

หัวข้อการค้นคว้าอิสระ	การประยุกต์ใช้ทฤษฎีเหมืองข้อมูลในบริการยืมคืนของห้องสมุด กรณีศึกษา สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ Applying Data Mining for Service in Library Case Study Thammasat University Library
ชื่อผู้เขียน	นายดุลยรัตน์ กรณธ์แสง Dulyarat Gronsang
แผนกวิชา/คณะ	สาขาวิชาการบริหารเทคโนโลยี วิทยาลัยนวัตกรรม มหาวิทยาลัยธรรมศาสตร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ดร. กมล เกียรติเรืองกมล
ปีการศึกษา	2552

บทสรุป

การค้นคว้าอิสระฉบับนี้ได้ศึกษาการประยุกต์ใช้ทฤษฎีเหมืองข้อมูลในบริการยืมคืนของห้องสมุด กรณีศึกษา สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ โดยในปัจจุบันห้องสมุดมีทรัพยากรเป็นจำนวนมหาศาล ทำให้ผู้ยืมหนังสือไม่สามารถใช้ทรัพยากรของห้องสมุดได้อย่างมีประสิทธิภาพ ดังนั้นจึงจำเป็นต้องค้นหากฎความสัมพันธ์ (Association Rule) และการจัดกลุ่ม (Cluster) ของข้อมูลการยืมคืนหนังสือ ซึ่งเป็นการเพิ่มศักยภาพของให้ห้องสมุดสามารถใช้ทรัพยากรที่มีอยู่ได้อย่างมีประสิทธิภาพมากขึ้น

งานวิจัยนี้เป็นงานวิจัยเชิงทดลองเพื่อค้นหารูปแบบและความรู้ที่ซ่อนอยู่ภายในข้อมูลการยืมคืนหนังสือระหว่างเดือนมกราคมถึงเดือนมิถุนายน พ.ศ.2552 ซึ่งวิเคราะห์ข้อมูลตามแบบจำลอง CRISP-DM โดยนำเทคนิคในการทำเหมืองข้อมูล (Data Mining) มาช่วยในการวิเคราะห์ข้อมูลเพื่อ

1) หาความสัมพันธ์ระหว่างหนังสือที่ถูกยืมกับสมาชิกห้องสมุดที่ยืมหนังสือนั้นด้วย Association Algorithm เมื่อพิจารณาความสัมพันธ์ระหว่างหนังสือที่ถูกยืมกับสมาชิกห้องสมุดที่ยืมหนังสือนั้นจะพบกฎอยู่ 3 กฎที่มีความเชื่อมั่นมากกว่า 80% และบรรณารักษ์มีความสนใจพบว่า

- หากยืมหนังสือหมวดรัฐศาสตร์และหมวดกฎหมายแล้วจะยืมหนังสือหมวดการศึกษา

- หากยืมหนังสือหมวดกฎหมายแล้วจะยืมหนังสือหมวดรัฐศาสตร์
- หากยืมหนังสือหมวดรัฐศาสตร์และหมวดเทคโนโลยีการจัดการแล้วจะยืมหนังสือหมวดวิทยาศาสตร์

2) แบ่งกลุ่มผู้ยืมหนังสือตามลักษณะของผู้ยืมและเลขหมู่ของหนังสือที่ยืมด้วย Clustering Algorithm ซึ่งสามารถแบ่งได้เป็น 4 กลุ่ม

ผู้วิจัยประมวลผลข้อมูลตามขั้นตอนของ CRISP-DM ซึ่งขั้นตอนของ CRISP-DM มีดังนี้

ขั้นตอนที่ 1 การทำความเข้าใจในธุรกิจ (Business/Research Understanding Phase)

สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ เป็นหน่วยงานในมหาวิทยาลัยธรรมศาสตร์ ที่จัดตั้งขึ้นมาเพื่อรองรับการบริการทางด้านการศึกษแก่นักศึกษา อาจารย์ ข้าราชการ ตลอดจนนักวิจัยของมหาวิทยาลัยธรรมศาสตร์ โดยปัจจุบัน สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ ประสบปัญหาในการจัดสรรทรัพยากรสารสนเทศแก่สมาชิกห้องสมุด เช่น สมาชิกค้นหาหนังสือที่ตนต้องการไม่เจอ ทำให้ต้องค้นชั้นเรียงหนังสือบ่อย เป็นผลให้ชั้นเรียงหนังสือกระจัดกระจายไม่เป็นระเบียบ หรือสมาชิกห้องสมุดไม่สามารถยืมหนังสือได้ เนื่องจากหนังสือที่ตนต้องการอยู่ที่ห้องสมุดอื่น งานวิจัยนี้จึงมีเป้าหมายเพื่อสร้างแบบจำลองในการพยากรณ์ความน่าจะเป็นในการยืมหนังสือของนักศึกษาจากคณะต่าง ๆ เพื่อเป็นแนวทางในการปรับปรุงการให้บริการยืมคืนหนังสือตลอดจนเป็นแนวทางในการดำเนินงานของสำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์

ขั้นตอนที่ 2 การทำความเข้าใจข้อมูล (Data Understanding Phase)

ข้อมูลที่ผู้วิจัยนำมาศึกษาแบ่งออกเป็น 3 ลักษณะคือ

1. ข้อมูลธุรกรรมการยืมคืนหนังสือ เป็นข้อมูลที่บ่งบอกว่าลักษณะของธุรกรรมการยืมคืนหนังสือ เช่น รหัสสมาชิก รหัสหนังสือ วันและเวลาที่ยืมหนังสือ เป็นต้น
2. ข้อมูลสมาชิกห้องสมุด เป็นข้อมูลที่บ่งบอกลักษณะของสมาชิกห้องสมุด เช่น ระดับการศึกษา และคณะ เป็นต้น
3. ข้อมูลหนังสือ เป็นข้อมูลที่บ่งบอกว่าลักษณะของหนังสือ เช่น หมวดหนังสือ เป็นต้น

ขั้นตอนที่ 3 การเตรียมข้อมูล (Data Preparation Phase)

- **คัดเลือกข้อมูล** ผู้วิจัยได้นำข้อมูลจากคลังข้อมูลบางส่วนไปทำเหมืองข้อมูล ข้อมูลดังกล่าวคือข้อมูลการยืมคืนที่องค์กรได้ทำการเก็บรวบรวมไว้ในคลังข้อมูลเป็นช่วงเวลา 2 ไตรมาส

คือ ระหว่างไตรมาสที่ 1 และไตรมาสที่ 2 ของ ปี พ.ศ.2552 (มกราคม พ.ศ. 2552 ถึง มิถุนายน พ.ศ. 2552)

- **ทำความสะอาดข้อมูล** ผู้วิจัยทำการจัดข้อมูลให้อยู่ในรูปแบบที่เหมาะสมก่อนทำเหมืองข้อมูล เช่น แก้ไขค่าว่างของข้อมูลโดยอาจใส่ค่า 0 ลงไป นอกจากนี้ผู้วิจัยยังไม่นำข้อมูลบางแถวที่มีข้อมูลไม่ครบถ้วนมาทำเหมืองข้อมูล เช่น ขาดรหัสหนังสือ หรือ ขาดรหัสผู้ยืมหนังสือ

- **จัดเตรียมข้อมูล** ผู้วิจัยเตรียมข้อมูลก่อนการวิเคราะห์เพื่อให้ข้อมูลเหมาะสมกับเทคนิคที่นำมาวิเคราะห์ ขั้นตอนนี้เป็นขั้นตอนที่สำคัญมากเพื่อให้การวิเคราะห์ข้อมูลมีคุณภาพโดยทั่วไปแล้วขั้นตอนนี้ส่วนมากจะใช้เวลานาน แต่ข้อมูลที่นำมาจากสำนักหอสมุดมหาวิทยาลัยธรรมศาสตร์ได้มีการทำความสะอาดข้อมูล (Data Cleaning) มาแล้วระดับหนึ่งเนื่องจากองค์กรได้จัดเก็บข้อมูลในคลังข้อมูล (Data Warehouse) จึงทำให้ใช้เวลาในกระบวนการนี้ลดลง เนื่องจากการทำคลังข้อมูลได้มีการทำความสะอาดข้อมูลมาแล้วระดับหนึ่งแล้ว จากนั้นผู้วิจัยจึงทำการแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมต่อการวิเคราะห์ และสอดคล้องกับวัตถุประสงค์ที่ต้องการ

- **ผสานข้อมูล** ผู้วิจัยลดขนาดข้อมูล และลดจำนวนมุมมองของข้อมูลที่ไม่จำเป็น นอกจากนี้ยังเปลี่ยนรูปแบบให้ง่ายต่อการวิเคราะห์ โดยการเลือกข้อมูลเฉพาะข้อมูลที่สนใจ ดังนั้นจึงมีการตัดสมรรถที่ไม่จำเป็นในการวิเคราะห์มีดังนี้

ขั้นตอนที่ 4 การสร้างแบบจำลองตัวอย่าง (Modeling Phase)

ผู้วิจัยเลือกลักษณะของการทำเหมืองข้อมูล ดังนี้

- **กฎความสัมพันธ์ (Association Rule)** เพื่อค้นหารูปแบบความสัมพันธ์ของพฤติกรรม การยืมหนังสือที่มีโอกาสเกิดพฤติกรรมการยืมหนังสือหมวดอื่นในอนาคต เพื่อที่จะได้จัดซื้อหรือจัดหาหนังสือเพื่อรองรับต่อความต้องการใช้งานของสมาชิกห้องสมุดซึ่งผลการวิเคราะห์สามารถสรุปได้ว่าเกิดกฎความสัมพันธ์ที่มีค่าความเชื่อมั่นมากกว่าร้อยละ 60

- **กฎการแบ่งกลุ่ม (Clustering)** เพื่อแบ่งกลุ่มสมาชิกห้องสมุดที่มีลักษณะการยืมหนังสือประจำหรือพฤติกรรมการยืมหนังสือคล้ายกันให้อยู่กลุ่มเดียวกัน และสมาชิกห้องสมุดที่มีลักษณะประจำหรือพฤติกรรมการยืมหนังสือต่างกันให้อยู่คนละกลุ่มกัน โดยผู้วิจัยใช้ขั้นตอนวิธีแบ่งกลุ่ม “K-Mean Clustering” ซึ่งผลการวิเคราะห์สามารถสรุปได้ดังนี้

- **ส่วนใหญ่สมาชิกห้องสมุดกลุ่มที่ 1** ประกอบด้วยนักศึกษาปริญญาตรีสังกัดคณะสายสังคมศาสตร์ เช่น คณะศิลปศาสตร์ คณะนิติศาสตร์ และคณะรัฐศาสตร์

มีความสนใจหนังสือทั่วไปภาษาไทยหมวดสังคมศาสตร์ หมวดภาษาศาสตร์และประวัติศาสตร์ทั่วไป โดยยืมประมาณ 1-3 เล่มต่อสัปดาห์ ในวันธรรมดา (จันทร์ถึงศุกร์) เวลาราชการ (8.00-16.00 น.)

- **ส่วนใหญ่สมาชิกห้องสมุดกลุ่มที่ 2** ประกอบด้วยนักศึกษาปริญญาตรีสังกัดคณะสายวิทยาศาสตร์ เช่น คณะวิทยาศาสตร์ คณะวิศวกรรมศาสตร์ คณะแพทยศาสตร์ มีความสนใจหนังสือทั่วไปภาษาไทยหมวดวิทยาศาสตร์ หมวดเทคโนโลยี และหมวดทางการแพทย์ โดยยืมประมาณ 2-3 เล่มต่อสัปดาห์ ในวันจันทร์ อังคาร และศุกร์) เวลากลางวันถึงเย็น (11.00-16.00 น.)
- **ส่วนใหญ่สมาชิกห้องสมุดกลุ่มที่ 3** ประกอบด้วยนักศึกษาปริญญาโท ปริญญาเอก และประกาศนียบัตร สังกัดคณะนิติศาสตร์ พาณิชยศาสตร์และการบัญชี และวิทยาศาสตร์ มีความสนใจหนังสือทั่วไปภาษาไทยหมวดสังคมศาสตร์ กฎหมาย และเทคโนโลยี โดยยืมประมาณ 2-4 เล่มต่อสัปดาห์ ในวันหยุดเสาร์ อาทิตย์ เวลากลางวันถึงหัวค่ำ (11.00-20.00 น.)
- **โดยส่วนใหญ่สมาชิกห้องสมุดกลุ่มที่ 4** ประกอบด้วยอาจารย์และข้าราชการ สังกัดคณะนิติศาสตร์ คณะวิทยาศาสตร์ และคณะพาณิชยศาสตร์และการบัญชี มีความสนใจหนังสือทั่วไปภาษาไทยหมวดสังคมศาสตร์ หมวดภาษาศาสตร์ หมวดวิทยาศาสตร์ โดยยืมประมาณ 2-4 เล่มต่อสัปดาห์ ในวันธรรมดา (จันทร์ถึงศุกร์) เวลากลางวันถึงหัวค่ำ (11.00-20.00 น.)

ขั้นตอนที่ 5 การประเมินผล (Evaluation Phase)

ผู้วิจัยประเมินแบบจำลองตามลักษณะของการทำเหมืองข้อมูลดังนี้

- **กฎความสัมพันธ์ (Association Rule)** แบบจำลองที่ได้จากการทำเหมืองข้อมูลโดยกฎความสัมพันธ์สามารถประเมินผลที่ได้จากการพิจารณาค่าความเชื่อมั่น (Confidence) ของกฎความสัมพันธ์ โดยจะนำกฎความสัมพันธ์ที่มีค่าความเชื่อมั่นสูงกว่าร้อยละ 80
- **การแบ่งกลุ่ม (Clustering)** แบบจำลองที่ได้จากการแบ่งกลุ่มสามารถประเมินผลที่ได้จากระยะห่างจากจุดศูนย์กลางของกลุ่ม ซึ่งพบว่าระยะห่างจากจุดศูนย์กลางของกลุ่มทุกกลุ่มมีค่าความคลาดเคลื่อนมาตรฐานไม่เกิน 0.003 นอกจากกลุ่ม 4 ที่มีค่าความคลาดเคลื่อนมาตรฐานที่สูงกว่ากลุ่มอื่นคือ 0.04

ขั้นตอนที่ 6 การนำไปใช้งาน (Deployment Phase)

การศึกษาพฤติกรรมการยืมหนังสือของสมาชิกห้องสมุดถือเป็นกุญแจสำคัญที่ทำให้ผู้บริหารห้องสมุดทราบถึงพฤติกรรมการยืมหนังสือของสมาชิกห้องสมุด ตลอดจนความต้องการใช้หนังสือของสมาชิกห้องสมุดโดยบรรณารักษ์สามารถนำข้อมูลเกี่ยวกับพฤติกรรมของสมาชิกห้องสมุดมาใช้ประโยชน์ในหลาย ๆ ด้าน ดังนี้

- **กฎความสัมพันธ์** จากการทำเหมืองข้อมูลของธุรกรรมการยืมคืนหนังสือด้วยกฎความสัมพันธ์ สามารถนำมาประยุกต์กับการทำงานภายในห้องสมุดได้ดังนี้

- **การจัดเรียงชั้นหนังสือ** เพื่อให้การบริการของห้องสมุดเป็นไปอย่างมีประสิทธิภาพมากยิ่งขึ้น บรรณารักษ์สามารถนำข้อมูลจากการวิเคราะห์รูปแบบการยืมหนังสือมาใช้ในการจัดเรียงชั้นหนังสือให้กับสมาชิกของห้องสมุดโดยจัดวางหนังสือใหม่ในตำแหน่งที่ใกล้เคียงกันในกรณีที่หมวดของหนังสือนั้นมีความสัมพันธ์กัน เช่น จากกฎความสัมพันธ์ที่ว่า ถ้าร้อยละ 95 ของสมาชิกห้องสมุดที่ยืมหนังสือหมวดรัฐศาสตร์ และหมวดกฎหมาย แล้วจะยืมหนังสือหมวดการศึกษาด้วย บรรณารักษ์สามารถนำข้อมูลรูปแบบการยืมข้างต้นมาใช้ประโยชน์โดยจัดเรียงหนังสือหมวดรัฐศาสตร์ หมวดกฎหมาย และหมวดการศึกษาให้อยู่ใกล้กัน
- **การปรับปรุงการสืบค้นหนังสือ** บรรณารักษ์สามารถปรับปรุงการสืบค้นหนังสือในระบบห้องสมุดอัตโนมัติ ให้สามารถแนะนำหมวดหมู่หนังสือที่เกี่ยวข้อง กล่าวคือปรับปรุงระบบสืบค้นหนังสือให้สามารถแนะนำสมาชิกห้องสมุดที่กำลังสืบค้นหนังสือว่ามีหมวดหมู่ที่มีความสัมพันธ์กับหมวดหมู่ที่สมาชิกห้องสมุดกำลังค้นหาอะไรบ้าง

- **กฎการแบ่งกลุ่ม** จากการทำเหมืองข้อมูลของข้อมูลธุรกรรมการยืมคืนหนังสือด้วยกฎการแบ่งกลุ่ม สามารถนำมาประยุกต์กับการทำงานภายในห้องสมุดได้ดังนี้

- **การแนะนำหนังสือใหม่** บรรณารักษ์สามารถแจ้งข้อมูลไปยังกลุ่มเป้าหมายได้อย่างถูกต้อง เช่น จากกฎการแบ่งกลุ่มที่ว่า นักศึกษาระดับปริญญาตรีสังกัดคณะศิลปศาสตร์ คณะนิติศาสตร์ และคณะรัฐศาสตร์ มีความสนใจหนังสือทั่วไป ภาษาไทยหมวดสังคมศาสตร์ หมวดภาษาศาสตร์และประวัติศาสตร์ ทำให้บรรณารักษ์ทราบว่าหากห้องสมุดมีหนังสือใหม่ที่เป็นหมวดสังคมศาสตร์ หมวดภาษาศาสตร์และประวัติศาสตร์ ต้องแจ้งไปยังนักศึกษาระดับปริญญาตรีสังกัด

คณะศิลปศาสตร์ คณะนิติศาสตร์ และคณะรัฐศาสตร์ ซึ่งบรรณารักษ์สามารถนำมาประยุกต์ใช้กับระบบห้องสมุดอัตโนมัติในส่วนของคุณข้อมูลผู้ยืม

- **การจัดกิจกรรมส่งเสริมการอ่าน** ในการจัดกิจกรรมส่งเสริมการอ่านนั้น บรรณารักษ์สามารถกำหนดช่วงวัน เวลา และเนื้อหาของกิจกรรมให้สอดคล้องกับกลุ่มเป้าหมายของกิจกรรมส่งเสริมการอ่านได้อย่างถูกต้อง เช่น จากกฎการแบ่งกลุ่มที่ว่า นักศึกษาระดับปริญญาโท ปริญญาเอก และประกาศนียบัตร ชอบยืมในวันหยุดเสาร์อาทิตย์ เวลากลางวันถึงหัวค่ำ (11.00-20.00 น.) ทำให้บรรณารักษ์ทราบว่า หากต้องการจัดกิจกรรมส่งเสริมการอ่านให้นักศึกษาปริญญาโท ต้องจัดในช่วงวันและเวลาดังกล่าว

การศึกษาวิจัยเรื่อง “การนำทฤษฎีเหมืองข้อมูลมาประยุกต์ใช้ในบริการยืมคืนของห้องสมุด กรณีศึกษา สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์” ในครั้งนี้ ผู้วิจัยนำข้อมูลธุรกรรมการยืมคืนหนังสือ ที่สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์จัดเก็บไว้แล้วมาประมวลผลด้วยขั้นตอนของ CRISP-DM เพื่อสร้างแบบจำลองพฤติกรรมการยืมหนังสือ โดยแบบจำลองที่ได้จะแสดงรูปแบบความสัมพันธ์ของพฤติกรรมการยืมหนังสือของผู้ยืมหนังสือ นอกจากนี้แบบจำลองยังสามารถแบ่งกลุ่มผู้ยืมหนังสือจากลักษณะประจำ หรือพฤติกรรมผู้ใช้ ซึ่งจะเป็นแนวทางในการปรับปรุงการให้บริการยืมคืนวัสดุสารนิเทศให้มีประสิทธิภาพให้มากยิ่งขึ้น