจนการแปลงรูปแบบข้อมูลที่มีอยู่เดิม	● ข้อมูลที่ถูกสร้างขึ้นมาใหม่บางส่วนจากข้อมูลเดิม (Derived Attribute) เช่น ข้อมูลความยาวและข้อมูลความกว้าง เป็นต้น ● ข้อมูลที่ถูกสร้างขึ้นมาใหม่ (Generated Record) เช่น ข้อมูลลูกค้าที่ไม่ได้ซื้อสินค้าเลยในปีที่ผ่านมา
	ผสานข้อมูล (Integrate Data) โดยเป็นการสร้างตารางข้อมูลใหม่จากการรวมกันของตารางหลายตาราง หรือระเบียนหลายระเบียน	● การรวมกันของตารางตั้งแต่สองตารางขึ้นไป
	เปลี่ยนรูปแบบของข้อมูล (Format Data) ให้สามารถทำงาน ภายในเครื่องมือที่ใช้ ในการ ออกแบบ (Modeling Tool) โดยที่ไม่มีการเปลี่ยนแปลงความหมาย	● รูปแบบข้อมูลซึ่งสามารถทำงาน ด้วย เครื่องมือที่ใช้ออกแบบได้ โดยเครื่องมือต้องมีการลำดับของข้อมูล เช่น ข้อมูลที่อยู่สดมภ์ ในส่วนแรกของแต่ละระเบียนต้องไม่ซ้ำกัน

ตารางที่ 2.1 (ต่อ)
ตารางแสดงขั้นตอนย่อยของ CRISP-DM

ขั้นตอน (Process)	วิธีการ (Input)	ผลลัพธ์ (Output)
4. ช่วงสร้างแบบจำลอง (Modeling Phase)	เลือกเทคนิคในการออกแบบ (Select Modeling Technique) เช่น สร้างต้นไม้แห่งการตัดสินใจ (Decision Tree) หรือโครงสร้างเครือข่ายประสาทเทียม (Neural Network)	● เทคนิ คที่ใช้ในการออกแบบเหมืองข้อมูล ● ข้อสมมุติฐานเกี่ยวกับแบบจำลองของการทำเหมืองข้อมูล
	สร้างแบบจำลอง (Build Model)	● แบบจำลอง จากการทำเหมืองข้อมูล (Models) และคุณลักษณะพื้นฐานของแบบจำลอง (Model Description)
	ประเมินแบบจำลอง (Assess Model) โดยเป็นการประเมินผลลัพธ์แบบจำลองที่ได้จากการทำเหมืองข้อมูล	● ผลการประเมิน (Model Assessment) โดยสรุปคุณภาพของแบบจำลองและจัดอันดับคุณภาพของแบบจำลอง ● ผลที่ได้จากการปรับแต่งค่าพารามิเตอร์ ต่างๆ ของแบบจำลอง
5. ช่วงประเมินผล (Evaluation Phase)	ประเมินผล (Evaluate Results) โดยเป็นการหาแบบจำลอง จากการ ทำ เหมืองข้อมูล ที่ตรงกับวัตถุประสงค์ทางธุรกิจมากที่สุด	● ผลการประเมินการทำเหมืองข้อมูลโดย ต้องคำนึงถึงกฎเกณฑ์ทางธุรกิจและการปรับปรุงแบบจำลอง ในการทำเหมืองข้อมูล

ตารางที่ 2.1 (ต่อ)
ตารางแสดงขั้นตอนย่อยของ CRISP-DM

ขั้นตอน (Process)	วิธีการ (Input)	ผลลัพธ์ (Output)
5. ช่วงประเมินผล (Evaluation Phase)	ทบทวนกระบวนการ ทำเหมืองข้อมูล (Review Process) โดยที่ทำการ ทบทวนกระบวนการทำเหมืองข้อมูลเพื่อใช้ในการตัดสินใจปัจจัยที่เกี่ยวข้องกับความต้องการทางธุรกิจ	● ข้อสรุปของกระบวนการและข้อเน้นย้ำกิจกรรมที่ผิดพลาดหรือกิจกรรมที่ควรทำซ้ำ
	ค้นหาขั้นตอนต่อไป (Determine Next Steps) โดยที่ทำหลังจากประเมินผล ที่ได้ และทบทวนกระบวนการทำเหมืองข้อมูลเพื่อให้ทราบว่าควร ที่จะสิ้นสุดโครงการหรือเริ่มกระบวนการทำเหมือนข้อมูลใหม่เมื่อใด โดยงานนี้เกี่ยวข้องกับการวิเคราะห์ทรัพยากรและงบประมาณที่มีอิทธิพลต่อการตัดสินใจ	● รายละเอียดการกระทำที่มีโอกาสเป็นไปได้ พร้อมทั้งเหตุผลในแต่ละการกระทำ
6. ช่วงนำไปใช้งาน (Deployment Phase)	วางแผนในการดำเนินงาน (Plan Deployment) โดยนำผลลัพธ์ ที่ได้จากการทำเหมืองข้อมูล มาทำการ สรุป ตลอดจน เป็นการวางแผนและ กำหนดกลยุทธ์ในการดำเนินงาน	● แผนการดำเนินงานที่ทำการกำหนด กลยุทธ์ในการดำเนินงาน รวมถึงขั้นตอนที่จำเป็นในการดำเนินงาน

ตารางที่ 2.1 (ต่อ)
ตารางแสดงขั้นตอนย่อยของ CRISP-DM

ขั้นตอน (Process)	วิธีการ (Input)	ผลลัพธ์ (Output)
6. ช่วงนำไปใช้งาน (Deployment Phase)	วางแผน การเพื่อ ตรวจสอบ และบำรุงรักษา (Plan Monitoring and Maintenance) โดยที่การตรวจสอบและบำรุงรักษาเป็นประเด็นสำคัญในการทำเหมืองข้อมูล	● ข้อสรุป ที่ได้ จากการตรวจสอบและ การบำรุงรักษา ตลอดจนกระบวนการ ในการดำเนินงาน
	สร้างรายงาน ผลลัพธ์ขั้นสุดท้าย (Produce Final Report) โดยที่รายงานขั้นสุดท้าย นั้นขึ้นอยู่กับแผนการนำไปใช้งาน จริง ซึ่งจะประกอบไปด้วยข้อสรุปของโครงการและประสบการณ์ที่ได้รับ	● รายงานขั้นสุดท้าย
	ทบทวน การทำเหมืองข้อมูล (Review Project) โดยที่ประเมินข้อประสิทธิภาพการทำเหมืองข้อมูล เพื่อหา แนว ทางในการปรับปรุงให้ดีขึ้น	● ข้อสรุป ที่ได้รับ จากประสบการณ์ในการทำเหมืองข้อมูล

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

2.2.1 An Extended Process Model of Knowledge Discovery in Database

ภาพที่ 2.3

กระบวนการค้นหาความรู้ในฐานข้อมูล

ที่มา : (Fayyad & Stolorz, 1997)

Fayyad and Stolorz (1997) ได้ทำการศึกษาเรื่องกระบวนการค้นหาความรู้ในฐานข้อมูล (Knowledge Discovery in Database หรือ KDD) โดยกล่าวว่ากระบวนการค้นหาความรู้ในฐานข้อมูล เป็นกระบวนการหารูปแบบหรือสารสนเทศที่เป็นประโยชน์ในฐานข้อมูลซึ่งมีขนาดใหญ่ โดยเริ่ม ต้นตั้งแต่กระบวนการเก็บรวบรวมข้อมูล (Collection) กระบวนการคัดเลือกข้อมูล (Selection) กระบวนการจัดการข้อมูลก่อนวิเคราะห์ (Preprocessing) กระบวนการแปลงข้อมูล (Transformation) การทำเหมืองข้อมูลหรือการใช้เครื่องมือทางสถิติในการวิเคราะห์ข้อมูล (Data Mining) รวมไปถึงการแปลและประเมินสารสนเทศหรือองค์ความรู้ที่ได้จากการวิเคราะห์ข้อมูล (Interpretation and Evaluation) โดยที่กระบวนการค้นหาความรู้ในฐานข้อมูล คือ รูปแบบโดยรวมของกระบวนการค้นหาความรู้จากข้อมูล ในขณะที่การทำเหมืองข้อมูล เป็นเพียงขั้นตอนหนึ่งในกระบวนการเหล่านี้ โดยการทำเหมืองข้อมูล เป็นเพียงการนำเอาขั้นตอนวิธีที่เฉพาะเจาะจงมาดึงรูปแบบความสัมพันธ์ที่ได้จากข้อมูลทั้งหมด

2.2.2 การทำเหมืองข้อมูลเพื่อพยากรณ์การล้มละลายขององค์กร (Dynamics of Modeling in Data Mining:Interpretive Approach to Bankruptcy Prediction)

Tae Kyung Sung (1999) ทำการศึกษาเรื่องการทำเหมืองข้อมูลเพื่อพยากรณ์การล้มละลายขององค์กร พบว่าตัวแปรสำคัญในคาดการณ์การล้มละลายขององค์กรคือ กระแสเงินสดภายในสินทรัพย์รวม (Cash Flow to Liabilities) และสินทรัพย์ที่ผู้ถือหุ้นตราสารทุนและหนี้สิน (Fixed Assets to Stockholders Equity)

1. Business/Research Understanding Phase

จากวิกฤติเศรษฐกิจในเอเชียตะวันออกก่อให้เกิดปรากฏการณ์การล้มละลายขององค์กรไปทั่วโลก งานวิจัยนี้จึงมีเป้าหมายเพื่อสร้างแบบจำลองในการพยากรณ์การล้มละลายของบริษัท เพื่อช่วยลดผลกระทบที่จะเกิดกับเศรษฐกิจการเงินภายในและภายนอกประเทศ

2. Data Understanding Phase

ข้อมูลที่ได้มาจาก 2 แหล่ง คือ องค์กรในยุคที่เศรษฐกิจมีความมั่นคง (ในช่วงปี ค.ศ. 1991-1995) และ องค์กรในยุคที่เกิดวิกฤติเศรษฐกิจ (ในช่วงปี ค .ศ. 1997-1998) จากองค์กรชั้นนำจำนวน 29 องค์กร โดยข้อมูลทางการเงินได้มาจากหน่วยงานแลกเปลี่ยนทางการเงินของประเทศเกาหลี (Korean Stock Exchange) และรับรองความถูกต้องของข้อมูลโดยธนาคารแห่งประเทศเกาหลี (Bank of Korea) และธนาคารเพื่ออุตสาหกรรมเกาหลี (Korea Industrial Bank)

3. Data Preparation Phase

เนื่องจากมีความซ้ำซ้อนของข้อมูลทำให้มีการคัดข้อมูลจาก 16 องค์กรออกจากชุดข้อมูลที่ใช้วิจัย โดยสนใจข้อมูลเกี่ยวกับอัตราการเจริญเติบโตทางธุรกิจ กิจกรรม ความปลอดภัย ผลกำไร และผลผลิตขององค์กร

4. Modeling Phase

ใช้แบบจำลองต้นไม้แห่งการตัดสินใจ (Decision Tree) กับข้อมูลในสภาพปกติ (Normal-Conditions) และข้อมูลในสภาพวิกฤติ (Crisis-Conditions) เช่น ถ้าการเพิ่ม

ผลผลิต (Productivity) มากกว่า 19.65 พยากรณ์ได้ว่ามีโอกาสไม่ล้มละลายประมาณร้อยละ 19.65

ถ้าอัตราการไหลเวียนของเงิน (Cash Flow) มากกว่า 5.65 สามารถพยากรณ์ได้ว่ามีโอกาสไม่ล้มละลายประมาณร้อยละ 95

ถ้าการเพิ่มผลผลิตน้อยกว่า 19.65 และอัตราการไหลเวียนของเงินน้อยกว่า 5.65 พยากรณ์ได้ว่ามีโอกาสล้มละลายประมาณร้อยละ 84

5. Evaluation Phase

จากการหารือของผู้วิจัยและผู้เชี่ยวชาญทางการเงิน จากผลลัพธ์ที่ได้จากการทำเหมืองข้อมูล ได้ข้อสรุปว่าปัจจัยทุน (Capital Productivity) เป็นปัจจัยสำคัญในการบ่งชี้ว่าองค์กรใดมีความเสี่ยงในการล้มละลาย เพื่อให้แน่ใจได้ว่าแบบจำลองนี้ใช้ได้กับองค์กรผู้ผลิตสินค้าทางด้านอุตสาหกรรมทุกๆแห่ง จึงมีการสุ่มตัวอย่างขององค์กรที่ไม่ล้มละลาย (Control Sample) เพื่อเปรียบเทียบกับบริษัทในชุดข้อมูล

6. Deployment Phase

ยังไม่มีการนำผลงานวิจัยนี้ไปใช้งานจริง เนื่องจากการใช้จริงต้องขึ้นอยู่กับดุลพินิจของผู้ใช้งาน อย่างไรก็ตามจากผลงานวิจัยนี้ทำให้สถาบันการเงินภายในประเทศเกาหลีต่างตระหนักเกี่ยวกับปัจจัยหรือเงื่อนไขอันนำไปสู่การล้มละลายขององค์กร

2.2.3 A New Efficient Approach for Data Clustering in Electronic Library Using Ant Colony Clustering Algorithm

Chen (2006) ได้ทำการศึกษาเรื่องการทำเหมืองข้อมูล โดย ได้นำข้อมูลการใช้สื่ออิเล็กทรอนิกส์ของผู้ใช้บริการห้องสมุดมาแบ่งกลุ่มผู้ใช้ด้วยเทคนิคการวิเคราะห์ข้อมูล การสร้างคลังข้อมูล และการทำเหมืองข้อมูล โดย จากการทำเหมืองข้อมูลนั้นได้ใช้ขั้นตอนวิธี Ant Colony Clustering และ K-means Clustering ในการแบ่งกลุ่มผู้ใช้บริการ ผลการวิเคราะห์พบว่ากลุ่มที่เกิดจากการแบ่งด้วยขั้นตอนวิธี Ant Colony Clustering นั้น สมาชิกในกลุ่มจะมีลักษณะคล้ายกันมากกว่ากลุ่มสมาชิกที่เกิดจากการแบ่งด้วยขั้นตอนวิธี K-Means Clustering ดังนั้นขั้นตอนวิธี Ant Colony Clustering จะมีประสิทธิภาพมากกว่าขั้นตอนวิธี K-means Clustering

2.2.4 Using Data Mining Technology to Provide a Recommendation Service in the Digital Library

Chen (2007) ได้ทำการศึกษาเรื่องการทำเหมืองข้อมูล โดยนำธุรกรรมการยืมคืนหนังสือ และวัสดุสารนิเทศของสมาชิกห้องสมุดมาทำเหมืองข้อมูล มาทำการแบ่งกลุ่มด้วยขั้นตอนวิธี Ant Colony Clustering จากนั้นจึงแบ่งกลุ่มของผู้ใช้บริการ ห้องสมุดออก เป็น 3 กลุ่ม แล้วนำข้อมูลในแต่ละกลุ่มมาหากฎความสัมพันธ์ (Association Rule) ของหนังสือและวัสดุสารนิเทศ โดยใช้ขั้นตอนวิธี Apriori ซึ่งผลการศึกษาที่ได้ไม่เพียงนำมาสร้างระบบให้คำแนะนำสำหรับการค้นหาหนังสือผ่านเว็บเพจเท่านั้น แต่ยังช่วยในการค้นหาหนังสือที่เหมาะสมกับสมาชิกห้องสมุด และยังเป็นแนวทางในการจัดสรรงบประมาณของห้องสมุดที่มีจำกัดให้สามารถเลือกซื้อเฉพาะหนังสือที่ตรงกับความต้องการของผู้ใช้ได้มากขึ้น

2.2.5 Organizational Data Mining: Leveraging Enterprise Data Resource for Optimal Performance

Nicholson (2003) ได้ทำการศึกษาเรื่องการทำเหมืองข้อมูล ภายในห้องสมุด โดยกล่าวถึงการประยุกต์ใช้หลักการเหมืองข้อมูล กับงานด้านต่าง ๆ ของห้องสมุด เช่น

- ใช้ปรับปรุงระบบ OPAC ให้สามารถแสดงข้อมูลหนังสือเล่มอื่น ๆ ที่มีเนื้อหาเกี่ยวข้องกับหนังสือที่ผู้ใช้ต้องการยืม เพื่อเป็นทางเลือกให้ผู้ใช้สามารถหาข้อมูลที่เกี่ยวข้องกันได้โดยไม่ต้องเสียเวลาในการค้น
- ใช้ในการทำนายความต้องการใช้หนังสือในอนาค ต โดยดูจากรูปแบบการยืมของหนังสือที่มีจำนวนครั้งของการยืมสูงแล้วนำมาตัดสินใจว่าควรจะมีสำเนาของหนังสือนั้นกี่เล่ม หรือนำมาประเมินว่าควรจะมีหนังสือในหมวดใด จำนวนใด เท่าใด จึงจะเพียงพอต่อการใช้งานของผู้ใช้บริการ
- ใช้ในการค้นหารูปแบบการใช้งานเว็บไซต์ของห้องสมุด แล้วนำข้อมูลที่ได้มาปรับปรุงเว็บไซต์ให้เหมาะสม เช่น การวางข้อความที่สำคัญไว้ในตำแหน่งที่ผู้ใช้มองเห็นง่าย หรือปรับปรุงให้เว็บไซต์น่าดึงดูดเพื่อไม่ให้ผู้ใช้ออกจากเว็บไซต์
- ใช้ในการค้นหาว่ามีปัจจัยใดที่ทำให้เกิดการใช้ทรัพยากรต่าง ๆ ในห้องสมุดน้อย ใช้ในการแบ่งกลุ่มผู้ใช้ห้องสมุดกับประเภทหนังสือที่มีการยืม

- ใช้เป็นแนวทางในการหารูปแบบการใช้บริการยืมคืนหนังสือ เพื่อจะได้ทราบว่าช่วงเวลาใดที่มีการยืมคืนหนังสือกันมากก็จัดให้มีเจ้าหน้าที่ประจำอยู่หลายคน หากเวลาใดที่มีการยืมคืนน้อยก็จัดให้เจ้าหน้าที่ประจำอยู่เพียงไม่กี่คน
- ใช้ในการวิเคราะห์ข้อมูลการจัดซื้อหนังสือเข้าห้องสมุดว่าควรจะซื้อจากใครเพื่อให้ได้ประโยชน์สูงสุด
- ใช้ในการ แบ่งกลุ่มของชนิดงานที่ทำในห้องสมุดตามเวลาที่ใช้งานเหล่านั้น เพื่อให้สามารถจัดการระบบการทำงานให้มีประสิทธิภาพที่สุด
- ใช้ในการค้นหารูปแบบหรือแนวคำถามที่ผู้ใช้ห้องสมุดมักจะถามบรรณารักษ์ จากนั้นจึงนำแนวคำถามดังกล่าวมาฝึกให้กับบรรณารักษ์คนใหม่ ซึ่งจะทำให้การฝึกนี้มีประสิทธิภาพ ประหยัดเวลา ประหยัดงบประมาณ และตอบสนองความต้องการของผู้ใช้ห้องสมุดได้
- ใช้ในการค้นหารูปแบบของระยะเวลาการยืมและคืนหนังสือ รวมถึงการต่ออายุการยืมหนังสือ ทำให้ผู้บริหารห้องสมุดสามารถวางนโยบายการยืมหนังสือได้สอดคล้องกับความต้องการของสมาชิกห้องสมุดมากขึ้น
- ใช้ในการ ค้นหารูปแบบว่าหนังสือเล่มใด ชนิดใด ที่มักถูกขโมย หรือส่งคืนช้ากว่ากำหนด เพื่อนำมาวางนโยบายต่าง ๆ ให้หนังสือหายน้อยลง หรือให้ผู้ใช้คืนหนังสือตรงเวลามากขึ้น