

การปรับปรุงเทคนิคการทำเหมืองข้อมูลเพื่อความเข้าใจพฤติกรรมการยืมหนังสือในห้องสมุดที่ดีขึ้น

Modification of Data Mining Technique for Better Understanding of Book-Loan Behaviors

อาจารย์ ดร.นพพน เลิศชูวงศา
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์
ผู้ช่วยศาสตราจารย์ ดร. นิพนธ์ ปริญาวุฒิชัย
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏราชนครินทร์

Abstract

Improving information source management in libraries has been one of the most prominent issues and among researchers' interest in the recent decades. However, the analyses leading to such improvement has their limitation as they usually focus on only some perspective and lack considerations of the characteristics of the borrowers. This article, hence, proposes a new concept that improves Apriori by integrating Wavelet and Histogram into the analysis of association rules. As a result, the researchers identify the association rules that explain behaviors of the borrowers as per person, instead of per borrowing transaction as found in most research. Not only this paper helps administrators design and improve information source management in the libraries, but it also discovers a novel concept that strengthens the abilities of a data mining technique using Apriori.

Keywords : Data Mining, Apriori, Wavelet, Histogram, Library Management

บทคัดย่อ

การเพิ่มประสิทธิภาพการบริหารจัดการทรัพยากรสารสนเทศในห้องสมุดถือเป็นเรื่องสำคัญและอยู่ภายใต้ความสนใจของนักวิจัยตลอดหลายทศวรรษที่ผ่านมา อย่างไรก็ตามการวิเคราะห์ข้อมูลเพื่อนำไปสู่การออกแบบกลยุทธ์ในการเพิ่มประสิทธิภาพดังกล่าวในปัจจุบันยังมีข้อจำกัด โดยข้อมูลที่นำมาวิเคราะห์นั้นไม่เพียงพอหรือเป็นเพียงแง่มุมใดแง่มุมหนึ่งเพียงด้านเดียว อีกทั้งยังขาดความเชื่อมโยงกับคุณสมบัติต่าง ๆ ของตัวผู้ยืมอีกด้วย บทความนี้จึงนำเสนอแนวคิดในการปรับปรุงกฎความสัมพันธ์ตามขั้นตอนเอไพริออริพื้นฐานโดยอาศัยเครื่องมืออย่างแบบจำลองทางเวลาด้วยเวฟเลตและกราฟแจกแจงความถี่ ทำให้คณะนักวิจัยค้นพบกฎความสัมพันธ์ที่อธิบายพฤติกรรมของการยืม “รายบุคคล” แทนรูปแบบ “รายรายการ” เหมือนกรณีงานวิจัยทั่วไป ผลลัพธ์ที่ได้จากการศึกษาในครั้งนี้นอกจากจะทำให้ผู้บริหารห้องสมุดกำหนดกลยุทธ์และงบประมาณได้อย่างเหมาะสมแล้ว ยังเป็นการเพิ่มประสิทธิภาพการสร้างกฎความสัมพันธ์ตามขั้นตอนเอไพริออริอีกด้วย

คำสำคัญ: เหมืองข้อมูล เอไพริออริ เวฟเลต กราฟแจกแจงความถี่ การจัดการห้องสมุด

ความเป็นมาและความสำคัญของปัญหา

ห้องสมุดถือเป็นแกนสำคัญในการขับเคลื่อนองค์ความรู้ขององค์กรและปัจเจกบุคคล โดยเฉพาะอย่างยิ่ง ในยุคปัจจุบันที่องค์ความรู้ด้านวิทยาศาสตร์ เทคโนโลยี และสื่อสารสนเทศมีการเปลี่ยนแปลงอย่างรวดเร็ว และก้าวหน้าอย่างไม่เคยเป็นมาก่อน (Long and Wu, 2012) หน้าที่สำคัญของห้องสมุด คือ การให้การบริการยืม-คืนทรัพยากรสารสนเทศ จึงมีใช้เรื่องแปลกประการใด

ที่ห้องสมุดจะต้องเผชิญกับการจัดการข้อมูลที่เกี่ยวข้องทั้งข้อมูลการบริการยืม-คืนทรัพยากรสารสนเทศ ในรูปแบบของหนังสือ วารสาร งานวิจัยทั้งประเภทกระดาษและอิเล็กทรอนิกส์ ข้อมูลจัดซื้อจัดจ้าง ข้อมูลทะเบียนทรัพยากรสารสนเทศ และข้อมูลของผู้รับบริการ จำนวนมหาศาล (สุภาพร ชัยธัมมะปกรณ์, 2014) ซึ่งแน่นอนที่สุดย่อมต้องอาศัยกลยุทธ์ทางการบริหารจัดการที่มีประสิทธิภาพสูง

ที่จะทำให้ผู้ใช้ในยุคปัจจุบันเกิดความพึงพอใจต่อการบริการยืม-คืนทรัพยากรสารสนเทศจากห้องสมุด ทั้งในแง่ของความสะดวกสบายในการค้นหาทรัพยากรสารสนเทศ การจัดวางตำแหน่งทรัพยากรสารสนเทศตามชั้นวางที่เหมาะสม รูปแบบของทรัพยากรสารสนเทศที่เข้ากับวิถีชีวิตของผู้ใช้ เป็นต้น (Rui, 2016; Uppal and Chindwani, 2003)

ในอดีตความรู้สึกของบรรณารักษ์เพียงลำพังอาจเพียงพอต่อการออกแบบกลยุทธ์ในการบริหารจัดการห้องสมุด ในปัจจุบันปัจจัยและองค์ประกอบมีความซับซ้อนกว่าเดิมมาก ทำให้ผู้บริหารห้องสมุดโดยเฉพาะอย่างยิ่งในมหาวิทยาลัยต่างๆ จำเป็นต้องนำข้อมูลจากแหล่งต่างๆ มาประกอบการตัดสินใจ หรือจัดทำรายงานที่มีเหตุและผลเพื่อเสนอขอของบประมาณในการปรับปรุงห้องสมุดให้มีความทันสมัย ตอบโจทย์ความต้องการของผู้ใช้ในมหาวิทยาลัย

วัตถุประสงค์การวิจัย

บทความนี้ได้นำเสนอรูปแบบการวิเคราะห์ข้อมูลด้วยเทคนิคการทำเหมืองข้อมูลที่พัฒนาต่อยอดจากการค้นหาความสัมพันธ์ตามขั้นตอนวิธีเอปรีออริ (The Apriori Algorithm) โดยมีวัตถุประสงค์ดังนี้

1. เพื่อค้นหาความสัมพันธ์ของพฤติกรรมการยืมตาม ‘รายชื่อผู้ใช้’ แทนการค้นหาความสัมพันธ์ของการยืมตาม ‘รายการยืม’ หนังสือของผู้ใช้ที่นิยมใช้กันโดยทั่วไป (Ru et al., 2015; Uppal and Chindwani, 2013)
2. ประยุกต์เทคนิคการทำเหมืองข้อมูลโดยอาศัยเครื่องมือทางคณิตศาสตร์และสถิติอย่างกระบวนการเวฟเลต (Wavelet) (Rui, 2016) และกราฟแจกแจงความถี่ที่แสดงความสัมพันธ์ระหว่างข้อมูลเป็นหมวดหมู่ (Histogram) เพื่อค้นหารูปแบบการยืมหนังสือตามรายชื่อผู้ใช้ และ
3. เพื่อนำเสนอแนวคิดในการปรับปรุงรูปแบบการทำเหมืองข้อมูลและเพื่อกำหนดกลยุทธ์ที่มีประสิทธิภาพให้แก่ผู้บริหารของมหาวิทยาลัยในการปรับปรุงคุณภาพการบริการของห้องสมุดให้เข้ากับพฤติกรรมของผู้ใช้ต่อไปในอนาคต โดยอ้างอิงกับข้อมูลการยืมหนังสือจริงของนักศึกษาแต่ละรายในหนึ่งปีการศึกษา (ปีการศึกษา 2558) ของมหาวิทยาลัยราชภัฏราชนครินทร์

ทบทวนวรรณกรรม

ในหัวข้อนี้ได้เริ่มต้นจากการทบทวนเกี่ยวกับบทบาทห้องสมุดในปัจจุบันและวรรณกรรมที่เกี่ยวข้องกับการวางแผนการบริหารห้องสมุดให้เกิดประโยชน์ หลังจากนั้นจะเป็นการอธิบายหลักการทำเหมืองข้อมูลซึ่งถูกนำมาใช้เป็นหลักในการวิเคราะห์รูปแบบและพฤติกรรมการยืมหนังสือในหอสมุดกลาง มหาวิทยาลัยราชภัฏราชนครินทร์ ซึ่งเป็นกรณีศึกษาของการวิจัยในครั้งนี้

บทบาทของห้องสมุด

โดยมากแล้ว เมื่อกล่าวถึงห้องสมุด บุคคลทั่วไปมักคิดถึงห้องสมุดในโรงเรียน หรือมหาวิทยาลัย เนื่องจากห้องสมุดถือเป็นแกนสำคัญในการขับเคลื่อนองค์ความรู้โดยเฉพาะอย่างยิ่ง ในยุคปัจจุบันที่องค์ความรู้ด้านวิทยาศาสตร์ เทคโนโลยี และสารสนเทศมีการเปลี่ยนแปลงอย่างรวดเร็ว และก้าวหน้าอย่างไม่เคยเป็นมาก่อน (Long and Wu, 2012) โดยในปัจจุบัน ห้องสมุดก็ได้ตั้งอยู่ในเฉพะสถานศึกษาเท่านั้น แต่ยังคงปรากฏอยู่ทั่วไปในหน่วยงานต่างๆ ทั้งของภาครัฐและเอกชน หน้าที่สำคัญของห้องสมุด คือ การให้บริการยืม-คืนทรัพยากรสารสนเทศ ดังนั้นจึงมิใช่เรื่องแปลกประการใด ที่ห้องสมุดจะต้องเผชิญกับการจัดการข้อมูลที่เกี่ยวข้องทั้งข้อมูลการบริการยืม-คืนทรัพยากรสารสนเทศ ทั้งหนังสือ วารสาร งานวิจัยทั้งประเภทกระดาษ และอิเล็กทรอนิกส์ ข้อมูลจัดซื้อจัดจ้าง ข้อมูลทะเบียนทรัพยากรสารสนเทศ และข้อมูลของผู้รับบริการ จำนวนมหาศาล (สุภาพร ชัยธัมมะปกรณ์, 2014) การศึกษาวิจัยเพื่อค้นหารูปแบบความต้องการของผู้ใช้บริการห้องสมุดผ่านข้อมูลที่มีกักขังอยู่ในรูปแบบของฐานข้อมูลในระบบสารสนเทศ จึงเป็นหนึ่งในเรื่องสำคัญของแวดวงวิจัยด้านสารสนเทศศาสตร์ เพราะนั่นจะเป็นการช่วยผู้บริหารห้องสมุดกำหนดกลยุทธ์และแนวทางในการเพิ่มประสิทธิภาพการใช้ทรัพยากรสารสนเทศของห้องสมุด (Kularbphetong et al., 2014; Long and Wu, 2012) เช่น การกำหนดแนวทางในเพิ่มจำนวนการยืม-คืนทรัพยากรสารสนเทศของห้องสมุด (Castro-Gessner et al., 2013) หรือแม้กระทั่งข้อเสนอแนะในการปรับปรุงหลักสูตร (Reeb and Gibbons, 2004) ต่อไปในอนาคตได้อีกด้วย

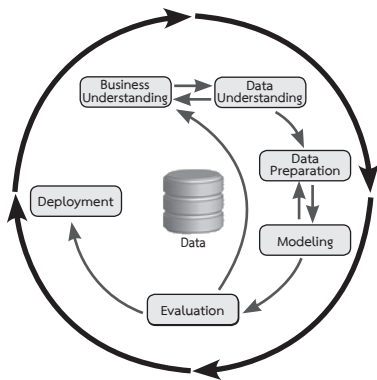
สารสนเทศศาสตร์กับการวิเคราะห์กฎความสัมพันธ์

การประยุกต์ใช้สารสนเทศศาสตร์ในการเพิ่มประสิทธิภาพการจัดการห้องสมุดสามารถกระทำได้หลายแนวทาง หนึ่งในแนวทางที่นักวิจัยนิยมใช้ คือ การทำเหมืองข้อมูล (Data Mining) ซึ่งในบางครั้งอาจเรียกว่า การค้นพบความรู้ (Knowledge Discovery) (Ahirwal et al., 2012) เหมืองข้อมูล หมายถึงกระบวนการนำข้อมูลจำนวนมากที่เกี่ยวข้องกับระบบบริหารจัดการทรัพยากรสารสนเทศของห้องสมุดที่เกิดขึ้นในแต่ละวันซึ่งโดยมากจะบรรจุอยู่ในรูปแบบของฐานข้อมูล หรือคลังข้อมูล (Data Warehouse) มาประมวลผล เพื่อค้นหารูปแบบ หรือกฎความสัมพันธ์ระหว่างองค์ประกอบที่เกี่ยวข้องเพื่อประโยชน์ในการวางแผนบริหารจัดการให้มีประสิทธิภาพ (สุภาพร ชัยธัมมะปกรณ์, 2014) ตัวอย่างของข้อมูลที่เกี่ยวข้องกับระบบบริหารจัดการทรัพยากรสารสนเทศของห้องสมุดที่สามารถนำไปใช้ในการวางแผนบริหารจัดการห้องสมุด ได้แก่ ข้อมูลประวัติการบริการทรัพยากร

สารสนเทศ ข้อมูลการสืบค้นทรัพยากรสารสนเทศ หรือประวัติการเข้าถึงฐานข้อมูลต่างๆ ของห้องสมุด เป็นต้น (สุภาพร ชัยธัมมะปกรณ์, 2014) ซึ่งโดยพื้นฐานแล้ว ผู้บริหารสามารถใช้สถิติในการประเมิน หรือรับรู้สถานการณ์ปัญหาอย่างง่ายจากข้อมูลเหล่านี้ได้ อย่างไรก็ตาม ด้วยเทคนิคการประมวลผลการทำเหมืองข้อมูล ผู้บริหารจะสามารถล่วงรู้รูปแบบการใช้บริการทรัพยากรสารสนเทศ หรือพฤติกรรมการใช้บริการของผู้ใช้บริการ นำไปสู่การวิเคราะห์ต่อไปได้ว่า ทรัพยากรสารสนเทศใดยังไม่ถูกนำไปใช้ประโยชน์อย่างเต็มศักยภาพ หรือทรัพยากรสารสนเทศใดมีความสัมพันธ์กับคุณสมบัติของกลุ่มผู้ใช้บริการเฉพาะกลุ่ม และสัมพันธ์กันด้วยเหตุผลใด ทั้งนี้ เพื่อให้การกำหนดกลยุทธ์การบริหารจัดการทรัพยากรสารสนเทศที่อยู่ในห้องสมุด รวมถึงการจัดซื้อทรัพยากรสารสนเทศที่จะดำเนินการต่อไปในอนาคตตอบโจทย์ความต้องการของผู้ใช้บริการมากที่สุด (สุภาพร ชัยธัมมะปกรณ์, 2014; Ahirwal et al., 2012; Mehta et al., 2013)

หลักการดำเนินงานขั้นพื้นฐานของการทำเหมืองข้อมูล

การทำเหมืองข้อมูลในปัจจุบันอาศัยแบบแนวคิดจากแบบจำลองการทำเหมืองข้อมูลของของคริสป์-ดีเอ็ม (CRISP-DM Model) เพื่อสกัดองค์ความรู้ หรือพัฒนาทฤษฎีความสัมพันธ์ระหว่างองค์ประกอบที่มีส่วนเกี่ยวข้องหรือสัมพันธ์กัน และนำความรู้ที่ได้ไปประยุกต์ใช้เพื่อก่อให้เกิดประโยชน์ต่อการให้บริการยืมทรัพยากรสารสนเทศของห้องสมุด ซึ่งขั้นตอนต่างๆ ของการทำเหมืองข้อมูลดังภาพที่ 1 จะประกอบไปด้วย (เอกลสิทธิ์, 2015)



ภาพที่ 1 The Data Mining Life Cycle
ที่มา Azevedo and Zantos (2008)

การทำความเข้าใจธุรกิจ (Business Understanding) หรือกระบวนการทำความเข้าใจเนื้องานที่จะวิเคราะห์ โดยกำหนดจุดมุ่งหมายและความต้องการของผู้ศึกษาวิจัย เช่น กรณีของห้องสมุด นักวิจัยอาจประเมินสถานการณ์ที่ผู้ใช้บริการไม่สามารถใช้ศักยภาพของห้องสมุดได้อย่างเต็มที่ เนื่องจากมีปริมาณการยืมทรัพยากรสารสนเทศที่จัดว่าอยู่ในเกณฑ์ต่ำ

การทำความเข้าใจข้อมูล (Data Understanding) เป็นกระบวนการแปลงจุดมุ่งหมายที่กำหนดไว้ในขั้นตอนการทำความเข้าใจธุรกิจ ให้อยู่ในรูปแบบของปัญหาของการทำเหมืองข้อมูล ตัวอย่างเช่น ปัญหาที่ห้องสมุดในสถาบันการศึกษาหนึ่งๆ มีปริมาณการยืมทรัพยากรสารสนเทศน้อย อาจเกิดจากการขาดระบบการนำเสนอ จัดเรียง และกำหนดรูปแบบของการเสนอแนะทรัพยากรสารสนเทศที่ไม่ดีพอ เป็นต้น จึงจำเป็นที่นักวิจัยจะต้องศึกษาความสัมพันธ์ของหน่วยข้อมูลต่างๆ ที่เกี่ยวข้องกับข้อมูล ประวัติการบริการทรัพยากรสารสนเทศ เช่น ประเภทของทรัพยากรสารสนเทศ ตำแหน่งของทรัพยากรสารสนเทศ ข้อมูลผู้ใช้บริการ และข้อมูลการใช้บริการยืมทรัพยากรสารสนเทศของผู้ใช้บริการ เป็นต้น

การเตรียมข้อมูล (Data Preparation) เป็นขั้นตอนการแปลงสภาพและคัดกรองของข้อมูลที่ได้รวบรวมมาให้เหมาะสมสำหรับการทำเหมืองข้อมูล โดยในขั้นตอนนี้อาจใช้ระยะเวลาในการดำเนินการยาวนานที่สุด

การสร้างแบบจำลอง (Modeling) เป็นการวิเคราะห์ข้อมูลที่ได้จัดเตรียมไว้ด้วยเทคนิคอย่าง การจำแนกประเภทข้อมูล การหาความสัมพันธ์ และการแบ่งกลุ่มข้อมูล เป็นต้น ทั้งนี้อาจมีการย้อนกลับไปยังขั้นตอนของการเตรียมข้อมูลอีก หากข้อมูลที่จัดเตรียมไว้ยังไม่สามารถนำไปใช้สร้างแบบจำลองตามเทคนิคต่างๆ ที่เลือกได้ โดยการสร้างแบบจำลองนี้อาจช่วยทำให้ค้นพบความสัมพันธ์ที่ “ซ่อนอยู่” ในเหมืองข้อมูลได้

การประเมินผล (Evaluation) เป็นการวัดประสิทธิภาพหรือประเมินผลการวิเคราะห์ข้อมูลจากการทำเหมืองข้อมูลด้วยเทคนิคต่าง ๆ เพื่อสร้างความน่าเชื่อถือให้กับการทำเหมืองข้อมูล

การทำไปใช้งาน (Deployment) เป็นกระบวนการนำแบบจำลองที่ค้นพบ ไปประยุกต์ใช้เพื่อแก้ไข หรือปรับปรุงสถานการณ์ปัญหาที่เกิดขึ้น

แบบจำลองจากเทคนิคการค้นหากฎความสัมพันธ์

การสร้างแบบจำลองด้วยเทคนิคการค้นหากฎความสัมพันธ์ (Association Rules) นับเป็นวิธีการสร้างแบบจำลองของการทำเหมืองข้อมูลวิธีการหนึ่งที่ได้รับนิยามเป็นอย่างมาก (วิภาวรรณบัวทอง, 2014) ทั้งนี้เทคนิคดังกล่าวจะค้นหาความสัมพันธ์ของข้อมูลจากเหมืองข้อมูลที่มีอยู่ เพื่อนำไปสังเคราะห์เป็นรูปแบบที่เกิดขึ้นบ่อย ๆ (Frequent pattern) เพื่อนำไปใช้ทำนายโอกาสที่จะเกิดรูปแบบเดิมซ้ำอีกในอนาคต ซึ่งโดยมากมักนำเสนอในรูปแบบของกฎความสัมพันธ์ที่สามารถเขียนได้ในรูปเซตของรายการที่เป็นเหตุไปสูเซตของรายการที่เป็นผล ทั้งนี้ในช่วงเริ่มต้นการใช้กฎความสัมพันธ์มักถูกนำมาใช้กับการวิเคราะห์ตะกร้าตลาด (Market

Basket Analysis) ก่อนจะถูกนำมาประยุกต์กับศาสตร์แขนงอื่นๆ อย่างที่พบเห็นในปัจจุบัน โดยการพัฒนากฎความสัมพันธ์ของอาศัยตัวเลข 2 ชุด คือ ค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) โดยค่าสนับสนุน คือ สัดส่วนของการเกิดขึ้นพร้อมๆ กันของรายการที่กำลังให้ความสนใจอยู่ต่อโอกาสทั้งหมดที่เป็นไปได้ ส่วนค่าความเชื่อมั่น คือ ตัววัดประสิทธิภาพของกฎความสัมพันธ์ เช่น ค่าความน่าเชื่อถือในการซื้อแอปเปิ้ลแล้วจะซื้อผ้าอ้อมเด็ก เกิดจากสัดส่วนของค่าสนับสนุน (“แอปเปิ้ล”, “ผ้าอ้อมเด็ก”) ต่อค่าสนับสนุน (“แอปเปิ้ล”) ซึ่งหากค่าความน่าเชื่อถือมีค่าเข้าใกล้ 1 ย่อมแสดงว่าองค์ประกอบภายใต้การสนับสนุนนั้น มีความสัมพันธ์กันอย่างใกล้ชิด เช่น หากค่าความน่าเชื่อถือในการซื้อแอปเปิ้ลแล้วจะซื้อผ้าอ้อมเด็กเท่ากับ 1 ย่อมหมายความว่า ทุกๆ ครั้งที่ซื้อแอปเปิ้ล ผู้ซื้อย่อมจะซื้อผ้าอ้อมเด็กด้วย

ขั้นตอนวิธีเอปพรอริ (The Apriori Algorithm) ซึ่งเป็นหนึ่งในขั้นตอนวิธีสำคัญที่ใช้ในการค้นหากฎความสัมพันธ์เกิดขึ้นในปี ค.ศ. 1994 โดย Rakesh Agrawal และ Ramakrishnan Srikant ผู้เป็นคณะนักวิจัยที่ร่วมกันปรับปรุงขั้นตอนวิธีในการค้นหากฎความสัมพันธ์ให้มีประสิทธิภาพยิ่งขึ้น โดยใช้เทคนิค “Support-based Pruning” เพื่อช่วยลดจำนวนการค้นหากฎความสัมพันธ์แทนการแจกแจงทั้งหมด (Agrawal and Srikant, 1994) โดยขั้นตอนวิธีดังกล่าวได้รับการยอมรับว่ามีสมรรถนะเหนือกว่าวิธีการอื่น (ลักษณะโฆโมนอย, 2548) ถึงแม้ว่าจะมีการปรับปรุงประสิทธิภาพการทำงานของขั้นตอนวิธีเอปพรอริเรื่อยๆ เช่น มีการปรับปรุงขั้นตอนวิธีจากเอปพรอริให้เป็น เอปพรอริทิด (AprioriTid) เอปพรอริไฮบริด (Apriori Hybrid) หรือเทอร์ทีอุส (Tertius) (Arora et al., 2013) ทั้งนี้หลักการของขั้นตอนวิธีเอปพรอริที่เป็นมาตรฐานจะเหมือนกัน กล่าวคือ จะค้นหาแบบแนวกว้างจากการสร้างและตรวจสอบไอเท็มเซตที่ปรากฏขึ้นบ่อยทีละขั้น เริ่มจากไอเท็มเซตที่มีจำนวนสมาชิกเท่ากับหนึ่ง ถ้าไอเท็มเซตใดมีค่าสนับสนุนน้อยกว่าที่กำหนดก็จะตัดไอเท็มเซตนั้นออกไปทันที โดยจะทำการวนซ้ำเรื่อยไปจนกว่าจะไม่เหลือเซตของไอเท็มที่จะสร้างในขั้นถัดไป ซึ่งขั้นตอนวิธีเอปพรอริจะอาศัยกระบวนการสำคัญสองกระบวนการ คือ ขั้นตอนการรวมและขั้นตอนการคัดออก (Agrawal and Srikant, 1994) เพื่อให้ได้ค่าความเชื่อมั่น และค่าความสัมพันธ์ (Lift) เป็นผลลัพธ์ของการค้นหากฎความสัมพันธ์ ทั้งนี้การใช้ขั้นตอนเอปพรอริโดยทั่วไปจะเน้นการวิเคราะห์เชิงรายการของการยืมหนังสือในแต่ละครั้งมิใช่การวิเคราะห์รายบุคคล (Tsuji et al., 2012; Mehta et al., 2013) ซึ่งการวิเคราะห์อย่างหลังนั้นสามารถนำไปประยุกต์เพื่อกำหนดพฤติกรรมการยืมหนังสือรายบุคคลที่เชื่อมโยงกับข้อมูลพื้นฐานของนักศึกษาแต่ละราย เช่น คณะที่สังกัด วิชาที่ลงทะเบียนเรียน รวมถึงผลการเรียนของนักศึกษาได้

ขั้นตอนและวิธีการวิเคราะห์

การเก็บข้อมูล

ขั้นตอนการดำเนินการวิจัยประกอบไปด้วยสองส่วนหลักๆ คือ ส่วนที่มีการสัมภาษณ์แบบกึ่งโครงสร้างเพื่อรวบรวมข้อมูลคำถามเกี่ยวกับประเด็นที่เจ้าหน้าที่ที่เกี่ยวข้องกับการยืมหนังสือสนใจจะหาคำตอบ โดยการสัมภาษณ์เชิงลึกนี้ได้ดำเนินการกับบรรณารักษ์ที่รับผิดชอบดูแลการวิเคราะห์เลขหมวด และจัดทำบรรณารายการของหอสมุดกลาง มหาวิทยาลัยราชภัฏราชชนนครินทร์ นอกจากนี้คณะนักวิจัยยังได้นำประวัติรายการยืมหนังสือของนักศึกษาจากทุกคณะจากฐานข้อมูลของหอสมุดกลางในหนึ่งปีการศึกษา (ปีการศึกษา 2558) ซึ่งมหาวิทยาลัยประกอบไปด้วยห้าคณะ ได้แก่ คณะครุศาสตร์ คณะมนุษยศาสตร์และสังคมศาสตร์ คณะเทคโนโลยีอุตสาหกรรม คณะวิทยาการจัดการ คณะวิทยาศาสตร์และเทคโนโลยี รวมจำนวนทั้งสิ้น 12,248 คน ทั้งนี้หมวดหมู่ของหนังสือที่นำมาวิเคราะห์เป็นไปตามรูปแบบของการจัดหมวดหมู่แบบดิวอี้ (Dewey Decimal Classification) แยกเป็น 16 หมวดใหญ่ (ตารางที่ 1) โดยตารางที่ 2 แสดงอัตราการยืมหนังสือแยกตามคณะในปีการศึกษาดังกล่าว

ทั้งนี้สิ่งที่สังเกตได้ว่า จำนวนการยืมหนังสือของนักศึกษายสายสังคมศาสตร์ (ครุศาสตร์ และมนุษยศาสตร์และสังคมศาสตร์) มีมากกว่านักศึกษายสายวิทยาศาสตร์ (เทคโนโลยีอุตสาหกรรม วิทยาการจัดการ และวิทยาศาสตร์และเทคโนโลยี) ซึ่งมาจากเหตุผลที่ว่าจำนวนนักศึกษายสายสังคมศาสตร์ของมหาวิทยาลัยมีจำนวนมากกว่า ดังนั้นเพื่อป้องกันความลำเอียงของข้อมูล คณะนักวิจัยจึงใช้อัตราการยืมของนักศึกษาประกอบการอธิบายเหตุการณ์ที่เกี่ยวข้องกับกฎความสัมพันธ์ที่ค้นพบจากการทำเหมืองข้อมูล ซึ่งเกิดจากการนำจำนวนหนังสือที่ถูกยืมหารด้วยจำนวนนักศึกษาที่ลงทะเบียนเรียนในปีการศึกษานั้นๆ ทั้งนี้ข้อมูลในเบื้องต้นที่สะท้อนให้เห็นถึงอัตราการยืมหนังสือที่ยังอยู่ในเกณฑ์ต่ำ ย่อมจุดประกายความน่าสนใจที่จะค้นหาโอกาสในการเพิ่มอัตราการยืมหนังสือให้มากขึ้นต่อไปในอนาคต

ข้อมูลที่นำมาวิเคราะห์จะเป็นรายการยืมหนังสือที่อยู่ในฐานข้อมูลการให้บริการยืม-คืนหนังสือของหอสมุดกลาง โดยผลลัพธ์ที่ได้จะเป็นความสัมพันธ์ของหมวดหนังสือในแต่ละการยืม ซึ่งจะแสดงความสัมพันธ์ของหมวดในภาพรายบุคคล ซึ่งสามารถอธิบายพฤติกรรมและความสนใจ รวมถึงความเชื่อมโยงกับข้อมูลที่คุณลักษณะและสถานะของนักศึกษาคนดังกล่าว เช่น คณะที่สังกัด วิชาที่เรียนได้อีกด้วย

การวิเคราะห์ข้อมูล

งานวิจัยนี้อาศัยเทคนิคการวิเคราะห์ข้อมูลสามเครื่องมือ กล่าวคือ การวิเคราะห์ข้อมูลโดยอาศัยแบบจำลองทางเวลาด้วยเวฟเลต การวิเคราะห์ข้อมูลด้วยกราฟแจกแจงความถี่ที่แสดงความสัมพันธ์ระหว่างข้อมูลเป็นหมวดหมู่ และการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธีเอโพรออริ

ตารางที่ 1 หมวดหมู่ของหนังสือในหอสมุดกลาง

เลขหมวด	ชื่อหมวด	เลขหมวด	ชื่อหมวด
000	วิทยาการคอมพิวเตอร์ สารสนเทศ และงานทั่วไป	800	วรรณกรรม
100	ปรัชญาและจิตวิทยา	900	ประวัติศาสตร์และภูมิศาสตร์
200	ศาสนา	CD	ซีดี
300	สังคมศาสตร์	Child	เด็กเล็ก
400	ภาษา	Novel	นวนิยาย
500	วิทยาศาสตร์	Pocket Book	หนังสือพอกเก็ต
600	เทคโนโลยี	Young	เด็กโต
700	ศิลปะและนันทนาการ	Short	เรื่องสั้น

ตารางที่ 2 จำนวนและอัตราการยืมหนังสือจากหอสมุดแยกตามรายคณะ

รายชื่อคณะ	จำนวนหนังสือยืม	จำนวนนักศึกษา	อัตราการยืมของนักศึกษา
คณะศึกษาศาสตร์	4,253	2,083	2.04
คณะมนุษยศาสตร์และสังคมศาสตร์	1,005	2,597	0.38
คณะเทคโนโลยีอุตสาหกรรม	170	1,074	0.16
คณะวิทยาการจัดการ	382	4,277	0.08
คณะวิทยาศาสตร์และเทคโนโลยี	681	2,217	0.31
รวม	6,491	12,248	1.89

แบบจำลองทางเวลาด้วยเวฟเลต

เวฟเลต (Wavelet) หรือ คลื่นเล็ก เป็นเทคนิคที่ใช้ในการประมวลสัญญาณเพื่อค้นหา หรือดึงข้อมูลที่ต้องการจากสัญญาณที่มีอยู่ โดยอาศัยฟังก์ชันพื้นฐานที่มีลักษณะเชิงตั้งฉากปกติ (Ortho-normal) เพื่อให้สัญญาณที่กระจัดกระจายอยู่ปรากฏในฐานที่กระชับที่สุด (Chui, 1992) โดยในบทความนี้จะใช้ฟังก์ชันพื้นฐาน daubechies db2 ซึ่งมีลักษณะสัญญาณใกล้เคียงกับรูปแบบการยืมหนังสือ โดยจะใช้ฟังก์ชันดังกล่าวช่วยกระชับรูปแบบการยืมหนังสือให้กระชับที่สุด ทั้งนี้คณะนักวิจัยได้อาศัยแบบจำลองทางเวลาด้วยเวฟเลต 2 ชุด เพื่อแยกสัญญาณออกเป็นความถี่ต่ำและความถี่สูงจากฐานของ DB2 ดังสมการที่ 1 และ 2 ด้านล่างนี้

$$Y_L(n) = \sum_{k=-\infty}^{\infty} x(k)g(n-k) \quad [1]$$

$$Y_H(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) \quad [2]$$

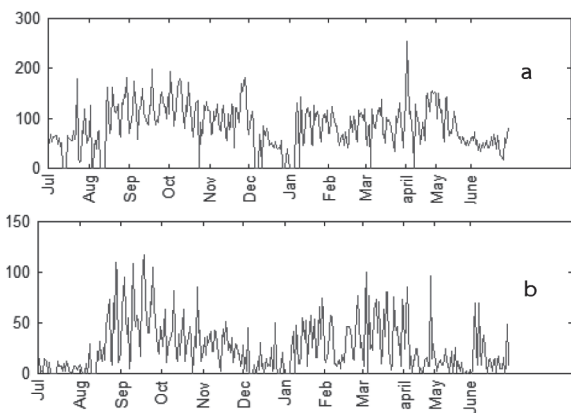
จากสมการข้างต้น $g(n)$ แทนสัญญาณไม่ต่อเนื่อง (Discrete signal) ความถี่ต่ำของ DB2 และ $h(n)$ เป็นสัญญาณไม่ต่อเนื่อง ความถี่สูงของ DB2 และ $x(n)$ แทนสัญญาณข้อมูล โดยมีค่า k แทนการเลื่อนของสัญญาณ

ในบทความนี้ได้นำสัญญาณความถี่ต่ำ หรือที่ได้จากการแยกส่วนที่ไม่ต้องการออก หรือออกไปจากการพิจารณา โดยผลลัพธ์ที่ได้จะเป็นการสร้างแนวโน้มเชิงเส้นที่ขจัดการกวัดแกว่งหรือเบี่ยงเบนอย่างรวดเร็วออกจากความสนใจ ซึ่งข้อมูลที่ถูกนำมาวิเคราะห์ด้วยเวฟเลตสำหรับงานวิจัยนี้ประกอบไปด้วย กราฟแสดงจำนวนนักศึกษาที่เข้าใช้บริการ ณ หอสมุดกลางของมหาวิทยาลัยในช่วงกรกฎาคม 2557 จนถึงปลายเดือนมิถุนายน 2558 เป็นระยะเวลาหนึ่งปีเต็ม (ดังภาพที่ 2 (a)) และกราฟแสดงจำนวนหนังสือที่ถูกยืมในช่วงระยะเวลาเดียวกัน โดยข้อมูลที่ปรากฏในกราฟทั้งสอง

จะถูกนำมาผ่านกระบวนการเวฟเลตที่มีฐานแบบ daubechies dB2 เพื่อสร้างส่วนประมาณค่า (Approximation) และนำไปวิเคราะห์ในสมการที่ [1] เพื่อสร้างแนวโน้มต่อไป

กราฟแจกแจงความถี่

กราฟแจกแจงความถี่เป็นแผนภาพแสดงความสัมพันธ์ระหว่างข้อมูลที่ต้องการที่อาจมีการรวบรวมเป็นหมวดหมู่ ทำให้ผู้วิจัยสามารถเข้าใจการกระจายของกลุ่มข้อมูลที่อาจนำไปสู่กระบวนการประมาณการ หรือการวางแผนสำหรับเหตุการณ์ที่คาดว่าจะเกิดขึ้นในอนาคต ตัวอย่างเช่น การหาค่าความน่าจะเป็นของตัวแปรสุ่มที่เกิดจากค่ากระจายตัวแบบปกติ (Normal distribution) การกระจายตัวแบบพิวซอง (Poisson distribution) การกระจายตัวแบบชี-สแควร์ (Chi-squared) เป็นต้น ทั้งนี้เนื่องจากการค้นหาความสัมพันธ์ตามขั้นตอนเอโพรออริของหมวดหมู่หนังสือในบทความนี้เกิดจากความสัมพันธ์พื้นฐานของความถี่ข้อมูลแยกตามหมวดหมู่ของรายการหนังสือที่ยืม ดังนั้นการนำกราฟแจกแจงความถี่มาวิเคราะห์ร่วมกับขั้นตอนวิธีเอโพรออริจึงเป็นสิ่งจำเป็นอย่างยิ่งต่อการวิเคราะห์ความสัมพันธ์ของการยืมหนังสือในมิติต่างๆ ซึ่งกลับถูกมองข้ามไปในงานวิจัยในปัจจุบัน (Mehta et al., 2013; Tsuji et al., 2012)



ภาพที่ 2 (a) กราฟแสดงจำนวนนักศึกษาที่เข้าใช้บริการหอสมุดกลางในปีการศึกษา 2558 โดยแกนแนวนอนแสดงระยะเวลา ขณะที่แกนแนวตั้งแสดงจำนวนนักศึกษา และ (b) กราฟแสดงจำนวนหนังสือที่ถูกยืมในปีการศึกษา 2558 โดยที่แกนแนวนอนแสดงระยะเวลา ขณะที่แกนแนวตั้งแสดงจำนวนเล่มหนังสือที่ถูกยืม

ขั้นตอนวิธีเอโพรออริ

ในการวิเคราะห์เพื่อค้นหาความสัมพันธ์ของข้อมูลที่เกี่ยวข้องกับการยืมหนังสือในห้องสมุด มักอาศัยการวิเคราะห์ด้วยขั้นตอนวิธีเอโพรออริจากความสัมพันธ์ของ “รายการยืมหนังสือ” ตามหมวดหมู่ของหนังสือ (Ru et al., 2015; Uppal and Chindwani, 2013) บทความวิจัยนี้ต่อยอดแนวคิดดังกล่าวด้วยการวิเคราะห์ตามขั้นตอน

เอโพรออริจากข้อมูลการยืมหนังสือของนักศึกษา “รายคน” ซึ่งแนวคิดดังกล่าวจะช่วยทำให้ทราบลักษณะนิสัย หรือพฤติกรรมของนักศึกษาในการยืมหนังสือจากหอสมุดกลางในปีการศึกษา 2558 โดยกำหนดนิยามและขั้นตอนที่เกี่ยวข้องดังต่อไปนี้

สัญลักษณ์ MinB หมายถึง ค่าเกณฑ์ขั้นต่ำของจำนวนหนังสือที่จะถูกนำมาวิเคราะห์การยืมหนังสือของนักศึกษาตลอดทั้งปีแยกตามหมวด

สัญลักษณ์ S หมายถึง นักศึกษาแต่ละคนในปีการศึกษา โดย S จะเท่ากับจำนวนนักศึกษาที่ยืมหนังสือ ดังนั้น S จึงมีค่าได้ตั้งแต่ 1 2 3 ... N (โดย N คือ จำนวนนักศึกษาทั้งหมดในปีการศึกษา 2558)

สัญลักษณ์ xB หมายถึง ข้อมูลหมวดหนังสือ มีค่าเป็น 001 002 ... จนถึง Short (อ้างอิงจากตารางที่ 1)

สัญลักษณ์ ตาราง T(S, xB) หมายถึง ตารางแสดงความสัมพันธ์ระหว่างการยืมหนังสือตลอดปีการศึกษาของนักศึกษา รายหนึ่งๆ สำหรับหมวดหนังสือที่กำหนดไว้ในหอสมุดกลางของมหาวิทยาลัย ตัวอย่างเช่น $T(100, \text{Short}) = 2$ หมายถึง นักศึกษา รายที่ 100 มีการยืมหนังสือหมวด Short ทั้งหมด 2 เล่มตลอดปีการศึกษา 2558

สัญลักษณ์ Pr(S, xB) หมายถึง ตาราง T(S, xB) ที่มีค่าในตารางเพียง 0 หรือ 1 ตามเงื่อนไขต่อไปนี้

ขั้นตอนที่ 1 : ตรวจสอบรายการข้อมูลยืมหนังสือจากการยืมแต่ละครั้ง รวบรวมและเรียบเรียงข้อมูลใหม่ให้เป็นรายการยืมของนักศึกษาตลอดทั้งปีในตาราง T(S, xB)

ขั้นตอนที่ 2 : ตรวจสอบรายการในตาราง T(S, xB) สำหรับค่าขั้นต่ำของการยืม MinB กำหนดเป็น 1 ซึ่งสามารถเปลี่ยนแปลงได้ เมื่อ $T(S, xB) \geq \text{MinB}$ กำหนดค่าในตาราง $\text{Pr}(S, xB) = 1$ มิเช่นนั้น $\text{Pr}(S, xB) = 0$

ขั้นตอนที่ 3 : ตรวจสอบจำนวนรายการทั้งหมดในตาราง Pr จากจำนวน N รายการ เพื่อค้นหาความถี่ของการยืมหนังสือในแต่ละหมวด A, B, C, ... เป็น $f(A), f(B), f(C)$

ขั้นตอนที่ 4 : พิจารณาการปรากฏของข้อมูลร่วมกันเช่น {A, B}, {A, C}, {B, C} ... และพิจารณาความถี่ของเหตุการณ์ของกลุ่มข้อมูลรวมนั้นปรากฏ $f\{A, B\}, f\{A, C\}, f\{B, C\}$... ขั้นตอนนี้จะคำนวณถี่ของข้อมูลรวมจากความถี่ของข้อมูลพื้นฐานตามกฎความน่าจะเป็นโดย (1) เพิ่มขนาดของสมาชิกกลุ่มข้อมูลขึ้นเรื่อยๆ เช่น {A, B, C} และหา $f\{A, B, C\}$ ค่าตัวแปรจะถูกเก็บไว้ในลักษณะต้นไม้ และ (2) นำความถี่ที่ได้มาหาความน่าจะเป็นของเหตุการณ์ที่มีการอ้างอิง เช่น $P(\{A, B\} | A)$ เป็นการหาความน่าจะเป็นที่เหตุการณ์ {A, B} จะเกิดขึ้นเมื่อเทียบกับเหตุการณ์ A ซึ่งสามารถนำไปตีความ

ได้ว่า เมื่อเกิดเหตุการณ์ A จะมีโอกาสที่จะเกิดเหตุการณ์ {A, B} พร้อมกันเท่าไร ซึ่งนำไปสร้างเป็นกฎความสัมพันธ์ต่อไป

ขั้นตอนวิธีข้างต้นจะทำให้คณะนักวิจัยสามารถค้นหาและวิเคราะห์พฤติกรรมการยืมหนังสือของนักศึกษารายบุคคลได้ซึ่งจะสามารถนำไปขยายผลต่อไปได้ เช่น วิชาที่นักศึกษาท่านดังกล่าวกำลังศึกษาอยู่ในช่วงเวลายืม หรือแม้แต่คณะที่นักศึกษาดังกล่าวสังกัดอยู่ ซึ่งผลลัพธ์ที่ได้จะนำไปสู่การยกระดับคุณภาพของการเรียนการสอนของนักศึกษาผ่านการสนับสนุนของหอสมุดกลางของมหาวิทยาลัย

นอกจากนี้ขั้นตอนวิธีเอโพรออริจะช่วยให้เกิดการวิเคราะห์และสร้างมุมมองเชิงลึกที่มากกว่าการพิจารณาเพียงความถี่ของการยืมหนังสือในหมวดเดียวกันอย่างกรณีที่ดำเนินการด้วยการวิเคราะห์ความสัมพันธ์ของข้อมูลผ่านกราฟแจกแจงความถี่ บทความนี้ได้นำ

รูปแบบการกระจายของข้อมูลจากกราฟแจกแจงความถี่มาสนับสนุนกฎความสัมพันธ์ตามขั้นตอนวิธีเอโพรออริ ส่งผลให้ผู้บริหารหอสมุดสามารถบริหารจัดการทรัพยากรหนังสือได้อย่างมีประสิทธิภาพยิ่งขึ้น ทั้งนี้สำหรับการทำเหมืองข้อมูลในงานวิจัยนี้จะอาศัยตัวแปรสำคัญสองค่า คือ ค่าสนับสนุน และค่าความเชื่อมั่น โดยค่าสนับสนุนในงานวิจัยนี้ หมายถึง จำนวนครั้งของการยืมหนังสือในหมวดที่สนใจ เทียบกับจำนวนรายการทั้งหมด เพื่อใช้เป็นเกณฑ์ขั้นต่ำของการปรากฏหมวดหมู่หนังสือที่ผู้บริหารสนใจ ในขณะที่ค่าความเชื่อมั่น หมายถึง ร้อยละของความน่าจะเป็นของเหตุการณ์หนึ่งที่สามารถเกิดขึ้น ตัวอย่างเช่น ร้อยละของนักศึกษาที่ยืมหนังสือหมวด 001 และ 002 แล้วจะมีการยืมหมวด 003 = จำนวนครั้งของการปรากฏของเหตุการณ์ {001, 002, 003} เทียบกับเหตุการณ์ {001, 002} ตารางที่ 3 สรุปเครื่องมือที่ใช้ในการวิเคราะห์เหมืองข้อมูล

ตารางที่ 3 เครื่องมือที่ใช้สำหรับการวิเคราะห์ข้อมูลสำหรับการทำเหมืองข้อมูล

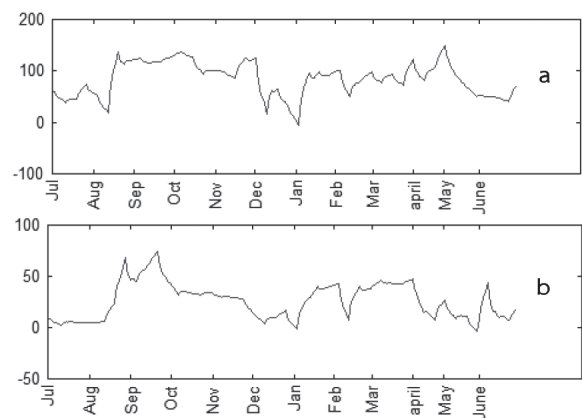
เครื่องมือ	ข้อมูลป้อนเข้า	ผลลัพธ์ที่คาดหวัง
เวฟเลต	ข้อมูลการเข้าห้องสมุด	แนวโน้มการเข้าห้องสมุดและการยืมหนังสือของนักศึกษา
กราฟแจกแจงความถี่	ข้อมูลการยืมหนังสือทั้งหมด	การกระจายของหมวดหนังสือที่ถูกยืม ค่าสูงสุด ค่าเบี่ยงเบนเพื่อวิเคราะห์ลักษณะข้อมูลป้อนเข้าร่วมกับกราฟจากเอโพรออริ
เอโพรออริ	รายการยืมหนังสือจำแนกเป็นของนักศึกษาแต่ละคน	ความสัมพันธ์จากการยืมหนังสือของนักศึกษาแต่ละราย

ผลการวิจัย

คณะนักวิจัยได้อาศัยเครื่องมือและวิธีการที่ได้อภิปรายในหัวข้อก่อนหน้า ในการวิเคราะห์ข้อมูลการยืมหนังสือที่อยู่ในฐานข้อมูลการให้บริการทรัพยากรสารสนเทศของหอสมุดกลางมหาวิทยาลัยราชภัฏราชนครินทร์ โดยสามารถสรุปผลการวิเคราะห์ข้อมูลได้ดังนี้

ผลการวิเคราะห์ด้วยแบบจำลองทางเวลาเวฟเลต

ภายหลังจากการนำกราฟแสดงจำนวนนักศึกษาที่เข้าใช้บริการในหอสมุดกลาง และกราฟแสดงจำนวนหนังสือที่ถูกยืมในปีการศึกษา 2558 ดังที่ปรากฏในรูป 2(a) และ 2(b) ตามลำดับมาวิเคราะห์ด้วยเวฟเลตที่มีฐานแบบ daubechies dB2 เพื่อใช้ค้นหาส่วนประมาณค่า คณะนักวิจัยค้นพบกราฟแสดงแนวโน้มที่สะดวกต่อการพิจารณาทิศทางและประเมินสถานการณ์การยืมคืนในภาพรวม ดังแสดงในภาพที่ 3(a) และ 3(b)



ภาพที่ 3 กราฟผลลัพธ์ที่ผ่านกระบวนการเวฟเลตที่มีฐานแบบ daubechies dB2 เพื่อค้นหาแนวโน้มของข้อมูล โดยกราฟ (a) แสดงส่วนประมาณค่าจำนวนผู้เข้าใช้บริการหอสมุดกลาง และ (b) แสดงส่วนประมาณค่าจำนวนหนังสือที่ถูกยืม ตามลำดับ

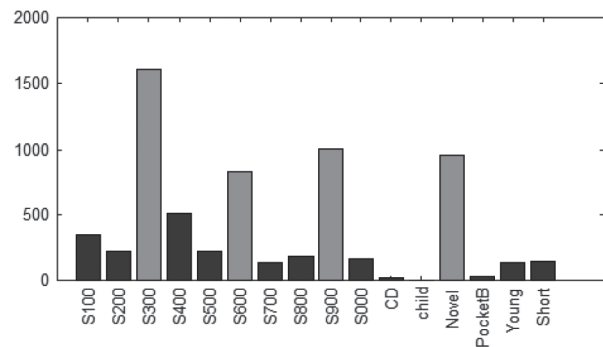
จากภาพที่ 3(a) จะเห็นได้ว่าจำนวนนักศึกษาเข้าใช้บริการในหอสมุดกลางแบ่งออกได้เป็นสองช่วง คือ ช่วงสูงและช่วงต่ำ ซึ่งเมื่อพิจารณาจากแกนเวลาจะสอดคล้องกับช่วงระยะเวลาของการเปิดและปิดภาคการศึกษา โดยในช่วงเปิดภาคการศึกษานักศึกษาเข้าใช้บริการมากกว่าปิดเทอมเพียงไม่ถึงสองเท่า แสดงให้เห็นว่า การใช้บริการในหอสมุดกลางช่วงเปิดภาคการศึกษายังมีปริมาณน้อยเมื่อเทียบกับจำนวนนักศึกษาลงทะเบียนเรียนในช่วงเวลาดังกล่าว และหากจะย้อนกลับไปดูกราฟ 2(a) เพื่อศึกษาจำนวนผู้เข้าใช้บริการในระดับที่ลึกลงไป จะพบการแกว่งตัวขึ้น-ลงในรอบหนึ่งสัปดาห์ ซึ่งก็น่าจะใช่เรื่องแปลกอะไร เนื่องจากในช่วงวันเสาร์-อาทิตย์มหาวิทยาลัยไม่มีการเรียนการสอนในภาคปกติ ข้อเท็จจริงที่เป็นรูปธรรมดังกล่าวสามารถนำไปใช้ประกอบการกำหนดกลยุทธ์ในการส่งเสริมให้เกิดการเข้าใช้บริการในหอสมุดกลางต่อไปได้

ทำนองเดียวกับกราฟ 2(a) และ 3(a) งานวิจัยนี้พบว่าจำนวนการยืมหนังสือจากกราฟ 2(b) และ 3(b) มีความสอดคล้องกับระยะเวลาในการปิดภาคการศึกษาเช่นกัน โดยมีข้อสังเกตว่าในช่วงปิดภาคการศึกษามีนักศึกษาจำนวนหนึ่งเข้าใช้บริการหอสมุดกลางแต่แทบไม่มีการยืมหนังสือออกจากหอสมุด เช่น ช่วงปิดภาคการศึกษาฤดูร้อน (ปิดเทอมใหญ่) ซึ่งอาจเป็นการเข้าห้องสมุดเพื่อคืนหนังสือที่ยืมไปเพื่อใช้สอบปลายภาคในภาคการศึกษาก่อนหน้า และจากการใช้ Pearson Correlation เพื่อดูความสัมพันธ์ของกราฟ 3(a) และ 3(b) จะได้ค่าสหสัมพันธ์ (Correlation Coefficient) เท่ากับ 0.34 ซึ่งแสดงถึงความสัมพันธ์ไปในทิศทางเดียวกันในบางช่วงของปีการศึกษา โดยเฉพาะอย่างยิ่ง ช่วงปิดภาคการศึกษาฤดูร้อน และช่วงเปิดภาคการศึกษาแรกของปีการศึกษา นอกจากนี้จากการวิเคราะห์กราฟในรูปที่ 2 และ 3 จะพบข้อสังเกตดังต่อไปนี้

1. ช่วงกลางเดือนสิงหาคมถึงกลางเดือนกันยายน ซึ่งเป็นช่วงเปิดภาคการศึกษาแรก มีจำนวนหนังสือที่ถูกยืมจากหอสมุดมากที่สุด ซึ่งสอดคล้องกับแนวโน้มของจำนวนนักศึกษาที่เข้าใช้บริการห้องสมุด
2. ช่วงต้นเดือนธันวาคม ซึ่งเป็นช่วงก่อนการสอบปลายภาค (ภาคการศึกษาแรก) ปรากฏว่ามีจำนวนผู้เข้าใช้บริการเป็นจำนวนมาก แต่กลับมีจำนวนหนังสือที่ถูกยืมออกจากหอสมุดน้อย ซึ่งจากการตรวจสอบพบว่าจำนวนผู้เข้าใช้บริการที่เพิ่มขึ้นสืบเนื่องมาจากการเข้าใช้บริการเครื่องคอมพิวเตอร์ที่จัดให้โดยหอสมุดกลางที่มีจำนวนเพิ่มมากขึ้น ในช่วงระยะเวลาดังกล่าว
3. ช่วงเดือนมกราคมซึ่งเป็นช่วงเปิดภาคการศึกษาที่สอง นักศึกษาเริ่มกลับเข้าใช้บริการหอสมุดกลาง และมีจำนวนการยืมหนังสือที่เพิ่มมากขึ้นจากเดือนก่อน อย่างไรก็ตามตามจำนวนหนังสือที่ถูกยืมยังอยู่ในระดับที่ต่ำ เมื่อเทียบกับช่วงเปิดภาคการศึกษาแรก

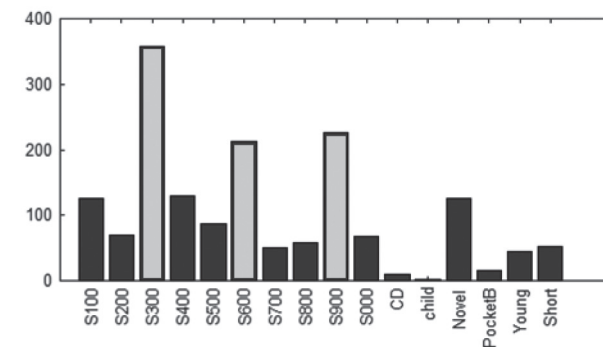
ผลการวิเคราะห์ด้วยกราฟแจกแจงความถี่

จากข้อมูลการยืมหนังสือ จำนวนรวมทั้งสิ้น 6,491 รายการ นำไปสู่การสร้างกราฟแจกแจงความถี่ (Histogram) ดังภาพที่ 4 ซึ่งเมื่อพิจารณาจากจำนวนของรายการยืมหนังสือแยกตามหมวดของหนังสือ จะสามารถแบ่งกลุ่มคร่าวๆ ออกได้เป็นสองกลุ่ม กลุ่มแรก คือ กลุ่มที่มีปริมาณการยืมสูง ประกอบไปด้วย หนังสือจากหมวด 300 600 900 และนวนิยาย และกลุ่มที่มีปริมาณการยืมต่ำ ประกอบไปด้วย หนังสือจากหมวดอื่นๆ นอกเหนือจากที่ปรากฏในกลุ่มที่มีปริมาณการยืมสูง โดยกลุ่มแรกจะมีค่าเฉลี่ยของจำนวนการยืมอยู่ที่ค่าเฉลี่ย 1,100 รายการ ขณะที่อีกกลุ่มมีค่าเฉลี่ยอยู่ที่ 174 รายการ ซึ่งความแตกต่างของจำนวนรายการยืมหนังสือของทั้งสองกลุ่มที่ห่างกันมากจะส่งผลต่อความลำเอียง (Bias) ในการวิเคราะห์ความสัมพันธ์เอโทรออริและกำหนดกลยุทธ์ในการจัดการ



ภาพที่ 4 ความถี่ของจำนวนหนังสือที่ถูกยืมในแต่ละหมวดหมู่

และเมื่อพิจารณาค่าสหสัมพันธ์ของกราฟในภาพที่ 4 เทียบกับกราฟแสดงจำนวนนักศึกษาที่ยืมหนังสือในแต่ละหมวด ดังภาพที่ 5 จะพบว่า ค่าสหสัมพันธ์ของข้อมูลในกราฟทั้งสองอยู่ที่ 0.957 ซึ่งมีความหมายว่า จำนวนนักศึกษาที่ยืมหนังสือสอดคล้องกับจำนวนหนังสือที่ถูกยืมกระจายไปตามแต่ละหมวด โดยหมวดของหนังสือที่มีจำนวนนักศึกษายืมอยู่ในเกณฑ์มาก 3 อันดับแรกได้แก่ หมวด 300 900 และ 600 ตามลำดับ



ภาพที่ 5 ความถี่ของจำนวนนักศึกษาที่ยืมหนังสือต่อหมวด

จากกราฟแจกแจงความถี่ข้างต้นทำให้ผู้วิจัยกำหนดข้อควรระวังในการวิเคราะห์ด้วยวิธีเอโพรอร์รีได้ดังต่อไปนี้

1. ความถี่ของการยืมหนังสือในหมวด 300 600 และ 900 มีค่าสูง ส่งผลให้ค่าสนับสนุนในดังกล่าวมีค่าสูง และมีความแตกต่างกับกลุ่มที่เหลือค่อนข้างมาก สำหรับหมวดนวนิยาย พบว่ามีจำนวนผู้ยืมที่ต่ำแต่จำนวนหนังสือที่ถูกยืมสูง ซึ่งเป็นไปได้ว่ามีกลุ่มผู้สนใจยืมหนังสือจากหมวดดังกล่าวมีจำนวนไม่มาก แต่มักจะมีการยืมอย่างต่อเนื่อง

2. เมื่อคำนวณค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานจะพบว่าหนังสือในหมวด 300 มีค่าเฉลี่ยของจำนวนหนังสือที่ถูกยืมต่อคนเท่ากับ 2.2 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 4.3 ขณะที่หนังสือในหมวด 600 มีค่าเฉลี่ยของจำนวนหนังสือที่ถูกยืมต่อคนเท่ากับ 1.1 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 2.6 และหนังสือในหมวด 900 มีค่าเฉลี่ยของจำนวนหนังสือที่ถูกยืมต่อคนเท่ากับ 1.3 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 4.3 ซึ่งค่าเฉลี่ยดังกล่าวจัดได้ว่าอยู่ในเกณฑ์ที่ค่อนข้างต่ำในขณะที่ค่าความเบี่ยงเบนค่อนข้างสูง ตีความได้ว่าข้อมูลมีการกระจายตัวมาก ซึ่งอาจแบ่งกลุ่มของนักศึกษาได้เป็นสองกลุ่ม คือ กลุ่มที่มีการยืมหนังสือจำนวนมากกว่าค่าเฉลี่ย และที่น้อยกว่าค่าเฉลี่ย

ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธีเอโพรอร์รี

จากการวิเคราะห์ความสัมพันธ์ของรายการยืมหนังสือข้ามหมวดเพื่อศึกษาพฤติกรรมการยืม ตลอดปีการศึกษาพบว่า มีนักศึกษากว่า 731 คนที่มีการยืมหนังสือข้ามหมวด โดยตารางที่ 4 แสดงรายการความสัมพันธ์ของการยืมหนังสือข้ามหมวดในทุกรอบเดือนที่มีค่าสนับสนุนสูงสุด 10 อันดับแรก (MinB = 1) และรายการความสัมพันธ์ข้ามหมวดในในรอบเดือนที่มีรายการยืมหนังสือข้ามหมวดอย่างน้อย 2 ครั้ง ตลอดปีการศึกษาที่มีค่าสนับสนุนสูงสุด 10 อันดับแรกของข้อมูลทั้งหมด (MinB = 2) จากตารางที่ 4 คณะนักวิจัยสามารถสรุปสาระสำคัญที่ได้ ดังนี้

1. เมื่อพิจารณาจากความสัมพันธ์จากขั้นตอนวิธีที่พัฒนาขึ้นจะพบว่า ความสัมพันธ์กับการยืมหนังสือจากหมวด 300 600 และ 900 เกิดขึ้นบ่อยครั้ง ทั้งนี้ผลลัพธ์ดังกล่าวไม่น่าประหลาดใจเนื่องจากหากพิจารณาจากกราฟแจกแจงความถี่จะพบว่าหมวดดังกล่าวที่มีค่าสนับสนุนมาก

2. ค่าเบี่ยงเบนมาตรฐานที่มีการกระจายตัวมากแสดงให้เห็นว่ากลุ่มข้อมูลที่นำมาวิเคราะห์มีลักษณะการยืมที่แตกต่างกันมาก ดังนั้นการปรับค่า MinB จะช่วยคัดกรองกลุ่มนักศึกษาที่คณะนักวิจัยสนใจได้ดีขึ้น โดยเฉพาะอย่างยิ่ง กรณีกลุ่มนักศึกษาที่มีรูปแบบของยืมหนังสือ (หมวดหนังสือ) เหมือนกันมากกว่าหนึ่งครั้งของรายการยืม

ตารางที่ 4 สรุป 10 อันดับความสัมพันธ์ของหมวดการยืมหนังสือที่มี MinB = 1 และ MinB = 2 ตามลำดับ

ลำดับที่	MinB=1				MinB=2			
	หมวดเริ่มต้น	หมวดที่ตามมา	ค่าความเชื่อมั่น	ค่าสนับสนุน	หมวดเริ่มต้น	หมวดที่ตามมา	ค่าความเชื่อมั่น	ค่าสนับสนุน
1	900	300	0.729	0.329	900	300	0.727	0.338
2	300	900	0.507	0.329	300	900	0.578	0.338
3	600	300	0.625	0.252	100	300	0.567	0.124
4	100	300	0.673	0.168	400	300	0.538	0.102
5	400	300	0.651	0.165	200	600	0.5	0.073
6	Novel	300	0.51	0.118	200	900	0.5	0.073
7	200	300	0.677	0.106	100 และ 900	300	0.731	0.069
8	600 และ 900	300	0.688	0.106	100 และ 300	900	0.559	0.069
9	500	300	0.557	0.094	600 และ 900	300	0.576	0.069
10	200	900	0.585	0.091	000	300	0.577	0.055

3. เมื่อพิจารณาจากความสัมพันธ์จากขั้นตอนวิธีที่พัฒนาขึ้น จะพบว่า ความสัมพันธ์กับการยืมหนังสือจากหมวด 300 600 และ 900 เกิดขึ้นบ่อยครั้ง ทั้งนี้ผลลัพธ์ดังกล่าวไม่น่าประหลาดใจ เนื่องจากหากพิจารณาจากกราฟแจกแจงความถี่จะพบว่าหมวดดังกล่าวที่มีค่านับสนุนมาก

4. MinB เปลี่ยนจาก 1 เป็น 2 จะทำให้กฎความสัมพันธ์ข้อที่ 3 และ 6 ในตาราง MinB = 1 หายไปจาก 10 อันดับความสัมพันธ์ที่มีค่านับสนุนสูงที่สุด ย่อมแสดงว่าโอกาสในการกลับมายืมหนังสือในหมวด 600 (เทคโนโลยี) และ Novel (นวนิยาย) นั้นมีน้อย

5. (MinB = 1) กฎความสัมพันธ์เกือบทุกข้อจะแสดงโอกาสที่นักศึกษาจะยืมหนังสือที่ตามมาในหมวด 300 แปลความได้ว่า นักศึกษาสนใจหนังสือในหมวดนี้มากเป็นพิเศษ อาจด้วยเพราะนักศึกษาทุกคนชอบหนังสือในหมวดนี้ หรือต้องค้นคว้าโดยอ้างอิงกับหนังสือในหมวดดังกล่าว

6. ทั้งสองกรณี (MinB = 1 และ 2) โอกาสที่นักศึกษาคนหนึ่งที่ยืมหนังสือในหมวด 900 (ประวัติศาสตร์และภูมิศาสตร์) แล้วจะยืมหนังสือในหมวด 300 (สังคมศาสตร์) มีมากถึงกว่าร้อยละ 72 แต่ในทางกลับกัน นักศึกษาที่ยืมหนังสือในหมวด 300 กลับมีความเป็นไปได้น้อยกว่า ซึ่งหากเชื่อมโยงกับกราฟแจกแจงความถี่ของการยืมหนังสือในรูปที่ 4 ก็คงพบว่า การที่หนังสือในหมวด 300 มีจำนวนมากอาจเพิ่มโอกาสในการเป็นหมวดที่ตามมาจากการยืมหนังสือ

7. กรณีพิจารณาการยืมข้ามหมวดทุกเงื่อนไข (MinB = 1) นักศึกษาที่ยืมหนังสือในหมวด 600 แล้วจะยืมหนังสือในหมวด 300 มีโอกาสเกิดขึ้นกว่าร้อยละ 63 แต่เมื่อนับจำนวนครั้งของการยืมข้ามหมวดที่มากกว่า 1 ครั้งตลอดปีการศึกษา (MinB = 2) จะพบว่าโอกาสที่เกิดเหตุการณ์ดังกล่าวซาลดลงมาก ตรงข้ามกับกรณีที่นักศึกษายืมหนังสือในหมวด 100 (ปรัชญาและจิตวิทยา) หรือ 400 (ภาษา) แล้วจะยืมหนังสือในหมวด 300 ซ้ำ ซึ่งอาจแปลความได้ว่า หนังสือในหมวด 100 300 และ 400 เป็นหมวดของหนังสือที่เกี่ยวข้องกับมนุษยศาสตร์และสังคมศาสตร์ ขณะที่หนังสือในหมวด 600 เป็นหมวดเทคโนโลยีซึ่งเกี่ยวข้องกับมนุษยศาสตร์และสังคมศาสตร์น้อยกว่า นอกจากนี้เมื่อตรวจสอบต้นสังกัดของนักศึกษาที่ยืมหนังสือในหมวดต่างๆ จะพบว่า นักศึกษาที่สังกัดคณะทางสายวิทยาศาสตร์ (คณะวิทยาศาสตร์และเทคโนโลยี และคณะเทคโนโลยีอุตสาหกรรม) จะยืมหนังสือจำนวนน้อยกว่าสายมนุษยศาสตร์และสังคมศาสตร์ ซึ่งนี่เองอาจเป็นเหตุผลสำคัญที่ทำให้หนังสือที่เกี่ยวข้องกับหมวดทางด้านวิทยาศาสตร์และเทคโนโลยีมีจำนวนการยืมน้อย

8. ด้วยจำนวนการยืมหนังสือในหมวด 300 ที่มีปริมาณมาก ทำให้การสรุปผลการวิเคราะห์ข้อมูลเป็นไปได้อย่างกว้างๆ เท่านั้น

อีกทั้งความถี่ของการยืมหนังสือในหมวดนี้มีสูง ทำให้บังบังความสัมพันธ์ระหว่างหมวดหนังสืออื่น ดังนั้น หากต้องการวิเคราะห์ข้อมูลการยืมหนังสือที่ละเอียดหรือเจาะจงลงไป จำเป็นที่หอสมุดจะกำหนดหมวดหมู่ย่อยให้กับหนังสือในหมวดดังกล่าว

อภิปรายผลการวิจัยและสรุปผลวิจัย

งานวิจัยที่เกี่ยวข้องกับการนำเทคนิคการทำเหมืองข้อมูลมาวิเคราะห์รูปแบบของการยืมหนังสือจากห้องสมุด โดยมากมักให้ความสำคัญกับการนำข้อมูลรายการยืมหนังสือทั้งหมดมาวิเคราะห์ โดยไม่คำนึงถึงคุณสมบัติของบุคคลที่ยืมหนังสือแต่ละราย และช่วงเวลาที่เกี่ยวข้องกับการยืมหนังสือ (Nicholson, 2006; Ru et al., 2015; Tsuji et al., 2012; Uppal and Chindwani, 2013) อย่างไรก็ตาม วิธีการดังกล่าวกลับไม่สามารถอธิบายพฤติกรรมการยืมหนังสือจากห้องสมุดในรูปแบบ “รายบุคคล” ที่เชื่อมโยงกับมิติเวลาและคุณสมบัติของบุคคลได้ งานวิจัยชิ้นนี้จึงได้ผนวกการวิเคราะห์ข้อมูลโดยอาศัยแบบจำลองทางเวลาด้วยเวฟเลต และกราฟกระจายความถี่เข้ากับเทคนิคการทำเหมืองข้อมูลตามขั้นตอนวิธีเอไพโรอริ โดยแนวคิดดังกล่าวทำให้คณะนักวิจัยสามารถอธิบายพฤติกรรมการยืมหนังสือรายบุคคลที่เจาะลึกลงไปเทียบกับมิติคุณสมบัติของบุคคลดังกล่าว และมิติของช่วงเวลาที่เกี่ยวข้อง โดยกฎความสัมพันธ์ที่ถูกค้นพบและอธิบายใหม่นี้จะช่วยทำให้ผู้บริหารของหอสมุดกลาง มหาวิทยาลัยราชภัฏราชชนครินทร์ สามารถออกแบบกลยุทธ์ที่มุ่งเน้นการตอบสนองความต้องการเฉพาะราย ซึ่งเหมาะสมกับลักษณะนิสัย คณะที่สังกัด และวิชาที่นักศึกษาลงทะเบียนต่อไปในอนาคต

ผลของการวิจัยทำให้ทราบว่า แบบจำลองทางเวลาด้วยเวฟเลตสามารถช่วยทำให้ผู้บริหารกำหนดกลยุทธ์การวางแผนส่งเสริมการอ่านโดยอาศัย “แนวโน้มน” ของช่วงเวลาที่ปริมาณการยืมหนังสือน้อย และวางแผนในการให้บริการยืมหนังสือให้มีประสิทธิภาพยิ่งขึ้นในช่วงเวลาที่มีแนวโน้มว่าจะมีปริมาณการยืมหนังสือมาก นอกจากนี้ งานวิจัยชิ้นนี้ยังแสดงให้เห็นความสำคัญของกราฟแจกแจงความถี่ที่มักจะถูกละเลยในการทำเหมืองข้อมูลทั่วไปว่าสามารถนำมาประยุกต์เข้ากับเทคนิคการทำเหมืองข้อมูลตามขั้นตอนวิธีเอไพโรอริ เพื่อสังเกตลักษณะของความเบี่ยงเบนจากระดับกัน (Threshold) ของประเด็นที่สนใจได้ ยกตัวอย่างเช่น จำนวนหนังสือในแต่ละหมวดที่มีการยืมตลอดปีการศึกษา และจำนวนนักศึกษาที่ยืมหนังสือในแต่ละหมวด ซึ่งจะนำไปสู่การแบ่งกลุ่มของข้อมูล เช่น กลุ่มหมวดที่มีจำนวนหนังสือที่ถูกยืมมากกับกลุ่มหมวดที่มีจำนวนหนังสือที่ถูกยืมน้อย เป็นต้น การแบ่งกลุ่มดังกล่าว นอกจากจะทำให้คณะนักวิจัยสามารถวิเคราะห์ที่มาของ

กฎความสัมพันธ์ที่ได้จากขั้นตอนวิธีเอโพรอรี้ได้แล้ว ในอนาคตผู้บริหารของหอสมุดกลางอาจประยุกต์ใช้การกำหนดค่าน้ำหนัก (Weighting) ให้กับประเด็นที่สนใจเพื่อรักษาคุณภาพของการให้บริการทรัพยากรสารสนเทศในมิติต่างๆ ดังแสดงในตารางที่ 5 ซึ่งเปรียบเทียบขั้นตอนวิธีเอโพรอรี้ทั่วไปกับขั้นตอนวิธีเอโพรอรี้ประยุกต์ที่ใช้ในงานวิจัยนี้

ในอนาคตนักวิจัยที่ประสงค์จะต่อยอดงานวิจัยนี้อาจพิจารณาใช้ข้อมูลคุณสมบัติของ “นักศึกษา” แต่ละรายเพื่อประกอบการอธิบายพฤติกรรมเชิงลึกที่ตอบโจทย์เฉพาะกรณี ซึ่งผลลัพธ์ดังกล่าวสามารถนำไปสู่การปรับแต่ง (Customised) กลยุทธ์ให้มี

ความเหมาะสมและประสิทธิภาพยิ่งขึ้น ตัวอย่างเช่น ผลลัพธ์จากการเชื่อมโยงผลการเรียนของนักศึกษาแต่ละรายกับรูปแบบของการยืมหนังสือจากหอสมุด อาจทำให้ผู้บริหารค้นพบแนวทางหรือกลยุทธ์ในการส่งเสริมการอ่าน หรือผลลัพธ์จากการศึกษาว่าเหตุใดนักศึกษาที่มาจากคณะทางสายวิทยาศาสตร์และเทคโนโลยีจึงมีความถี่ของการยืมหนังสือในระดับที่น้อยกว่านักศึกษาที่มาจากสายสังคม ซึ่งต้นเหตุของปัญหาดังกล่าวอาจมีได้เกิดจากตัวนักศึกษาโดยตรง หากแต่อาจเกิดจากองค์ประกอบร่วม เช่น หอสมุดอาจมิได้จัดเตรียมหนังสือที่นักศึกษาต้องการในปริมาณที่เหมาะสม และ/หรือ หนังสือที่จัดหาให้ขาดซึ่งคุณภาพและความทันสมัยก็เป็นได้

ตารางที่ 5 ผลการเปรียบเทียบระหว่างขั้นตอนวิธีเอโพรอรี้ทั่วไปและแบบประยุกต์

ประเด็นการพิจารณา	ขั้นตอนวิธีเอโพรอรี้ทั่วไป	ขั้นตอนวิธีเอโพรอรี้ประยุกต์
กระบวนการ	ค้นหาความสัมพันธ์ระหว่างหมวดต่างๆ ในรายการยืมทั้งหมด	ค้นหาความสัมพันธ์ระหว่างหมวดต่าง ๆ ภายใต้เงื่อนไขที่กำหนด เช่น พารามิเตอร์ MinB กรอบเวลาที่สนใจ คุณสมบัติของผู้ยืม การกระจายน้ำหนักของการยืมตามสังกัดของผู้ยืม เป็นต้น
มุมมอง	รายการยืมหนังสือแต่ละครั้ง	รายการยืมหนังสือรายบุคคล
ประโยชน์	ค้นพบรูปแบบของกฎความสัมพันธ์เพื่อนำไปประยุกต์ใช้ในการจัดเรียงและกำหนดตำแหน่งของหนังสือในหอสมุด รวมถึงการให้คำแนะนำ ในการยืมหนังสือ	<ul style="list-style-type: none"> - สามารถนำไปออกแบบกลยุทธ์ที่สอดคล้องกับพฤติกรรมของนักศึกษาตามช่วงเวลาและข้อมูลความถี่ต่างๆ - ค้นพบความสัมพันธ์ที่สามารถนำไปออกแบบกลยุทธ์ที่ตอบสนองความต้องการรายบุคคล - ในอนาคตสามารถเชื่อมโยงผลลัพธ์จากกฎความสัมพันธ์กับคุณสมบัติของบุคคลในเชิงลึก เช่น ผลการเรียน คณะที่สังกัด วิชาที่ลงทะเบียนเรียน เป็นต้น ซึ่งจะนำไปสู่การจัดการเรียนการสอนที่มีประสิทธิภาพยิ่งขึ้น

เอกสารอ้างอิง

วิภาวรรณ บัวทอง. (2014). **Association Rule**. <https://wipawanblog.files.wordpress.com/2014/06/chapter-4-association-rule.pdf>. ลงวันที่ 19 พฤษภาคม 2014

ลักขมิ โขมโนทัย. (2548). **การค้นหาความสัมพันธ์ที่แข็งแกร่ง**. วิทยานิพนธ์ มหาวิทยาลัยสุรนารี.

สุภาพร ชัยธัมมะปกรณ์. (2014). **Bibliomining เพื่อประโยชน์ในการบริหารงานห้องสมุด**. <https://supapornhuang.wordpress.com/2014/01/14/bibliomining/> ลงวันที่ 14 มกราคม 2014.

เอกสิทธิ์ พัทธวงศ์ศักดิ์. (2015). **กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM และตัวอย่างการประยุกต์ใช้ทางการศึกษา**. <http://dataminingtrend.com/2014/data-mining-techniques/crisp-dm-example/> ลงวันที่ 3 ตุลาคม 2015

Agrawal, R. and Srikant, R. (1994). **Fast algorithms for mining association rules**. Proceedings of the International Conference on Very Large Databases. pp. 487-499.

- Ahirwal, Ramratan, Kumar Kori, Neelesh and Jain, Y.K. (2012). Improved Data mining approach to Find Frequent Itemset Using Support count table. **IJETTCS**. 1(2), July – August 2012.
- Arora, Jyoti, Bhalla, Nidhi, and Rao, Sanjeev. (2013). A Review on Association Rule Mining Algorithms. **International Journal of Innovative Research in Computer and Communication Engineering**. 1(5). pp. 1246-1251.
- Azevedo, Ana and Zantos, Manuel Filipe. (2008). **KDD, SEMMA and CRISP-DM: a parallel overview**.
- Castro- Gessner, Gabriela, Wilcox, Wendy and Chandler, Adam. (2013). **Hidden Patterns of LibGuides Usage: Another Facet of Usability**. ACRL 2013. pp. 254-261.
- Chui, Charles K. (1992). **An Introduction to Wavelets**. San Diego: Academic Press.
- Kularbphetong, K., Tenprakhon, K. and Roonrakwit, P. (2014). Development a Recommendation Library System Based On Android Application. **International Journal of Computer, Electrical, Automation, Control and Information Engineering**. 8(5). pp. 834-836.
- Long, Xiaojian and Wu, Yuchun. (2012). **Borrowing Data Mining Based on Association Rules**. 2012 International Conference on Computer Science and Electronics Engineering. 239-242.
- Mehta, G., Sharma, D. and Chauhan, E. (2013). Application of Incremental Mining and Apriori Algorithm on Library Transactional Database. **International Journal of Computer Applications**. 73(8). pp. 13-18.
- Nicholson, Scott (2006), The basis for bibliomining: frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services, **Journal Information Processing and Management**, Vol. 42, Issue 3, pp. 785-804.
- Reeb, Brenda and Gibbons, Susan. (2004). Students, Librarians, and Subject Guides, **Portal: Libraries and the Academy**, 4, no. 1 (2004): pp. 123-130.
- Ru, Wen, Xin, Zhanhong, Gan, Di and Xing, Jun. (2015). A Study on Associations between Different Classifications of Library Circulation Data, **Journal of Advances in Information Technology**, Vol.6, No.3, 2015.
- Rui, Z. (2016), “The study of wind power predict model based on wavelet transform and Elman neural network”, **IEEE CCDC.**, pp. 6026 – 6030.
- Tsuji, Keita, Kuroo, Erika, Sato, Sho, Ikeuchi, Ui, Ikeuchi, Atsushi, Yoshikane, Fuyuki, and Itsumura, Hiroshi. (2012), **Use of Library Loan Records for Book Recommendation**, Proceedings of IIAI 2012, pp. 30-35.
- Uppal, Veepu and Chindwani, Gunjan (2013), “An Empirical Study of Application of Data Mining Techniques in Library System”, **International Journal of Computer Applications**, Volume 74, July, 2013, pp. 42-46.