

การขุดค้นความสัมพันธ์หมวดหมู่ของเพจบนเฟซบุ๊ก

Mining Page Category Association on Facebook

พัชรารณณ์ ช่วยเจริญ (Patcharapond Chouyjaroen)* และ พนิดา ทรงรัมย์ (Panida Songram)**

Received : October 20, 2017

Revised : May 21, 2018

Accepted : July 18, 2018

บทคัดย่อ

งานวิจัยนี้เสนอการขุดค้นความสัมพันธ์หมวดหมู่ของเพจบนเฟซบุ๊ก โดยประยุกต์ใช้อัลกอริทึมเอฟพี-โกรธ (FP-Growth Algorithm) ในการขุดค้นหมวดหมู่ของเพจที่เกิดขึ้นร่วมกันบ่อย จากนั้นสร้างกฎความสัมพันธ์จากหมวดหมู่ของเพจที่เกิดขึ้นร่วมกันบ่อย ข้อมูลที่ใช้ในการศึกษาหาความสัมพันธ์หมวดหมู่ของเพจบนเฟซบุ๊กเป็นข้อมูลการกดถูกใจเพจของผู้ใช้จำนวน 1,780 คน ทำการทดลองหาค่าสนับสนุนขั้นต่ำ (Minimum Support) และค่าความเชื่อมั่นขั้นต่ำ (Minimum Confidence) ที่เหมาะสมในการสร้างกฎ ซึ่งผลการทดลองแสดงให้เห็นว่าการกำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 10% ค่าความเชื่อมั่นขั้นต่ำเท่ากับ 50% ให้ค่าความถูกต้องมากที่สุด คือ 93.71% โดยกฎที่สร้างได้มีจำนวน 7,749 กฎ และสามารถนำมาใช้ได้จริง 892 กฎ กฎที่ได้สามารถนำมาใช้เพื่อเป็นแนวทางในการวางแผนโฆษณาหรือนำเสนอเพจให้ตรงตามความสนใจมากขึ้น

คำสำคัญ: เพจบนเฟซบุ๊ก กฎความสัมพันธ์ เอฟพี-โกรธ เหมือนข้อมูล เครือข่ายสังคมออนไลน์

Abstract

This paper was proposed to mine association of page category on Facebook. The FP-Growth algorithm co-occurrence is adopted to find frequently of page categories.

Then they are used to create association rules. The dataset used to study page category association is the set of page likes of 1,780 users. The minimum support and confidence thresholds are investigated to fine the appropriate values for generating the rules. From the investigation, the appropriate

minimum sup-port and confidence values are 10% and 50%, respectively. They give 93.71% of accuracy and use only 892 rules from 7,749 rules for prediction. The association rules can be used to plan or propose pages to interested users.

Keywords: Facebook Page Category, Association Rule, FP-Growth Algorithm, Data Mining, Online Social Networking.

1. บทนำ

ปัจจุบันมีการใช้เครือข่ายสังคมออนไลน์ (Social Network) ในการสื่อสารมากขึ้น เฟซบุ๊ก (Facebook) เป็นหนึ่งในเครือข่ายสังคมออนไลน์ที่ได้รับความนิยมมากที่สุดในประเทศไทย โซเชียลลิงค์ (Zocial Inc) [1] ได้เปิดเผยข้อมูลเกี่ยวกับสถิติและพฤติกรรมการใช้งานเครือข่ายสังคมออนไลน์ว่าประเทศไทยใช้เฟซบุ๊กมากเป็นอันดับ 9 ของโลก เท่ากับประเทศเยอรมนี โดยมีจำนวนผู้ใช้งานมากถึง 28 ล้านราย คิดเป็น 42% ของประชากรทั้งประเทศไทยและมีการขยายตัวเพิ่มขึ้นอย่างรวดเร็ว เนื่องจากผู้ใช้หรือสมาชิกสามารถใช้เฟซบุ๊กเป็นช่องทางในการสื่อสารพูดคุย แลกเปลี่ยนข้อมูล และทำกิจกรรมต่าง ๆ มากมายร่วมกันและสามารถใช้งานได้ฟรี เฟซบุ๊กจึงถือว่าเป็นโครงสร้างสังคมที่มีเป้าหมายเดียวกัน ซึ่งอยู่ในรูปแบบของการเชื่อมโยงที่เป็นเครือข่ายได้อย่างมีประสิทธิภาพ [2]

เฟซบุ๊กมีเครื่องมือหลากหลายที่อำนวยความสะดวกในการติดต่อสื่อสาร ทำให้มีผู้ใช้เฟซบุ๊กจำนวนมาก ดังนั้นเฟซบุ๊กจึงเป็นแหล่งทำการตลาดในยุคดิจิทัล เพจ (Page) บนพื้นที่บนเฟซบุ๊กที่นิยมใช้ในการตลาด สามารถประชาสัมพันธ์หรือโฆษณาสินค้าให้กับผู้ที่สนใจ ไม่ว่าจะเป็นในรูปแบบเชิงส่วนตัวหรือเชิงธุรกิจ [3] เมื่อมีผู้ใช้ให้ความสนใจใน

* ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

** Computer Science, Faculty of Informatics Mahasarakham University.

เพจนั้นๆ จะมีการแสดงความชื่นชอบหรือสนใจด้วยการกดถูกใจ (Like) ซึ่งทำให้ผู้ใช้สามารถรับข่าวสารจากเพจที่กดถูกใจ [4] จึงได้มีการโฆษณาเพจในเฟซบุ๊กเพื่อให้ผู้ใช้กดถูกใจ โดยปัจจุบันเพจถูกแนะนำให้ผู้ใช้งาน โดยอาศัยข้อมูลที่ได้จากการกำหนดข้อมูลเป้าหมายเองของผู้ให้บริการ [5] ซึ่งวิธีการนี้ไม่สามารถบ่งบอกได้ว่าเพจที่แนะนำจะตรงต่อความสนใจของผู้ใช้หรือไม่

ดังนั้นเพื่อนำเสนอหรือแนะนำเพจให้ตรงกับกลุ่มที่สนใจงานวิจัยนี้จึงศึกษาและสืบค้นความสัมพันธ์หมวดหมู่ของเพจในเฟซบุ๊ก โดยนำข้อมูลการกดถูกใจเพจในอดีตของผู้ใช้บนเฟซบุ๊กมาใช้ในการสร้างกฎความสัมพันธ์ด้วยอัลกอริทึมเอพี-กโรซ โดยกฎที่ได้สามารถแนะนำหมวดหมู่ของเพจที่ตรงกับความสนใจของผู้ใช้ได้ ซึ่งจะส่งผลดีทั้งผู้ใช้และผู้ให้บริการเพจ

2. การทบทวนวรรณกรรม

2.1 เฟซบุ๊ก

เฟซบุ๊ก คือ เครือข่ายสังคมออนไลน์ที่ใช้ในการติดต่อสื่อสาร และทำกิจกรรมร่วมกัน [6] เฟซบุ๊กได้เปิดตัวเมื่อวันที่ 4 กุมภาพันธ์ ปีพ.ศ. 2548 [7] โดย Mark Zuckerberg ถือเป็นเครือข่ายทางสังคมออนไลน์ที่ช่วยให้คนติดต่อสื่อสารกับเพื่อน คนในครอบครัว เพื่อนร่วมงาน และคนรู้จักได้อย่างมีประสิทธิภาพ เฟซบุ๊กได้มีพัฒนาเทคโนโลยีที่อำนวยความสะดวกในการแบ่งปันข้อมูล การส่งข้อความ การสร้างกลุ่ม การสร้างกิจกรรม การแชร์วิดีโอ และการสร้างพื้นที่แชร์ความรู้ อีกทั้งพัฒนาไปถึงการเข้าถึง เฟซบุ๊กผ่านมือถือ นอกจากนี้ยังได้มีการแปลเว็บไซต์เฟซบุ๊กเป็นภาษาต่างๆ ในปัจจุบัน เฟซบุ๊กได้มีการเจริญเติบโตขึ้นอย่างรวดเร็วและต่อเนื่อง ทำให้เฟซบุ๊ก กลายเป็นพื้นที่ที่นิยมในการทำประชาสัมพันธ์ เพื่อให้สามารถเข้าถึงผู้ใช้บริการได้อย่างกว้างขวาง

2.2 การทำเหมืองข้อมูลบนเว็บ (Web Mining)

การทำเหมืองข้อมูลบนเว็บ (Web Mining) เป็นการขุดค้นข้อมูลที่เป็นองค์ความรู้จากข้อมูลเว็บโดยใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) การทำเหมืองข้อมูลบนเว็บสามารถทำบนเนื้อหาข้อมูลบนเว็บหรือข้อมูลจากการใช้งานของผู้ใช้บนเว็บ การทำเหมืองข้อมูลบนเว็บถูกนำมาประยุกต์ใช้ในด้านต่างๆ เช่น การสำรวจความเห็นที่มีต่อสินค้าและบริการจากข้อความความคิดเห็นที่อยู่บนเว็บ การขุดค้นช่วงระยะเวลา

เวลาและลิงค์ที่ให้บริการเพื่อปรับเปลี่ยนการให้บริการที่เหมาะสมกับกลุ่มผู้ใช้ การทำเหมืองข้อมูลบนเว็บมีกระบวนการเหมือนการทำเหมืองข้อมูลทั่วไป ซึ่งประกอบไปด้วย [8] การทำความเข้าใจกับปัญหา การทำความเข้าใจข้อมูล การเตรียมข้อมูล การสร้างแบบจำลอง และการประเมินผล

การทำความเข้าใจประกอบไปด้วยการตั้งเป้าหมายว่าการทำเหมืองข้อมูลครั้งนี้ต้องการที่จะแก้ปัญหาใด มีการตั้งเกณฑ์วัดความสำเร็จในการทำเหมืองข้อมูล ซึ่งอาจเป็นได้ทั้งความสำเร็จในด้านรูปธรรม และความสำเร็จในด้านนามธรรม

การทำความเข้าใจข้อมูลประกอบด้วยการเก็บรวบรวมข้อมูลที่ต้องการใช้ กำหนดคุณลักษณะของข้อมูลที่เก็บรวบรวมได้ ตรวจสอบข้อมูลพร้อมตรวจสอบความสมบูรณ์และความถูกต้องของข้อมูล

การเตรียมข้อมูลประกอบด้วยการคัดเลือกข้อมูลในส่วนที่จำเป็นต้องใช้ ปรับเปลี่ยนรูปแบบข้อมูล เช่น นำข้อมูลจากหลายตารางมาเชื่อมกัน ทำความสะอาดข้อมูล เตรียมข้อมูลให้เหมาะสมที่สุด ซึ่งมีหลายวิธี เช่น การแก้ไขข้อมูลที่ ไม่สมบูรณ์หรือผิดพลาด การเติมค่าว่างของข้อมูล เป็นต้น จากนั้นทำการแปลงข้อมูลให้เหมาะสมกับโปรแกรมที่ประมวลผล เลือกลักษณะเฉพาะข้อมูลที่ต้องการวิเคราะห์ ลดความซ้ำซ้อนของข้อมูล เพื่อเพิ่มเนื้อที่และลดระยะเวลาในการประมวลผล

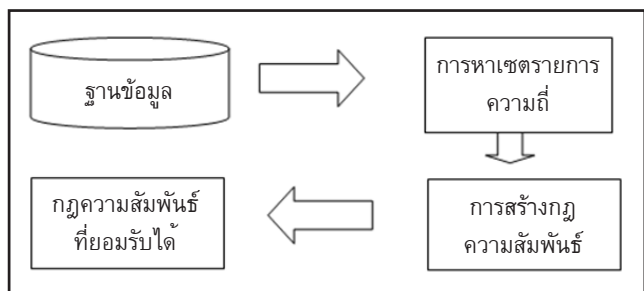
การสร้างแบบจำลองประกอบด้วยการเลือกอัลกอริทึมที่เหมาะสมกับข้อมูล กำหนดรูปแบบการทดสอบผลลัพธ์ของข้อมูล ทดสอบข้อมูลกับแบบจำลอง วัดความถูกต้องและความน่าเชื่อถือ ว่าไปในทิศทางไหน

การประเมินผลแบบจำลองที่ทำการสร้างขึ้นโดยการนำไปทดสอบใช้กับสถานการณ์จริง หรือเหตุการณ์จำลอง เพื่อดูว่าได้ผลมากน้อยเพียงใด มีข้อผิดพลาดเกิดขึ้นหรือไม่ เมื่อผิดพลาดต้องมีการแก้ไขด้วยวิธีใด และดำเนินการแก้ไขก่อนนำมาใช้งานจริง เพื่อตรวจสอบว่าผลที่ได้บรรลุเป้าหมายมากน้อยเพียงใด

2.3 การสร้างกฎความสัมพันธ์ (Association Rule)

กฎความสัมพันธ์ [9] เป็นการค้นหาความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่ที่มีอยู่ เพื่อนำไปใช้ในการวิเคราะห์หรือทำนายปรากฏการณ์ต่างๆ เช่น การวิเคราะห์การซื้อสินค้าของลูกค้า กระบวนการการหากฎความสัมพันธ์

สามารถแบ่งได้เป็น 2 ขั้นตอนหลัก คือ การหาเซตรายการความถี่ของกลุ่มข้อมูลที่มีความถี่ร่วมกันทั้งหมด และการสร้างกฎ ซึ่งแสดงได้ดังภาพที่ 1



ภาพที่ 1 กระบวนการค้นหากฎความสัมพันธ์

การหาเซตรายการความถี่ (Frequent Itemset Mining) เป็นการค้นหาเซตรายการที่มีค่านับสนับสนุนมากกว่าหรือเท่ากับค่านับสนับสนุนขั้นต่ำที่ผู้ใช้เป็นผู้กำหนดเองตามความเหมาะสมของแต่ละงาน จากนั้นนำเซตรายการที่มีความถี่รวมกันมาสร้างกฎความสัมพันธ์ โดยกฎความสัมพันธ์ที่ยอมรับได้เป็นกฎที่มีค่าความเชื่อมั่นมากกว่าหรือเท่ากับค่าความเชื่อมั่นขั้นต่ำที่ผู้ใช้กำหนดขึ้น นิยามที่เกี่ยวข้องกับการหากฎความสัมพันธ์สามารถแสดงได้ดังนี้

กำหนดให้ $I = \{i_1, i_2, \dots, i_m\}$ คือ เซตของรายการทั้งหมด และ $D = \{t_1, t_2, \dots, t_m\}$ คือ เซตของทรานแซกชันทั้งหมด (Transaction) และกำหนดให้ X คือ เซตรายการ โดยที่ $X \subseteq I$ บทนิยามที่ 1 ค่านับสนับสนุนของ X สามารถคำนวณได้ดังสมการ 1

$$supp(X) = \frac{frequency(X)}{|D|} \cdot 100 \quad (1)$$

บทนิยามที่ 2 กฎความสัมพันธ์ คือ กฎที่สร้างจากกลุ่มเซตรายการที่มีความถี่รวมกัน โดยอยู่ในรูปของ $X \rightarrow Y$ แทนเหตุการณ์ของการเกิด X แล้วเกิดเหตุการณ์ Y ร่วมกัน

บทนิยามที่ 3 ค่าความเชื่อมั่น คือ ความน่าเชื่อถือของกฎความสัมพันธ์ที่สร้างขึ้น โดยสามารถคำนวณได้ดังสมการ 2

$$conf(X) = \frac{supp(X \dot{\cup} Y)}{supp(X)} \cdot 100 \quad (2)$$

ปัจจุบันมีการพัฒนาอัลกอริทึมสำหรับสืบค้นเซตรายการความถี่จำนวนมาก อัลกอริทึมเอพี-ทรี (FP-Tree) เป็นอัลกอริทึมหนึ่งที่มีประสิทธิภาพในการหาเซตรายการความถี่หรือการหารูปแบบของข้อมูลที่เกิดขึ้นร่วมกันบ่อยๆ ในฐานข้อมูล เนื่องจากอัลกอริทึมนี้ไม่ได้สร้างแคนดิเดต (Candidate) และ

มีการอ่านฐานข้อมูลเพียงสองครั้ง โดยนำข้อมูลไปเก็บในโครงสร้างเอพี-ทรี (FP-Tree-Frequent-Pattern-Tree) จากนั้นทำการสร้างเซตของคอนดิชันนอลแพทเทินเบซ (Conditional Base Pattern) จากเอพี-ทรี แล้วนำค่าที่ได้ไปหาเซตรายการความถี่

เมื่อได้เซตรายการความถี่แล้วจะทำการสร้างกฎความสัมพันธ์ ซึ่งสร้างจากเซตรายการความถี่ที่มีความยาว 2 รายการขึ้นไป โดยจำนวนกฎที่ได้จากเซตรายการความถี่แต่ละอันจะเท่ากับ $2^n - 2$ โดยที่ n คือจำนวนรายการ

2.4 งานวิจัยที่เกี่ยวข้อง

ณรงค์ศักดิ์ และจิรัฐฐา [10] นำเสนอการประยุกต์ใช้เอพี-ทรีกับงานแนะแนวการศึกษาต่อในระดับอุดมศึกษา โดยนำผลการเรียน 7 รายวิชาหลัก คือ คณิต ไทย ชีวศึกษา ฟิสิกส์ เคมี สังคมศึกษา อังกฤษ ของนักเรียนชั้นมัธยมศึกษาตอนปลายเฉพาะสายวิทย์-คณิตมาทำการค้นหากฎความสัมพันธ์ เพื่อใช้ในการแนะนำแนวทางศึกษาต่อให้กับนักเรียน ผลจากการทดลองแสดงให้เห็นว่าสามารถแนะแนวการศึกษาได้ถูกต้องที่ 89.87%

ณัฐธิดา และอันธิกา [11] เสนอการหาปัจจัยที่ส่งผลต่อผลการเรียนของนักศึกษาเรียนอ่อน โดยใช้เทคนิคกฎความสัมพันธ์ เพื่อทำนายความเสี่ยงของการเรียนอ่อนของนักศึกษาชั้นปีที่ 1 ในคณะต่างๆ ของมหาวิทยาลัยสงขลานครินทร์ งานวิจัยนี้ใช้ข้อมูลของนักศึกษาที่ เก็บไว้แล้วในฐานข้อมูลของมหาวิทยาลัยตั้งแต่ปี พ.ศ. 2550-2552 ผลการศึกษาพบว่าปัจจัยที่ส่งผลทางการศึกษาของนักศึกษาในแต่ละคณะแตกต่างกันและความถูกต้องของผลลัพธ์ก็แตกต่างกัน อันเนื่องมาจากข้อมูลนำเข้าของนักศึกษาในแต่ละคณะต่างกัน ดังนั้นผลลัพธ์และความถูกต้องของกฎความสัมพันธ์จึงขึ้นอยู่กับข้อมูลที่ใช้ในการเรียนรู้ เมื่อมองในภาพรวมแล้วพบว่าปัจจัยที่ส่งผลต่อผลการเรียนของนักศึกษาเรียนอ่อนคือ วิธีการเข้าศึกษา คะแนนสอบเข้าศึกษาของวิชาต่างๆ และเพศของนักศึกษา

กฤษณพร และกมล [12] หากฎความสัมพันธ์ของผลิตภัณฑ์ประกันภัยกับพฤติกรรมของลูกค้าที่ถือครองผลิตภัณฑ์ ซึ่งสามารถนำความสัมพันธ์ดังกล่าวมาช่วยตัดสินใจการขายผลิตภัณฑ์ประกันภัยให้กับลูกค้าผู้สูงวัย กลุ่มบัญชีออมทรัพย์ของธนาคาร ทั้งนี้ยังสามารถนำเอาความรู้ที่ได้มาประยุกต์กับระบบเดิมที่มีอยู่ในองค์กรหรือ

พัฒนาซอฟต์แวร์ใหม่ เพื่อเพิ่มโอกาสในการขายและกระบวนการทำงานให้มีประสิทธิภาพมากขึ้น

กฤษณะ และคณะ [13] นำเสนอการใช้เทคนิคเหมืองข้อมูลเพื่อพัฒนาคุณภาพการศึกษาการศึกษาคณะวิศวกรรมศาสตร์ โดยการนำโครงสร้างต้นไม้ การพยากรณ์ข้อมูล และการหารูปแบบความสัมพันธ์ มาประยุกต์ใช้ในการแนะนำแนวทางให้กับนิสิตและการตัดสินใจในการลงทะเบียน

ชฎิกภรณ์ และคณะ [14] นำเสนอการพัฒนาแบบจำลองพยากรณ์ภาวะแทรกซ้อนของโรคอื่นในผู้ป่วยโรคเบาหวานกรณีศึกษาโรงพยาบาลปทุมธานี เปรียบเทียบผลการพยากรณ์ระหว่างเทคนิคเอพริออรี (Apriori Algorithm) และเอพี-โกรธ โดยกำหนดค่าสนับสนุน 1% ความเชื่อมั่น 20%

กิตติศักดิ์ [15] นำเสนอการคัดกรองสุขภาพเบื้องต้นโดยใช้เทคนิคเอพริออรีและเอพี-โกรธ ในการหาความสัมพันธ์ของโรคต่างๆ จากผลการทดลองให้ค่าความถูกต้อง 88.60%

Hsua [16] นำการทำนายประสิทธิภาพการเรียนของนักศึกษา ซึ่งใช้ข้อมูลประวัติการศึกษาและข้อมูลรายวิชาที่ได้รับจากมหาวิทยาลัยในไต้หวัน เพื่อที่จะทำนายแนวโน้มที่จะผ่านหรือไม่ผ่านรายวิชาหนึ่งๆ เมื่อผ่านการสอบกลางภาคไปแล้วหนึ่งสัปดาห์ โดยประยุกต์ใช้การสร้างต้นไม้การค้นหากฎความสัมพันธ์ และวิธีทางพันธุกรรม (Genetic Algorithm) รวมกัน

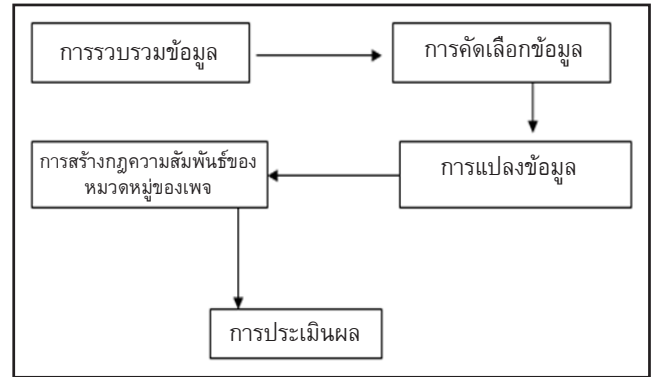
Choi YK [17] นำเสนอระบบแนะนำสินค้าด้วยวิธี Collaborative Filtering และวิธี Item-based Collaborative Filtering ร่วมกับการแนะนำสินค้าที่เกี่ยวข้องโดยใช้กฎความสัมพันธ์ โดยมุ่งไปที่สินค้าที่เกี่ยวข้องกับกลุ่มผู้ซื้อเป็นหลัก

Shunmin [18] นำเสนอการทำเหมืองข้อมูลในโรงพยาบาลโดยใช้เทคนิคกฎความสัมพันธ์ ทำให้ค้นพบกฎที่เกี่ยวข้องระหว่างโรคและปัจจัยที่สัมพันธ์กัน อย่างไรก็ตามงานวิจัยนี้ยังต้องการศึกษาเพิ่มเติมเพื่อให้สามารถแสดงกฎที่มีความสำคัญ แต่มีค่าสนับสนุนน้อยทำให้ไม่สามารถแสดงกฎดังกล่าวได้

จากงานวิจัยที่เกี่ยวข้องจะเห็นได้ว่ามีการนำกฎความสัมพันธ์ไปประยุกต์กับงานด้านต่างๆ และให้ผลค่าความถูกต้องที่สูง นอกจากนี้กฎที่ได้เป็นกฎที่เข้าใจง่าย สามารถนำไปใช้ได้เลย ดังนั้นงานวิจัยนี้จึงหาความสัมพันธ์หมวดหมู่ของเพจโดยประยุกต์ใช้การหากฎความสัมพันธ์

3. วิธีดำเนินการวิจัย

งานวิจัยนี้นำเสนอการขุดค้นความสัมพันธ์หมวดหมู่ของเพจบนเฟซบุ๊ก โดยมีวิธีการดำเนินการวิจัยแสดงดังภาพที่ 2 และมีรายละเอียดดังหัวข้อต่อไปนี้



ภาพที่ 2 วิธีดำเนินการวิจัย

3.1 การรวบรวมข้อมูล

งานวิจัยนี้ได้ทำการสร้างกฎความสัมพันธ์ของหมวดหมู่ของเพจ โดยรวบรวมข้อมูลการกดถูกใจเพจของผู้ใช้จำนวน 1,780 คนในช่วงวันที่ 25 เมษายน - 30 เมษายน พ.ศ. 2560 ซึ่งการกดถูกใจเป็นการแสดงความพึงพอใจ ความสนใจหรือความชอบตามความรู้สึกจริงและมีความน่าเชื่อถือ [13] การดึงข้อมูลเพจจากการกดถูกใจของผู้ใช้ในเฟซบุ๊ก สามารถดึงได้โดยใช้เฟซบุ๊กกราฟเอพีไอ (Facebook Graph Api) เมื่อมีการส่งคำร้องขอข้อมูลผ่านกราฟเอพีไอแล้ว มีการนำข้อมูลแสดงผลในรูปแบบของเจสัน (Json) ดังภาพที่ 3 [14] ใช้คำสั่งของ PHP เพื่อนำข้อมูลที่ได้ออกมาเก็บรวบรวมลงในฐานข้อมูลมายเอสคิวแอล (MySQL) เพื่อไว้ใช้ในการประมวลผลต่อไป

```

{
  "id": "100000346893678",
  "likes": {
    "data" [
      {
        "category": "Community",
        "name": "Marvel Thailand Fanpage",
        "create_time": "2015-01-06T07:55:47+0000",
        "id": "259216647506584"
      },
      {
        "category": "Sports team",
        "name": "ฟุตบอลทีมชาติไทย",
        "create_time": "2014-12-23T14:41:26+0000",
        "id": "189387414416552"
      },
      {
        "category": "Entertainer",
        "name": "ข้อความโดนๆ",
        "create_time": "2014-12-16T11:04:16+0000",
        "id": "164486926939395"
      }
    ]
  }
}
  
```

ภาพที่ 3 ตัวอย่างข้อมูลที่ดึงจากกราฟเอพีไอ

3.2 การคัดเลือกข้อมูล

ขั้นตอนแรกทำการดึงหมวดหมู่ของเพจที่ผู้ใช้สนใจมาทั้งหมด จากนั้นทำการคัดเลือกหมวดหมู่ของเพจที่ผู้ใช้กดถูกใจจำนวน 10 เพจขึ้นไปในหมวดหมู่เดียวกันออกมาเก็บในฐานะข้อมูล ยกตัวอย่างเช่น ผู้ใช้กดถูกใจเพจทั้งหมด 10 เพจ ซึ่งเพจดังกล่าวอยู่ในหมวดหมู่เดียวกัน แสดงว่าผู้ใช้สนใจหมวดหมู่นั้นจริง หมวดหมู่นั้นก็จะถูกเลือกนำมาใช้เป็นข้อมูลในงานวิจัย

3.3 การแปลงข้อมูล

งานวิจัยนี้ทำการค้นหาความสัมพันธ์ของหมวดหมู่ที่ผู้ใช้สนใจโดยเทคนิคการหาความสัมพันธ์ เพื่อให้ข้อมูลสามารถนำไปใช้กับเทคนิคการหาความสัมพันธ์ จึงทำการแปลงข้อมูลให้อยู่ในรูปแบบตัวเลขที่สามารถใช้ได้ด้วยเทคนิคการหาความสัมพันธ์ ซึ่งตัวอย่างการแปลงข้อมูลในตารางที่ 1

ตารางที่ 1 ตัวอย่างการแปลงสัญลักษณ์ข้อมูล

สัญลักษณ์ค่าแทนที่	หมวดหมู่
1	Transport
2	Sport
3	Shopping / Retail
4	Government
5	Address
...
75	Bank

โดยในงานวิจัยนี้ กำหนดให้ รายการ คือ หมวดหมู่ ดังนั้น 1 รายการ คือ 1 หมวดหมู่ และเซตรายการหมายถึงเซตของหมวดหมู่ ดังตัวอย่างตารางที่ 2

3.4 การสร้างกฎความสัมพันธ์ของหมวดหมู่เพจ

การสร้างกฎความสัมพันธ์ในงานวิจัยนี้ [15] ประกอบไปด้วยขั้นตอนใหญ่ 2 ขั้นตอน คือ ขั้นตอนแรกหาเซตรายการความถี่ ซึ่งเป็นเซตรายการที่มีค่าสนับสนุนมากกว่าหรือเท่ากับค่าสนับสนุนขั้นต่ำ งานวิจัยนี้ได้ประยุกต์ใช้อัลกอริทึมเอพี-กิโรในการหาเซตรายการความถี่ ซึ่งมีขั้นตอนดังต่อไปนี้

1) อ่านข้อมูลจากฐานข้อมูล เพื่อหาเซตรายการความถี่ที่มี 1 รายการที่พบว่ามีค่าสนับสนุนมากกว่าหรือเท่ากับ

ตารางที่ 2 ตัวอย่างผลการแปลงข้อมูล

Transaction	ITEMSET																				
	1	2	3	5	4	8	9	10	12	13	11										
1																					
2	15	13	9	16	5	17	10	7	6	11	21	23	3	18	8	1	19	22			
3	17	14	13	6	15	11	22	19	2	1	18	8	21	16	10	20					
4	2	11	21	8																	
5	13	17	15	8	2	4	5	18	16	22	25	6	11	1	21						
6	18	13	17	4	6	10	14	2	11	15	16	3									
7	5	3	6	18	19	17	8	12	2	5	6	13	11	10	24	22	15	4	13		
8	16	14	1	2	3	11	17	15	22	12	3	19	9	8	19	4	24				
9	22	8	5	18	24	16	17	13	2	12	13	1	15								
10	6	5	17	3	22	4	2	11	10	8	16	21	18	15							

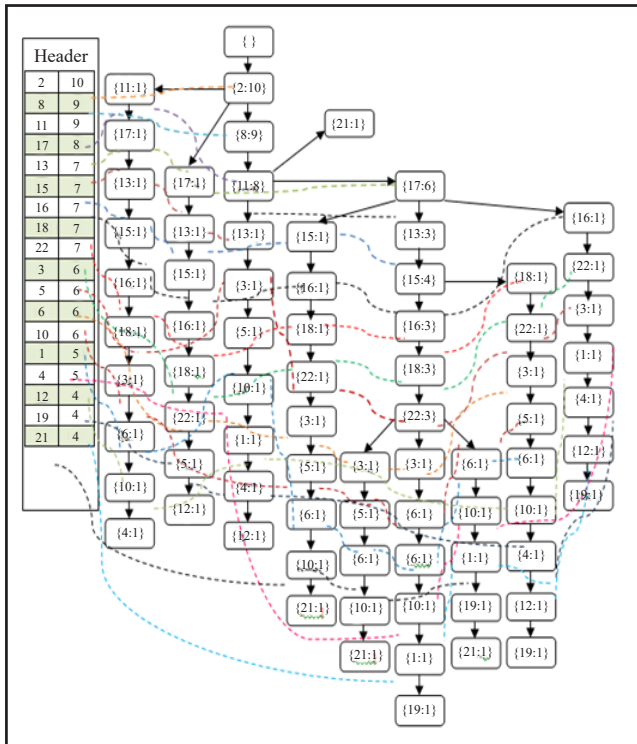
ตารางที่ 3 เซตรายการความถี่

รายการ	ความถี่	รายการ	ความถี่
1	5	16	7
2	10	13	7
3	6	14	3
4	5	15	7
5	6	16	7
6	6	17	8
7	1	18	7
8	9	19	4
9	3	20	1
10	6	21	4
11	9	22	7
12	4	23	1
13	7	24	3
14	3	25	1
15	7		

ค่าสนับสนุนขั้นต่ำที่กำหนด จากตารางที่ 3 สมมุติกำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 40% หรือความถี่เท่ากับ 4

2) เรียงเซตรายการความถี่ที่มี 1 รายการจากค่าความถี่มากไปน้อยใน f-list จากนั้นเรียงเซตรายการในแต่ละทรานแซคชันใหม่ตามรายการใน f-list

3) อ่านเซตรายการที่ละทรานแซคชันเพื่อสร้างเอฟพี-ทรี ซึ่งจะได้ดังภาพที่ 4



ภาพที่ 4 FP-Tree

4) ทำการขุดค้นเซตรายการความถี่ โดยเริ่มจากพิจารณา โหนดที่อยู่ใน Header Table ลำดับสุดท้าย ทำการหาเซต รายการความถี่ทุกโหนดจนครบทุกโหนดรายการ สามารถ สรุปรายการความถี่ที่หาได้ในรูปแบบ เซตรายการ ความถี่:ค่าสนับสนุน (%) ได้ดังนี้ {21}:40% {11, 21}:40% {8, 11, 21}:40% {2, 8, 11, 21}:40% {8, 21}:40% {2, 11, 21}:40% {12, 21}:40% {2, 8, 21}:40% {19}:40% {17, 19}:40% {11, 17, 19}:40% {2, 11, 17, 19}:40% {8, 2, 11, 17, 19}:40% {11, 19}:40% {17, 2, 19}:40% {8, 17, 2, 19}:40% {8, 19}:40% {2, 8, 19}:40% 17, 2, 8, 19}:40% {2, 19}:40% {8, 11, 19}:40% {2, 8, 11, 19}:40% {12}:40% {8,12}:40% {12, 8, 12}:% {2, 12}:44% {1}:50% {11, 1}:50% {8, 11, 1}:50% {2, 8, 11, 1}:50% {8, 1}:50% {2, 11, 1}:50% {2, 1}:50% {2, 8, 1}:50% {10}:60% {2, 10}:60% {6}:60% {2, 16}:60% {5}:60% {8, 5}:60% {2, 8, 5}:60% {2, 5}:60% {3}:60% {2, 3}:60% {22}:60% {8, 22}:60% {2, 8, 22}:60% {2, 22}:60% {18}:70% {2, 18}:70% {16}:70% {12,16}:70% {15}:70% {2, 15}:70% {13}:70% {2, 13}:70% {17}:80% {2, 17}:80% {11}:90% {2,11}:80% {8}:90%

ขั้นตอนที่สอง หากกฎความสัมพันธ์จากเซตรายการ ความถี่ที่มีความยาว 2 รายการขึ้นไป โดยจำนวนกฎที่ได้ จากเซตรายการความถี่แต่ละอันจะเท่ากับ $2^n - 2$ โดยที่ n คือ จำนวนรายการ เช่น เซตรายการความถี่ {8, 11, 21} สามารถ สร้างกฎได้ $2^3 - 2 = 6$ กฎ ดังตารางที่ 4 สมมติกำหนด ค่าความเชื่อมั่นขั้นต่ำ 60% กฎที่ยอมรับได้ คือ 8, 12 \rightarrow 11 และ 11, 21 \rightarrow 8

ตารางที่ 4 ตัวอย่างกฎความสัมพันธ์

กฎ	ค่าความเชื่อมั่น
8 \rightarrow 11, 21	33%
11 \rightarrow 8, 21	44%
21 \rightarrow 8,11	33%
8, 11 \rightarrow 21	50%
8, 21 \rightarrow 11	100%
11, 21 \rightarrow 8	80%

3.5 การเรียงกฎ

หลังจากได้กฎความสัมพันธ์ที่มีค่าความเชื่อมั่นมากกว่า หรือเท่ากับค่าความเชื่อมั่นขั้นต่ำทั้งหมดแล้ว กฎจะถูกนำมา พิจารณาเพื่อเรียงกฎแล้วนำไปใช้ โดยการเรียงกฎมี ดังต่อไปนี้

กำหนดให้กฎสองกฎ คือ R_i และ R_j โดย R_i มีลำดับ การเกิดก่อน R_j โดยกฎ R_i จะโดนนำไปเรียงกฎก่อน R_j เมื่อ ตรงเงื่อนไขต่อไปนี้

- 1) $conf(R_i) > conf(R_j)$ หรือ
- 2) ถ้า $conf(R_i) = conf(R_j)$ แต่ $supp(R_i) > supp(R_j)$ หรือ
- 3) $supp(R_i) = supp(R_j)$ แต่ $size(R_i) < size(R_j)$ ซึ่งหมายถึง ความยาวของ Antecedent ของ R_i จะต้องสั้นกว่าความยาว ของ Antecedent ของ R_j หรือ

4) ถ้า $size(R_i) = size(R_j)$ แต่ R_i มีลำดับการเกิดก่อน R_j ยกตัวอย่างจากตารางที่ 4 กฎ 8, 12 \rightarrow 11 จะถูกเรียง ก่อนกฎ 11,21 \rightarrow 8 เนื่องจากค่าความเชื่อมั่นของกฎ 8, 12 \rightarrow 11 มากกว่ากฎ 11,21 \rightarrow 8

3.6 การวัดประสิทธิภาพ (Evaluation)

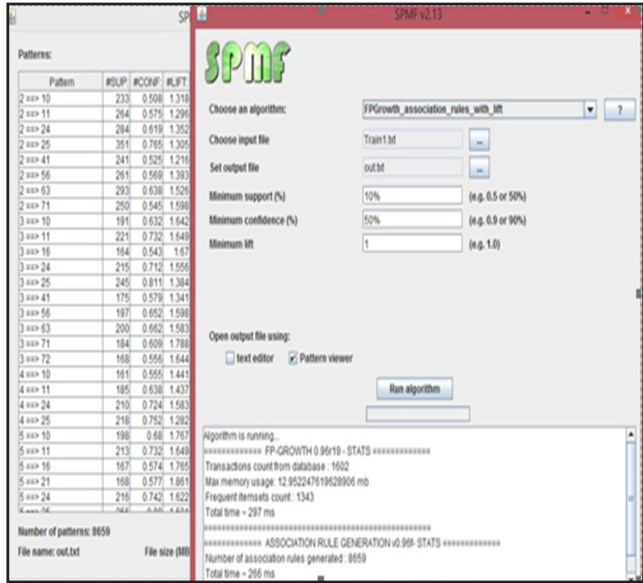
งานวิจัยนี้ใช้ k-fold Cross Validation ในการแบ่งข้อมูล เรียงรู้ (Training) และข้อมูลทดสอบ (Testing Set) โดย กำหนดค่า $K = 10$ ซึ่งเป็นค่า K ที่นิยมใช้ในการแบ่งข้อมูล และเหมาะกับการวัดประสิทธิภาพข้อมูลที่มีจำนวนไม่มาก

ทำการวัดประสิทธิภาพทั้งหมด 10 รอบ ในรอบแรกข้อมูลชุดแรกเป็นข้อมูลทดสอบประสิทธิภาพของโมเดลและข้อมูลชุดที่เหลือเป็นชุดเรียนรู้ ในรอบที่สองข้อมูลชุดที่สองเป็นข้อมูลทดสอบและข้อมูลชุดที่เหลือสร้างตัวจำแนก ทำวนไปเช่นนี้จนครบ 10 รอบ ในแต่ละรอบจะวัดค่าความถูกต้องดังสมการที่ 3 แล้วคำนวณค่าเฉลี่ยค่าความถูกต้องโดยคำนวณได้ดังสมการที่ 4

$$Accuracy = (|Ra|/|R|) \times 100 \quad (3)$$

โดยที่ $|Ra|$ คือ จำนวนข้อมูลที่ตรงกับกฎที่ได้
 $|R|$ คือ จำนวนข้อมูลทั้งหมดที่ใช้ในการทดสอบ

$$AVG = \left(\frac{\sum_{i=1}^k Accuracy}{K} \right) \quad (4)$$



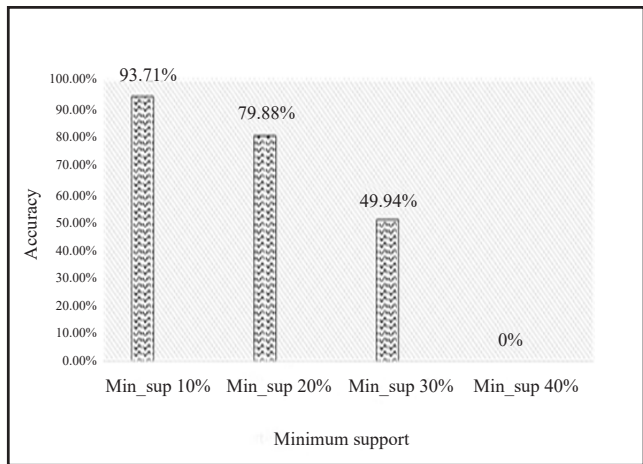
ภาพที่ 5 ตัวอย่างโปรแกรม SPMF

4. ผลการดำเนินการวิจัย

การค้นหาค่าความสัมพันธ์ของหมวดหมู่ของเพจที่ผู้ใช้ในเฟซบุ๊กให้ความสนใจนั้น เป็นข้อมูลที่ได้จากการเก็บข้อมูลความชอบหรือความสนใจในอดีตในเฟซบุ๊ก ข้อมูลรวบรวมผ่านการดึงข้อมูลผ่านกราฟ API ของเฟซบุ๊ก (www.facebook.com) เวอร์ชัน 2.8 (Graph API V.2.8) ข้อมูลที่ถูกดึงมาจากกราฟ API และยังไม่ผ่านการคัดกรองข้อมูลมีจำนวนมาก เนื่องจากข้อมูลยังมีการกระจายและซ้ำกันอยู่เป็นจำนวนมาก จึงต้องมีการคัดเลือกและจัดกลุ่มเพจให้อยู่ในกลุ่มหมวดหมู่เดียวกัน และเลือกเฉพาะหมวดหมู่ที่มีจำนวนกดถูกใจ 10 ครั้งขึ้นไป เพื่อเป็นการบ่งบอกว่าเป็นสิ่งที่ผู้ใช้สนใจจริงๆ

ข้อมูลที่ผ่านการเตรียมข้อมูลและแปลงข้อมูลในรูปแบบทรานแซคชันเรียบร้อยแล้วมาทำการทดลองกับ SPMF (<http://www.philippe-fournierviger.com/spmf/>) ดังภาพที่ 5 โปรแกรม โอเพ่นซอส ดาต้าไมนิ่งไลบรารี (Open – Source Data Mining Library) V.2.13

ในงานวิจัยนี้ได้ทำการศึกษาค่าสนับสนุนขั้นต่ำและค่าความเชื่อมั่นขั้นต่ำที่เหมาะสม โดยทดลองกำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 10% 20% 30% และ 40% และกำหนดค่าความเชื่อมั่นขั้นต่ำเท่ากับ 50% ผลการทดลองวัดประสิทธิภาพความถูกต้องแสดงได้ดังภาพที่ 4 และจำนวนกฎที่ได้จากการผลการทดลองแสดงได้ดังภาพที่ 6 จากภาพที่ 6 และตารางที่ 5 จะเห็นได้ว่าเมื่อกำหนด



ภาพที่ 6 ความถูกต้องเมื่อกำหนดค่าสนับสนุนขั้นต่ำเปลี่ยนแปลงและความเชื่อมั่นขั้นต่ำ 50%

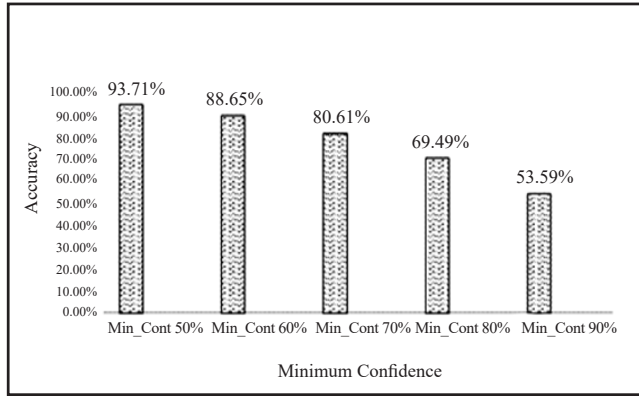
ตารางที่ 5 จำนวนกฎที่สร้างขึ้นเมื่อกำหนดค่าสนับสนุนขั้นต่ำเปลี่ยนแปลงและค่าความเชื่อมั่นขั้นต่ำ 50%

Min_Sup	Min_Conf	จำนวนกฎทั้งหมด	กฎที่นำมาใช้	กฎที่นำมาใช้ (%)
10%	50%	7,749	892	11.64%
20%	50%	94	77	81%
30%	50%	4	4	66%
40%	50%	0	0	0

ค่าสนับสนุนขั้นต่ำเท่ากับ 10% จะให้ค่าความถูกต้องสูงที่สุดคือ 93.71% แต่อย่างไรก็ตามจำนวนกฎที่สร้างขึ้นมีจำนวนมากและถูกนำไปใช้แค่ 11.64%



การทดลองที่สองเป็นการศึกษาค่าความเชื่อมั่นขั้นต่ำที่เหมาะสม โดยกำหนดค่าความเชื่อมั่นขั้นต่ำเท่ากับ 50% 60% 70% 80% และ 90% ค่าสนับสนุนขั้นต่ำเท่ากับ 10% จากผลการทดลองวัดประสิทธิภาพความถูกต้อง และจำนวนกฎที่ได้จากการผลการทดลองแสดงได้ดังภาพที่ 7 และตารางที่ 6



ภาพที่ 7 ความถูกต้องของการกำหนดค่าความเชื่อมั่นขั้นต่ำที่เปลี่ยนแปลงและค่าสนับสนุนขั้นต่ำ 10%

ตารางที่ 6 จำนวนกฎที่สร้างขึ้นเมื่อค่าความเชื่อมั่นขั้นต่ำเปลี่ยนแปลงและค่าสนับสนุนขั้นต่ำ 10%

Min_Sup	Min_Conf	จำนวนกฎทั้งหมด	กฎที่นำมาใช้	กฎที่นำมาใช้ (%)
10%	50%	7,749	892	11.64%
10%	60%	5,609	747	13%
10%	70%	3,579	612	17%
10%	80%	1,731	504	29%
10%	90%	405	278	68%

จากภาพที่ 7 และตารางที่ 6 จะเห็นได้ว่าค่าความเชื่อมั่นขั้นต่ำเท่ากับ 50% ให้ค่าความถูกต้องมากที่สุด ซึ่งการกำหนดค่าความเชื่อมั่นไม่ควรจะกำหนดให้มีค่าต่ำกว่า 50% เนื่องจากจะทำให้กฎขาดความน่าเชื่อถือในการนำไปพิจารณาความสัมพันธ์ของกฎหมวดหมู่ของเพจ แต่ถ้าค่าความเชื่อมั่นขั้นต่ำยิ่งสูง ยิ่งจะทำให้กฎความสัมพันธ์ที่มีความน่าเชื่อถือสูงตามไปด้วย

จากการทดลองในงานวิจัยนี้ จะเห็นได้ว่าเมื่อกำหนดค่าสนับสนุนขั้นต่ำที่ 10% ได้ค่าความถูกต้องที่สูงกว่าค่าสนับสนุนขั้นต่ำอื่น เนื่องจากยังกำหนดค่าสนับสนุนขั้นต่ำที่น้อยลง ทำให้ได้กฎเป็นจำนวนมากและความถูกต้อง

สูงตามไปด้วย ดังนั้นงานวิจัยนี้พิจารณาเลือกค่าสนับสนุนขั้นต่ำเท่ากับ 10% ค่าความเชื่อมั่น 50% เนื่องจากให้ประสิทธิภาพความถูกต้องที่สูงกว่าค่าอื่น

จากการทดสอบความถูกต้อง กฎที่อยู่อันดับต้นๆ ที่ได้จากการเรียงกฎเป็นกฎที่ถูกนำไปใช้มากที่สุด ดังนั้นงานวิจัยนี้จึงรายงานกฎ 10 อันดับแรกที่ได้จากการกำหนดค่าสนับสนุนขั้นต่ำเท่ากับ 10% และค่าความเชื่อมั่นขั้นต่ำเท่ากับ 50% ดังตารางที่ 7

ตารางที่ 7 กฎแรกเมื่อกำหนดค่าสนับสนุนขั้นต่ำ 10% และค่าความเชื่อมั่นขั้นต่ำ 50%

ลำดับกฎ	กฎที่นำไปใช้	ค่าสนับสนุน (%)	ค่าความเชื่อมั่น (%)
1	11 51 56 >> 41	15%	0.95
2	24 25 68 >> 41	14%	0.95
3	24 68 >> 41	13%	0.93
4	10 21 25 72 >> 24	12%	0.93
5	25 33 68 >> 41	12%	0.93
6	10 11 25 56 72 >> 24	11%	0.93
7	25 33 63 72 >> 56	11%	0.93
8	21 25 63 72 >> 24	10%	0.92
9	10 11 25 63 72 >> 24	10%	0.92
10	11 51 63 >> 41	10%	0.91

จากตารางที่ 7 สามารถอธิบายกฎได้ดังนี้

กฎที่ 1 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่การบริการ หมวดหมู่สินค้าและการบริการ, หมวดหมู่บุคคลสาธารณะ >> มักจะมีการกดถูกใจหมวดหมู่ชุมชน ด้วยความเชื่อมั่น 95%

กฎที่ 2 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่เว็บไซต์ หมวดหมู่บริษัท หมวดหมู่นิตยสาร >> มักจะมีการกดถูกใจหมวดหมู่ชุมชน ด้วยความเชื่อมั่น 95%

กฎที่ 3 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่เว็บไซต์, หมวดหมู่นิตยสาร >> มักจะมีการกดถูกใจหมวดหมู่ชุมชน ด้วยความเชื่อมั่น 93%

กฎที่ 4 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่ธุรกิจ หมวดหมู่สถาบันการศึกษา หมวดหมู่บริษัท หมวดหมู่อาชีพ >> มักจะมีการกดถูกใจหมวดหมู่เว็บไซต์ ด้วยความเชื่อมั่น 93%

กฎที่ 5 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่บริษัท หมวดหมู่สื่อ/ข่าวสาร/สิ่งพิมพ์ หมวดหมู่นิตยสาร >> มักจะมีการกดถูกใจหมวดหมู่ชุมชน ด้วยความเชื่อมั่น 93%

กฎที่ 6 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่ธุรกิจ หมวดหมู่บริการ หมวดหมู่บริษัท หมวดหมู่บุคคลสาธารณะ หมวดหมู่อาชีพ >> มักจะมีการกดถูกใจหมวดหมู่เว็บไซต์ ด้วยความเชื่อมั่น 93%

กฎที่ 7 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่บริษัท หมวดหมู่สื่อ/ข่าวสาร/สิ่งพิมพ์, หมวดหมู่โทรทัศน์/ภาพยนตร์ หมวดหมู่อาชีพ >> มักจะมีการกดถูกใจหมวดหมู่บุคคลสาธารณะ ด้วยความเชื่อมั่น 92%

กฎที่ 8 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่สถาบันการศึกษา หมวดหมู่บริษัท หมวดหมู่โทรทัศน์/ภาพยนตร์, หมวดหมู่อาชีพ >> มักจะมีการกดถูกใจหมวดหมู่เว็บไซต์ ด้วยความเชื่อมั่น 92%

กฎที่ 9 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่ธุรกิจ หมวดหมู่บริการ หมวดหมู่บริษัท, หมวดหมู่โทรทัศน์/ภาพยนตร์ หมวดหมู่อาชีพ >> มักจะมีการกดถูกใจหมวดหมู่เว็บไซต์ ด้วยความเชื่อมั่น 92%

กฎที่ 10 หมายถึง ผู้ใช้ที่กดถูกใจหมวดหมู่บริการ หมวดหมู่สินค้าและบริการ หมวดหมู่โทรทัศน์/ภาพยนตร์ >> มักจะมีการกดถูกใจหมวดหมู่ชุมชน ด้วยความเชื่อมั่น 91%

5. สรุปผลการวิจัย

งานวิจัยนี้นำเสนอแนวคิดในหาความสัมพันธ์ของหมวดหมู่ของเพจบนเฟซบุ๊ก โดยศึกษาความสัมพันธ์ของหมวดหมู่จากประวัติการกดถูกใจเพจของผู้ใช้เฟซบุ๊ก ทำการทดลองเพื่อศึกษาประสิทธิภาพความถูกต้องและจำนวนกฎที่สร้างขึ้นเพื่อหาค่าความสนับสนุนขั้นต่ำและค่าความเชื่อมั่นที่เหมาะสมซึ่งจากผลการทดลองที่ได้ในงานวิจัยครั้งนี้ แสดงให้เห็นได้ว่าการกำหนดค่าสนับสนุนขั้นต่ำที่ 10% และค่าความเชื่อมั่นขั้นต่ำที่ 50% ให้ค่าความถูกต้องสูงที่สุด คือ 93.71% โดยจำนวนกฎที่ใช้จริง คือ 892 กฎ คิดเป็น 11.64% ของจำนวนกฎที่สร้างขึ้นทั้งหมด เนื่องจากกฎที่ถูกเอามาใช้จริงเป็นกฎเดิมจึงทำให้จำนวนกฎที่เอามาใช้จริงมีจำนวนน้อยมากเมื่อเทียบกับกฎที่สร้างขึ้นทั้งหมด

6. ข้อเสนอแนะ

1) ผลการวิจัยนี้สามารถไปใช้ในการพัฒนาแอปพลิเคชันสำหรับแนะนำเพจที่ผู้ใช้สนใจได้ โดยเมื่อผู้ใช้กดถูกใจเพจใดเพจหนึ่ง แอปพลิเคชันสามารถแนะนำเพจอื่นที่ผู้ใช้สนใจได้ และนอกจากนี้ยังสามารถใช้เป็นข้อมูลสำหรับผู้ให้บริการที่ต้องการเพิ่มยอดการกดถูกใจ

2) การรวบรวมข้อมูลบนเฟซบุ๊ก ผู้วิจัยจำเป็นต้องใช้ Facebook Graph API ซึ่งมีการปรับเปลี่ยนคำสั่งบ่อย ดังนั้นผู้วิจัยจำเป็นต้องศึกษาคำสั่งเวอร์ชันล่าสุดของ Facebook Graph API ก่อนใช้คำสั่งในการดึงข้อมูล

3) จากผลการวิจัยจะเห็นได้ว่า จำนวนกฎที่สร้างมีจำนวนมาก แต่กฎที่นำไปใช้จริงมีอยู่จำนวนน้อย ดังนั้นการค้นหาวิธีการขุดค้นเฉพาะกฎที่สามารถนำไปใช้จริงจึงเป็นงานที่ท้าทายและควรทำการศึกษาในอนาคต

7. เอกสารอ้างอิง

- [1] P. Pongwittayapanu. *Information about statistics and behavior of using social network in Thailand.* Available Online at https://www.m-culture.go.th/surveillance/ewt_news.php?nid=897&filename=index accessed on 2 June 2014.
- [2] P. Nilagupta. *What's SocialNetwork.* Available Online at <https://sites.google.com/site/socialnetwork057/khwam-hmay-khxng-social-network>. accessed on 26 December 2014
- [3] T. Nithiaurai. *Online Shop on Facebook.* Bangkok: Dream & Passion; 2012.
- [4] Ngoc PT, M. Yoo "The Lexicon-based Sentiment Analysis for Fan Page Ranking in Facebook." *The International Conference on Information Networking (ICOIN)*, Phuket, Thailand. pp. 444 - 448, 10 - 12 February, 2014.
- [5] Kapook. *What's Fanpage.* Available Online at http://fbguide.kapook.com/howto/fan_page.php. accessed on 2 November 2014.



- [6] TechCrunch. *What is Facebook? Facebook profile and how to Facebook*. Available Online at http://www.dmc.tv/pages/top_of_week. accessed on 13 June 2014.
- [7] American Majority. *Facebook a Beginner's Guide*. Purcellville.
- [8] S. Komkong. *Data Mining*. Documentation Report Information and Communication Technology 2013.
- [9] G. Savi. "A Survey on Association Rule Mining in Market Basket Analysis." *International Journal of Information and Computation Technology*, No. 4, 2014.
- [10] N. Khongthim and J. Phuboon-ob "Applied FP-Growth algorithm to guide the higher education study." *National Conference on Computer Information Technologies 2011 (CIT2011) and UniNet Network Operation and Management Workshop 2011 (UniNOMS2011)* pp. 13-17. Nakhonpathom : Mahidol University:Thailand, 2011.
- [11] N. Suwanno and A. Singlam. "The use of Association rules to find the factors effected to the risk of low proficiency students Case study : Prince of Songkla University." *Journal of Management Sciences*, Vol. 28, No. 1, January - June 2011.
- [12] K. Suriyabantoeng and K. Keatruangkamala. *To Sell Insurance Product Modeling for Elderly Customer of Saving Account by Using DataMining*. Bangkok : Thammasat University; 2014.
- [13] K. Waiyamai, C.Songsiri, and T. Raktamanon. "Using Data Mining Techniques to Improve the Quality of Education in the Faculty of Engineering" *NECTEC Technical Journal*, Vol. 13, July-October 2001.
- [14] C. Saimro, V. Jaroenpuntaruk, and W. Pornpatcharapong. *The Development of Predictive Model for Complications of Disease In Diabetic Patients Using Data Mining Techniques : A Case of Pathum Thani Hospital*. pp. 26-27. Nonthaburi : Sukhothai Thammathirat Open University : Thailand, 2014.
- [15] K. Sumamal. *Basic Health Screening by Using Data Mining Techniques*. Master of Science in Web Engineering. Bangkok : Dhurakij Pundit University, 2012.
- [16] PL. Hsua. "The hybrid of association rule algorithms and genetic algorithms for tree induction: an example of predicting the student course performance." *Expert Systems with Applications 2003*, pp. 51-62, 2003.
- [17] Y. K. Choi and S. K. Kim. "An auxiliary recommendation system for repetitively purchasing items in E-commerce." *International Conference on Big Data and Smart Computing (BIGCOMP)*, Seoul, Korea. pp. 96 - 98.
- [18] W. Shunmin. "The Study of Association Rule Mining Technology in Hospital Information System Analysis." *5th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC) 2013*, pp. 245-248, 2013.