

บทที่ 3

ระเบียบวิธีวิจัย

3.1 รูปแบบงานวิจัย

งานวิจัยนี้เกี่ยวกับการวิจัยเชิงทดลอง (Experimental Research) โดยเป็นการนำฐานข้อมูลธุรกรรมการยืมคืนหนังสือที่สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ ได้ทำการจัดเก็บไว้แล้วซึ่งเป็นข้อมูลทุติยภูมิ (Secondary Data) มาทำการศึกษาพฤติกรรมของผู้ใช้ โดยการคำนวณทางสถิติ (Descriptive Statistics) ในการอธิบายลักษณะประจำของกลุ่มตัวอย่างที่นำมาศึกษา และทำเหมืองข้อมูลตามขั้นตอนของ CRISP-DM

3.2 ฐานข้อมูลที่ศึกษา

ข้อมูลที่นำมาศึกษาเป็นข้อมูลจากแหล่งข้อมูลทุติยภูมิ คือ ฐานข้อมูล การยืมคืนหนังสือจากโปรแกรมห้องสมุดอัตโนมัติ HORIZON ซึ่งทำการจัดเก็บในฐานข้อมูล SYBASE โดยที่สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ ได้ทำการเก็บไว้ในคลังข้อมูล ตั้งแต่ไตรมาสที่ 1 ถึงไตรมาสที่ 2 ของปี พ.ศ. 2552 (เมษายน พ.ศ. 2552 ถึงมิถุนายน พ.ศ. 2552)

3.3 ขนาดของฐานข้อมูลที่ศึกษา

ธุรกรรมการยืมคืนหนังสือ (Transaction) ของสมาชิก สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ มีประมาณ 6 หมื่นรายการ และสมาชิกห้องสมุดมีประมาณ 2 หมื่นราย

3.4 เครื่องมือที่ใช้ในการวิจัย

3.4.1 การใช้เครื่องมือทำเหมืองข้อมูล (Data Mining Tool)

การใช้เครื่องมือทำเหมืองข้อมูลเพื่อวิเคราะห์หารูปแบบของข้อมูล และความสัมพันธ์ของข้อมูล โดยนำข้อมูลที่ต้องการได้ทำการจัดเก็บไว้แล้ว ซึ่งเป็นข้อมูลทุติยภูมิ เช่น ข้อมูลประเภทหนังสือที่ยืมคืน รายการหนังสือที่ยืมคืนในแต่ละครั้ง และวันเวลาในการยืมคืนหนังสือ เป็นต้น มาทำการศึกษาพฤติกรรม โดยการใช้ซอฟต์แวร์ทำเหมืองข้อมูล และคำนวณผลทางสถิติ

โปรแกรมที่ใช้ประมวลผล (Software Tool)

ผู้วิจัยนำข้อมูลของ สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ มาประมวลผลด้วยโปรแกรม Microsoft SQL Server 2005 ซึ่งพัฒนาโดย บริษัท Microsoft โดยโปรแกรม Microsoft SQL Server 2005 ประกอบด้วยเทคโนโลยีต่อไปนี้

- SQL Server Database Engine
- SQL Server Analysis Services
- SQL Server Integration Services
- SQL Server Replication
- SQL Server Reporting Services
- SQL Server Notification Services
- SQL Server Service Broker
- Full – Text Search
- SQL Server Tools and Utilities

สามารถอธิบายขอบเขตของการทำงานของเทคโนโลยีดังกล่าว ได้ดังนี้

SQL Server Database Engine

เป็นเทคโนโลยีที่ใช้ในการเก็บการประมวลผล และรักษาความปลอดภัยให้ข้อมูล อีกทั้งยังทำให้สามารถควบคุมการเข้าถึง และการประมวลผลที่รวดเร็วเพื่อสนองต่อความต้องการขององค์กร

SQL Server Analysis Services

เป็นเทคโนโลยีที่ใช้ในการทำ Online Analytical Processing (OLAP) และ การทำเหมืองข้อมูลเพื่อนำไปใช้ด้านธุรกิจ Analysis Service สนับสนุนการทำ OLAP โดยให้ผู้ใช้สามารถออกแบบ สร้าง และจัดการข้อมูลที่ได้จากการรวบรวมข้อมูลจากแหล่งอื่น ๆ เข้าด้วยกัน สำหรับการทำให้เหมืองข้อมูลแล้ว Analysis Service ให้ผู้ใช้สามารถออกแบบ สร้าง และดูตัวแบบเหมืองข้อมูลที่สร้างได้ โดยสามารถนำข้อมูลจากแหล่งอื่น ๆ มาทำเหมืองข้อมูลด้วย ขั้นตอนวิธี ที่เป็นมาตรฐานและใช้กันอย่างกว้างขวาง

SQL Server Integration Services

เป็นเทคโนโลยีในการแปลงและรวบรวมข้อมูลระดับองค์กร ซึ่งผู้ใช้สามารถนำเครื่องมือไปทำการสกัด แปลง และรวบรวมข้อมูลจากแหล่งข้อมูลหลายแหล่งที่อยู่แยกจากกันได้ แล้วนำข้อมูลนี้ไปไว้ปลายทางที่เดียวกัน หรือแยกกันหลาย ๆ ที่ก็ได้

SQL Server Replication

Replication เป็นชุดของเทคโนโลยีที่มีไว้สำหรับการสำเนา และกระจายข้อมูลหรือวัตถุในฐานข้อมูลจากฐานข้อมูลหนึ่งไปยังฐานข้อมูลอีกแห่งหนึ่ง และทำให้ข้อมูลในฐานข้อมูลหลาย ๆ แห่งมีความต้องกัน (Consistency) นอกจากนี้ยังช่วยให้สามารถกระจายข้อมูลไปยังที่ต่าง ๆ และสามารถ ถูใช้งานข้อมูลเหล่านั้น ผ่านระบบเครือข่าย เชื่อมโยงระยะใกล้ , เครือข่ายเชื่อมโยงระยะไกล, เครือข่ายไร้สาย หรือเครือข่ายอินเทอร์เน็ตได้

SQL Server Reporting Services

เป็นเทคโนโลยีที่ทำให้สามารถสร้างรายงานจากข้อมูลที่อยู่ในแหล่งข้อมูลเชิงสัมพันธ์หรือแหล่งข้อมูลหลายมิติได้ โดยสามารถสร้างและจัดการรายงาน ที่อยู่ในรูปแบบ Tabular หรือ Matrix หรือ Graphic หรือ Free-Form ได้ ซึ่งรายงานที่สร้างขึ้นมานี้สามารถดูและจัดการผ่านเว็บเพจได้

SQL Server Notification Services

เป็นเทคโนโลยีสำหรับสร้างข้อความแจ้งเตือนและนำส่งข้อความนั้นไปให้ผู้ใช้งานมากในเวลาอันรวดเร็ว

SQL Server Service Broker

เป็นเทคโนโลยีที่ช่วยให้งานที่กระทำบนฐานข้อมูลนั้นมีความไว้วางใจได้ สามารถปรับขนาดงานได้ และทำให้งานมีความปลอดภัย ซึ่ง Service Broker นั้นเป็นเทคโนโลยีที่อยู่ในส่วนของ Database Engine อีกที่ Service Broker จะใช้การสื่อสารด้วยข้อความเพื่องานต่าง ๆ ที่อยู่แยกกันให้สามารถทำงานร่วมกันได้ นอกจากนี้ Service Broker ยังมีเครื่องมือต่าง ๆ ที่จำเป็นสำหรับสร้างงานบนฐานข้อมูลแบบกระจาย และเครื่องมือเหล่านี้ช่วยลดระยะเวลาในการพัฒนาฐานข้อมูลลงได้มาก นอกจากนี้ Service Broker ยังช่วยให้ผู้ใช้สามารถปรับขนาดของงาน ให้เหมาะสมกับปริมาณงานที่ได้รับ

SQL Server Tools and Utilities

SQL Server มีเครื่องมือต่าง ๆ ที่ช่วยให้ผู้ใช้ออกแบบ พัฒนา นำไปประยุกต์ใช้ และบริหารฐานข้อมูลเชิงสัมพันธ์ , Analysis Services Cube, Data Transformation Packages, Replication Technologies, Reporting Servers และ Notification Servers ได้

3.5 ผู้มีส่วนได้ส่วนเสีย

ผู้มีส่วนได้ส่วนเสียจากการทำเหมืองข้อมูลตามขั้นตอน CRISP-DM ประกอบด้วย

- สมาชิกห้องสมุด ซึ่งประกอบด้วย อาจารย์ นักศึกษา ข้าราชการ พนักงานมหาวิทยาลัย และลูกจ้างของมหาวิทยาลัยธรรมศาสตร์
- ข้าราชการและพนักงานมหาวิทยาลัยสังกัด สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์
- บุคคลภายนอกและสมาชิกสมทบ

3.6 การรวบรวมข้อมูล

ในการวิจัยครั้งนี้มีการเก็บรวบรวมข้อมูล ในการยืมคืบหนังสือ จากสำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ โดยมีการเก็บรวบรวมข้อมูลดังนี้

3.6.1 การเก็บรวบรวมข้อมูลจากเอกสาร

ทำการรวบรวมทฤษฎี และแนวคิดที่เกี่ยวข้องกับขั้นตอนของ CRISP-DM และการทำเหมืองข้อมูลทั้งจากงานพิมพ์และข้อมูลทางอินเทอร์เน็ต

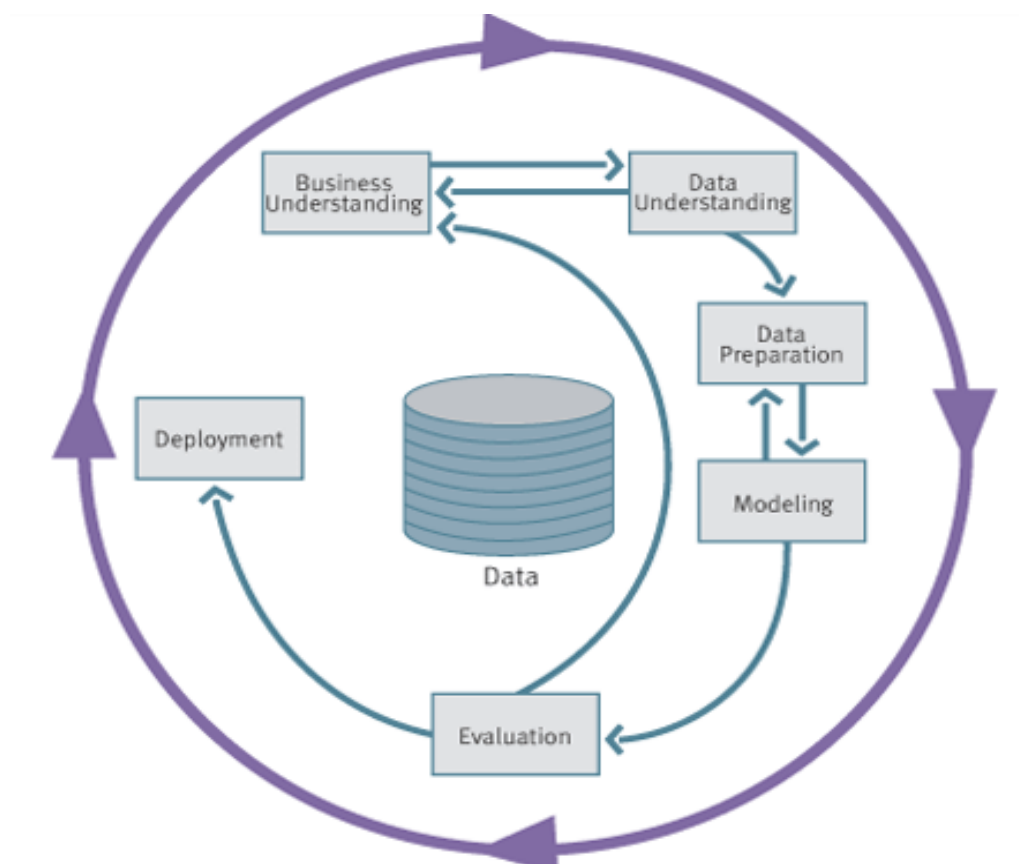
3.6.2 การเก็บข้อมูลจากฐานข้อมูล

ผู้วิจัยนำข้อมูลธุรกรรมการยืมคืบหนังสือของผู้ใช้บริการที่สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ ได้เก็บไว้ในฐานข้อมูลมาทำการวิเคราะห์ตามขั้นตอนของ CRISP-DM

3.7 การประมวลผลข้อมูล

ผู้วิจัยประมวลผลข้อมูลตามขั้นตอนของ CRISP-DM เพื่อให้ได้ความรู้ และนำความรู้ที่ได้มาประยุกต์ใช้ให้เกิดประโยชน์ต่อธุรกิจ ดังภาพที่ 3.1

ภาพที่ 3.1
การประยุกต์ใช้แบบจำลอง CRISP



ที่มา : (Pete Chapman)

ในการดำเนินโครงการนี้จะใช้การทำเหมืองข้อมูลโดยใช้ กฎความสัมพันธ์ (Association Rule) และกฎการแบ่งกลุ่ม (Clustering) ได้ดังนี้

3.7.1 ช่วงทำความเข้าใจธุรกิจ (Business/Research Understanding Phase)

ทำความเข้าใจเกี่ยวกับสายงาน หรือ ธุรกิจที่ต้องการศึกษา โดยศึกษาจากวิธีได้มาซึ่งความรู้ในการดำเนินงานในอดีต และศึกษาแนวทางในการดำเนินธุรกิจให้ประสบความสำเร็จ โดยสำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ ประสบกับปัญหาดังนี้

- สำนักหอสมุดสิ้นเปลืองงบประมาณในการจัดซื้อวัสดุสารสนเทศ เนื่องจากจำนวนและประเภทของวัสดุสารสนเทศไม่สอดคล้องต่อความต้องการในการใช้งาน

- วัสดุสารนิเทศไม่เพียงพอต่อความต้องการของผู้ใช้ในช่วงเหตุการณ์ เช่น ช่วงเวลาสอบ ห้องสมุด มีวัสดุสารนิเทศน้อยเกินไป ทำให้ไม่เพียงพอต่อการให้บริการแก่ผู้ใช้บริการ และหากไม่อยู่ในช่วงเวลาสอบ ผู้ใช้บริการจะไม่ยืมวัสดุสารนิเทศนั้นเลย นอกจากนี้แล้วยังส่งผลให้สำนักหอสมุดสิ้นเปลืองงบประมาณเป็นจำนวนมากในการจัดเก็บวัสดุสารนิเทศเข้าคลังหนังสือ
- วัสดุสารนิเทศของสำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ มีการยืมผ่านบริการยืมระหว่างห้องสมุดเป็นจำนวนมาก เช่น นักศึกษาที่อยู่วิทยาเขตท่าพระจันทร์ มีการยืม วัสดุสารนิเทศ บางชิ้นที่เก็บอยู่ที่วิทยาเขตรังสิตเป็นจำนวนบ่อยครั้ง ซึ่งนักศึกษาที่อยู่วิทยาเขตรังสิตเอง มีการยืม วัสดุสารนิเทศ ชิ้นนั้นเป็นจำนวนน้อยมาก

3.7.2 ช่วงทำความเข้าใจข้อมูล (Data Understanding Phase)

ข้อมูลที่ใช้ในการทำเหมืองข้อมูลได้มาจากธุรกรรมการยืมคืนหนังสือของสำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ ระหว่างเดือน มกราคม ถึงเดือนมิถุนายน พ.ศ.2552 จากฐานข้อมูล SYBASE อย่างไรก็ตามเนื่องจากฐานข้อมูลของสำนักหอสมุดมีขนาดใหญ่ ทำให้สำนักหอสมุดสามารถจัดเก็บข้อมูลธุรกรรมการยืมคืนหนังสือ ย้อนหลังได้เพียง 6 เดือนล่าสุดเท่านั้น ซึ่งหากต้องการข้อมูลที่มีระยะเวลานานหรือมีรายละเอียดมากกว่านั้น จำเป็นต้องเขียนคำสั่ง SQL เพื่อดึงข้อมูลขึ้นเอง ซึ่งหากทำเช่นนั้นจะกระทบต่อผู้ใช้ เพราะจะทำให้ระบบการยืมคืนเกิดความล่าช้า ดังนั้นจึงใช้ข้อมูลการยืมคืนของสำนักหอสมุดเป็นหลัก

3.7.3 ช่วงเตรียมข้อมูล (Data Preparation Phase)

ข้อมูลธุรกรรมการยืมคืนหนังสือของสมาชิกห้องสมุด ที่สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์จัดเก็บ แสดงได้ดังตารางที่ 3.1 และข้อมูลที่ใช้ภายในงานวิจัยแสดงได้ดังตารางที่ 3.2 ตามลำดับ

ตารางที่ 3.1

แสดงรายละเอียดข้อมูลสำหรับหอสมุดจัดเก็บ

Col#	Name	Description
1	BORROWERID	รหัสสมาชิก
2	ITEMID	รหัสหนังสือ
3	TDATE	วันและเวลาที่ยืมหนังสือ
4	TRDATE	วันที่ยืมหนังสือ
5	TRTIME	เวลาในการยืมหนังสือ
6	TRTYPE	รูปแบบในการให้บริการ
7	TRUSER	ชื่อเจ้าหน้าที่ที่ให้บริการยืมหนังสือ
8	LOCATION	สถานที่ยืมหนังสือ

ตารางที่ 3.2

แสดงรายละเอียดข้อมูลที่ใช้ภายในงานวิจัย

Col#	Name	Description
1	ITEMCOUNT	จำนวนวัสดุสารนิเทศที่กำลังยืม
2	TDATE	วันและเวลาที่ยืมหนังสือ
3	TRDATE	วันที่ยืมหนังสือ
4	TRTIME	เวลา
5	TRTYPE	รูปแบบในการให้บริการ
6	TRUSER	สถานที่คืน
7	LOCATION	หน่วยงานที่สังกัด
8	USER_STATUS	สถานะสมาชิกห้องสมุด
9	USER_YEAR	ชั้นปี
10	ITEMCOUNT	จำนวนวัสดุสารนิเทศที่กำลังยืม

3.7.4 ช่วงสร้างแบบจำลอง (Modeling Phase)

ผู้วิจัยเลือกลักษณะของการทำเหมืองข้อมูล ดังนี้

1. กฎความสัมพันธ์ (Association Rule) เพื่อค้นหารูปแบบความสัมพันธ์ของพฤติกรรม การยืมหนังสือที่มีโอกาสเกิดพฤติกรรมการยืมหนังสือหมวดอื่นในอนาคต เพื่อที่จะได้ จัดซื้อหรือจัดหาหนังสือเพื่อรองรับต่อความต้องการใช้งานของสมาชิกห้องสมุด
2. กฎการแบ่งกลุ่ม (Clustering) เพื่อทำการแบ่งกลุ่มของสมาชิกห้องสมุด ที่มีลักษณะ การยืมหนังสือประจำหรือพฤติกรรมการยืมหนังสือคล้ายกันให้อยู่กลุ่มเดียวกัน และ สมาชิกห้องสมุดที่มีลักษณะประจำหรือพฤติกรรมการยืมหนังสือต่างกันให้อยู่คนละ กลุ่มกัน โดยผู้วิจัยใช้ขั้นตอนวิธี แบ่งกลุ่ม “K-Mean Clustering”

3.7.5 ช่วงประเมินผล (Evaluation Phase)

ประเมินแบบจำลองที่ได้จากการทำเหมืองข้อมูลเพื่อพิจารณาถึงความเหมาะสม ในการนำแบบจำลองไปประยุกต์ใช้ว่าแบบจำลองที่ได้นั้นมีความแม่นยำในการทำนายมากน้อย เพียงใด โดยในการประเมินแบบจำลอง นั้นจะแบ่งตามลักษณะของการทำเหมืองข้อมูล เช่น กฎ ความสัมพันธ์สามารถประเมินผลที่ได้จากการพิจารณาค่าความเชื่อมั่น และค่าสนับสนุน ส่วนการ แบ่งกลุ่มสามารถประเมินผลที่ได้จากระยะห่างจากจุดศูนย์กลางของกลุ่ม เป็นต้น

3.7.6 ช่วงนำไปใช้งาน (Deployment Phase)

นำแบบจำลองที่ได้จากการทำเหมืองข้อมูล เป็นข้อมูลสำหรับการกำหนดแนวทาง ในการดำเนินงานภายในห้องสมุด เช่น การแนะนำหนังสือใหม่ การกำหนดนโยบายหนังสือสำรอง การจัดชั้นเรียงหนังสือ การจัดกิจกรรมส่งเสริมรักการอ่าน เป็นต้น