

## บทที่ 2

### แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัย ได้ศึกษา ค้นคว้าตำรา และบทความที่เกี่ยวข้องกับการพยากรณ์การซื้อสินค้าผลิตภัณฑ์สมุนไพร โดยวิธีทฤษฎีความสัมพันธ์ ใช้เทคนิค FP-Growth ในการหาทฤษฎีความสัมพันธ์ ตลอดจนวรรณกรรมหรืองานวิจัยที่เกี่ยวข้อง ดังนี้

#### 2.1 แนวคิด ทฤษฎี

- 2.1.1 สินค้า (product)
- 2.1.2 โมเดลพฤติกรรมผู้บริโภค (Consumer Behavior Model)
- 2.1.3 การตอบสนองของผู้ซื้อ
- 2.1.4 ทฤษฎีอุปสงค์
- 2.1.5 ทฤษฎีการบริโภค
- 2.1.6 นิยามเหมืองข้อมูล
- 2.1.7 ขั้นตอนการทำเหมืองข้อมูล
- 2.1.8 เทคนิคการทำเหมืองข้อมูล
- 2.1.9 อัลกอริทึม FP-Growth

#### 2.2 งานวิจัยที่เกี่ยวข้อง

- 2.2.1 การพยากรณ์การซื้อสินค้าประเภทกระเบื้องปูพื้นสำหรับร้านค้าปลีกโดยวิธีทฤษฎีความสัมพันธ์
- 2.2.2 การประยุกต์ใช้ทฤษฎีเหมืองข้อมูลในบริการยืมคืนของห้องสมุด กรณีศึกษา สำนักหอสมุดมหาวิทยาลัยธรรมศาสตร์
- 2.2.3 การพัฒนาระบบสารสนเทศเพื่อการพยากรณ์ผู้เข้าศึกษา โดยผ่านเครือข่ายอินเทอร์เน็ต
- 2.2.4 การพยากรณ์ราคาข้าวเปลือกโดยใช้เทคนิคการทำเหมืองข้อมูล
- 2.2.5 การศึกษาพฤติกรรมการใช้บริการสนามของนักกอล์ฟ โดยใช้อัลกอริทึมของทฤษฎีความสัมพันธ์ กรณีศึกษา สนามกอล์ฟคริสตัล เบย์ กอล์ฟคลับ
- 2.2.6 ระบบเหมืองข้อความจากคลังความรู้สมุนไพรไทย
- 2.2.7 การสร้างโมเดลการจัดการ ระบบนักศึกษาสัมพันธ์โดยใช้เทคนิคเหมืองข้อมูล

2.2.8 การศึกษาการหาพฤติกรรมการยกเลิกการใช้บริการธุรกรรมทางการเงินผ่านโทรศัพท์เคลื่อนที่ จากรายการบันทึกการทำธุรกรรมโดยการทำเหมืองข้อมูล

2.2.9 การใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษาวิทยาลัยราชพฤกษ์

2.2.10 การประยุกต์ใช้เหมืองข้อมูลเพื่อการทำนายสถานภาพนักศึกษา

## 2.1 แนวคิด ทฤษฎี

### 2.1.1 สินค้า

สินค้า คือ วัสดุ อุปกรณ์ ข้าวของเครื่องใช้ที่จับต้องได้ สามารถขนส่งหรือเคลื่อนย้ายไปมาได้ โดยสามารถนำไปเป็นกรรมสิทธิ์ได้ ในอดีตจะเกิดการแลกเปลี่ยนกัน ส่วนปัจจุบันจะเป็นการใช้เงินตราแลกเปลี่ยนกัน

### 2.1.2 โมเดลพฤติกรรมผู้บริโภค (Consumer Behavior Model)

โมเดลพฤติกรรมผู้บริโภค เป็นการศึกษาถึงเหตุจูงใจ ที่ทำให้เกิดการตัดสินใจซื้อผลิตภัณฑ์ โดยมีจุดเริ่มต้นจาก การเกิดสิ่งกระตุ้นที่ทำให้เกิดความต้องการ เปรียบเหมือนกลองดำที่ผู้ผลิตหรือผู้ขายไม่สามารถเดาได้ ความรู้สึกของผู้ซื้อจะได้รับอิทธิพลจากลักษณะต่างๆ ประสบการณ์ต่างๆ ในชีวิตของผู้ซื้อ จะมีผลต่อการตอบสนองหรือการตัดสินใจซื้อสินค้า ดังรายละเอียดต่อไปนี้

1. สิ่งกระตุ้น (stimulus) เป็นเหตุจูงใจที่ทำให้ผู้บริโภคซื้อสินค้า อาจเกิดจากความต้องการภายในร่างกาย และสิ่งกระตุ้นภายนอก ประกอบด้วย

- สิ่งกระตุ้นทางการตลาด เป็นสิ่งที่นักการตลาดสามารถควบคุมและจัดให้มีขึ้น เพื่อกระตุ้นให้ผู้บริโภคเกิดความต้องการ ซื้อผลิตภัณฑ์ ประกอบด้วย ตัวผลิตภัณฑ์ ราคา ช่องทางจัดจำหน่าย และช่องทางการส่งเสริมการตลาด

- สิ่งกระตุ้นอื่นๆ เป็นสิ่งกระตุ้นความต้องการ ของผู้บริโภคที่อยู่นอกองค์กร ที่ผู้ผลิตหรือผู้ขายไม่สามารถควบคุมได้ ได้แก่ สิ่งกระตุ้นทางเทคโนโลยี สิ่งกระตุ้นทางด้านกฎหมาย สิ่งกระตุ้นทางด้านวัฒนธรรม สิ่งกระตุ้นทางเศรษฐกิจ เช่น รายได้ของผู้บริโภค ภาวะเศรษฐกิจ การเก็บภาษีมูลค่าเพิ่ม แนวโน้มการปรับราคาสินค้า เป็นต้น

2. ความรู้สึกนึกคิดของผู้ซื้อ หลังจากได้รับการกระตุ้นทางการตลาด และสิ่งกระตุ้นอื่นๆ แล้ว ความรู้สึกนึกคิดของผู้ซื้อเป็นสิ่งที่ผู้ผลิตหรือผู้ขายไม่สามารถคาดคะเนได้ ดังนั้น จึงมีความพยายามค้นหาถึงความรู้สึกนึกคิดของผู้ซื้อให้พบ เพื่อจะได้จัดส่วนผสมทางการตลาดต่างๆ ให้ตอบสนองความต้องการผู้ซื้อได้ถูกต้อง ความรู้สึกนึกคิดได้รับอิทธิพลจาก

- ปัจจัยด้านวัฒนธรรม วัฒนธรรมเป็นพฤติกรรมและสิ่งที่คนกลุ่มหนึ่งสร้างขึ้นจากการเรียนรู้ในกลุ่มของตน และมีอิทธิพลอย่างมากต่อพฤติกรรมการซื้อขายของผู้ซื้อ เนื่องจากวัฒนธรรมได้แทรกซึมอยู่ในการดำเนินชีวิตประจำวัน

- ปัจจัยด้านสังคม มีอิทธิพลต่อกระบวนการตัดสินใจของผู้ซื้อ ประกอบด้วย ครอบครัว สถานภาพ และบทบาทของผู้บริโภคในสังคม

- ปัจจัยส่วนบุคคล ได้แก่ อายุ วัฏจักรชีวิต ครอบครัว อาชีพ รายได้ การศึกษา บุคลิกภาพ แนวคิดส่วนบุคคล

- ปัจจัยทางจิตวิทยา ได้แก่ การสนใจในการรับรู้ การเรียนรู้ ความเชื่อและทัศนคติ

3. กระบวนการตัดสินใจซื้อของผู้บริโภค จะประกอบด้วยขั้นตอน การรับรู้ปัญหา การหาข้อมูล การประเมินผลทางเลือก การตัดสินใจซื้อ และการประเมินค่าภายหลังการซื้อ

### 2.1.3 การตอบสนองของผู้ซื้อ

การตัดสินใจของผู้ซื้อ แต่ละรายแตกต่างกัน ขึ้นอยู่กับการเลือกผลิตภัณฑ์ การเลือกตราสินค้า การเลือกผู้ขาย การเลือกเวลาการซื้อ เป็นต้น

1. ทฤษฎีส่วนผสมทางการตลาด เป็นตัวแปรทางการตลาดที่สามารถควบคุมได้ 4 ประการ ซึ่งเป็นสิ่งกระตุ้นทางการตลาดที่มีผลต่อการตัดสินใจของผู้บริโภค ประกอบด้วย

- ผลิตภัณฑ์ เป็นสิ่งสนองตอบความจำเป็น และความต้องการของผู้ซื้อ

- ราคา จะต้องเหมาะสมทั้งลักษณะ การแข่งขันภายในตลาด และปฏิกิริยาของลูกค้า

- การจัดจำหน่าย เป็นการนำผลิตภัณฑ์ ไปสู่ตลาดเป้าหมายได้ทันเวลาและในสถานที่ที่ต้องการ จึงต้องมีการพิจารณาถึง เวลา สถานที่ สินค้า และบริการ

- การส่งเสริมการตลาด เป็นการแจ้ง บอกร สร้างความเข้าใจและการรับรู้ที่มีต่อสินค้า แก่ลูกค้าอย่างถูกต้อง การส่งเสริมการตลาดสามารถทำได้โดย

- การโฆษณา เป็นการเสนอขายโดยผ่านสื่อต่างๆ เช่น การโฆษณาทางวิทยุ โทรทัศน์ หนังสือพิมพ์ ป้ายโฆษณา โปสเตอร์และใบปลิว เป็นต้น

- การขายโดยพนักงานขาย เป็นการเสนอขายโดยใช้พนักงานขาย เข้าไปติดต่อลูกค้าโดยตรง

- การส่งเสริมการขาย เป็นกิจกรรมที่กระตุ้นการซื้อของผู้บริโภคและเพิ่มประสิทธิภาพของผู้ขาย

- การประชาสัมพันธ์ เป็นการสร้างความคิดเห็น หรือทัศนคติที่ดีต่อธุรกิจ เช่น การสร้างความสัมพันธ์อันดีกับชุมชน เป็นต้น

2. กระบวนการตัดสินใจซื้อของผู้บริโภค เป็นการตัดสินใจที่ต้องใช้เวลา และมีขั้นตอนที่จะนำมาช่วยในการพิจารณาตัดสินใจ ดังนี้

- การรับรู้ความต้องการ หรือการรับรู้ปัญหา ผู้บริโภครู้ว่าตนเองต้องการผลิตภัณฑ์หรือบริการใด

- การค้นหาข้อมูล เกี่ยวกับทางเลือกซื้อผลิตภัณฑ์ จากแหล่งต่างๆ เช่น แหล่งบุคคล แหล่งการค้า แหล่งสาธารณะ แหล่งประสบการณ์

3. การประเมินค่าทางเลือก ผู้บริโภคจะทำการวัดและเปรียบเทียบข้อมูลที่ได้รวบรวมมาเกณฑ์เหล่านี้คือ คุณสมบัติที่พึงประสงค์และไม่พึงประสงค์ของผลิตภัณฑ์ โดยให้คะแนนแก่เกณฑ์ต่างๆ สำหรับข้อมูลทุกๆ ส่วน แล้วนำมาเรียงลำดับ เปรียบเทียบ เพื่อใช้เป็นข้อมูลในการตัดสินใจ

4. การตัดสินใจซื้อ เป็นขั้นตอนที่ผู้บริโภคจะซื้อผลิตภัณฑ์ โดยอาศัยผลการประเมินทางเลือก

5. การประเมินทางเลือกซื้อ หลังการซื้อมาแล้ว ผู้บริโภคจะประเมินว่าผลิตภัณฑ์ที่ซื้อมานั้นมีคุณสมบัติที่ดี สามารถตอบสนองความต้องการหรือสร้างความพึงพอใจมากน้อยเพียงใด เกณฑ์การวัดความพอใจนั้นผู้บริโภคจะใช้ความคาดหวังที่ตั้งเอาไว้ก่อนการซื้อเป็นมาตรฐาน ถ้าผลิตภัณฑ์ให้ประโยชน์กว่าที่คาดหวัง ผู้บริโภคจะรู้สึกพอใจมาก ถ้าให้ประโยชน์น้อยกว่าที่คาดหวัง ผู้บริโภคจะรู้สึกไม่พอใจ

#### 2.1.4 ทฤษฎีอุปสงค์

อุปสงค์ หมายถึง ความต้องการสินค้าและบริการของผู้บริโภค โดยที่ผู้บริโภคจะต้องมีความสามารถที่จะจ่ายเงินซื้อ และเต็มใจที่จะซื้อสินค้าและบริการนั้น โดยมีปัจจัยที่จะส่งผลให้ผู้บริโภคซื้อสินค้ามากหรือน้อย ดังนี้

1. ราคาสินค้า ปกติสินค้าราคาสูง ปริมาณซื้อสินค้าจะลดลง แต่ถ้าราคาลดลง ปริมาณซื้อก็จะมาก เป็นไปตามกฎของอุปสงค์

2. รสนิยมของผู้บริโภค สิ่งที่กำหนดรสนิยมของผู้บริโภค เช่น อายุ เพศ ความเชื่อ ค่านิยม การศึกษา และอิทธิพลของการโฆษณา ดังนั้น หน่วยงานธุรกิจหลายแห่ง จึงยอมทุ่มเงินจำนวนมหาศาลในการโฆษณา เพื่อหวังผลในการเปลี่ยนแปลงรสนิยม

3. ปริมาณการซื้อขึ้นกับจำนวนประชากร โดยเฉพาะจำนวนประชากรที่มีกำลังซื้อ มักจะทำให้ปริมาณการซื้อเพิ่มมากขึ้น

4. ปริมาณการซื้อขึ้นอยู่กับรายได้เฉลี่ยของครัวเรือน ประชากรที่มีรายได้โดยเฉลี่ยเพิ่มสูงขึ้น พฤติกรรมความต้องการสินค้าและบริการจะเปลี่ยนไป คือมักจะลดการบริโภคสินค้าราคาถูก และในขณะเดียวกันก็หันไปบริโภคสินค้าราคาแพง หรือสินค้าที่มีคุณภาพดีกว่าเดิมเพิ่มขึ้น

5. ปริมาณการซื้อขึ้นอยู่กับสภาพการกระจายรายได้ ในระบบเศรษฐกิจและสังคม เช่น ประเทศที่มีการกระจายรายได้ทัดเทียมกัน การบริโภคสินค้าจะไม่แตกต่างกัน

6. ปริมาณการซื้อขึ้นอยู่กับราคาสินค้าอื่นๆ ที่เกี่ยวข้อง ในกรณีที่สินค้าที่ต้องใช้ร่วมกัน เมื่อผู้บริโภคซื้อสินค้าหรือบริการชนิดหนึ่งเพิ่มขึ้น ก็จะต้องซื้อสินค้าอีกชนิดที่ต้องใช้ร่วมกันเพิ่มขึ้นตามไปด้วย

7. ปริมาณการซื้อขึ้นอยู่กับฤดูกาล เช่น ฤดูหนาว จะมีการบริโภคสินค้าประเภทน้ำมันสูงขึ้น

8. ปริมาณการซื้อขึ้นอยู่กับลักษณะทางกายภาพ เช่น ลักษณะที่อยู่อาศัย ที่ตั้ง สาธารณูปโภคขั้นพื้นฐาน และสิ่งอำนวยความสะดวกอื่นๆ ด้วย หากสินค้าใดมีลักษณะทางกายภาพของสินค้าที่ดีกว่าสินค้าของคู่แข่งอื่นๆ ย่อมทำให้ความต้องการซื้อีมากกว่าสินค้าของผู้ผลิตรายอื่น

### 2.1.5 ทฤษฎีการบริโภค

เคนส์ (Keynes) ได้ศึกษาความสัมพันธ์ระหว่างการบริโภคและรายได้จากข้อเท็จจริงและสรุปเป็นกฎว่าด้วยการบริโภค ดังนี้

1. แม้ไม่มีรายได้เลยการบริโภค ก็จำเป็นต้องมีเพื่อยังชีวิตให้อยู่รอด
2. เมื่อรายได้เพิ่มขึ้นการบริโภคจะเพิ่มขึ้น เมื่อรายได้ลดลงการบริโภคก็จะลดลง
3. อัตราการแลกเปลี่ยนของค่าใช้จ่ายที่ใช้ในการบริโภค จะน้อยกว่าอัตราการเปลี่ยนแปลงของรายได้
4. รายได้ที่จ่ายได้ของบุคคล จะถูกแบ่งออกเป็นค่าใช้จ่ายเพื่อการบริโภคส่วนหนึ่ง และเป็นเงินออมอีกส่วนหนึ่งด้วย
5. ความต้องการถือเงินเพื่อใช้จ่ายยามฉุกเฉิน เช่น เจ็บป่วย เกิดอุบัติเหตุ ถูกออกจากงาน เป็นต้น ความต้องการถือเงินประเภทนี้จะขึ้นอยู่กับรายได้ กล่าวคือ ถ้ามีรายได้มากก็ถือเงินประเภทนี้มากด้วย

### 2.1.6 นิยามเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) เป็นกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ของข้อมูลที่ซ่อนอยู่ในชุดข้อมูลนั้น การทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และการแพทย์รวมทั้งในด้านเศรษฐกิจและสังคม

การทำเหมืองข้อมูล (Data Mining) เป็นวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลอย่างง่าย ๆ มาสู่การจัดเก็บในรูปแบบฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนถึงการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล เช่น กระบวนการหรือการเรียงลำดับของการค้นข้อมูลจำนวนมาก การนำมาใช้โดยหน่วยงานทางธุรกิจและนักวิเคราะห์ทางการเงินหรือการนำมาใช้งานในด้านวิทยาศาสตร์ เพื่อเอาข้อมูลขนาดใหญ่ที่สร้างโดยวิธีการทดลองและการสังเกตการณ์ที่ทันสมัย การสกัดหรือแยกข้อมูลที่เป็นประโยชน์จากข้อมูลขนาดใหญ่หรือฐานข้อมูล หรือการวางแผนทรัพยากรขององค์กร โดยสามารถวิเคราะห์ทางสถิติและตรรกะของข้อมูลขนาดใหญ่ เป็นการมองหารูปแบบที่สามารถช่วยการตัดสินใจได้ ( Visit, 2554:4 )

### 2.1.7 ขั้นตอนการทำเหมืองข้อมูล

การทำเหมืองข้อมูล ประกอบด้วย ขั้นตอนต่อไปนี้

1. Data Cleaning เป็นขั้นตอนสำหรับการคัดข้อมูลที่ไม่เกี่ยวข้องออกไป
2. Data Integration เป็นขั้นตอนการรวมข้อมูลที่มีหลายแหล่งให้เป็นข้อมูลชุดเดียวกัน
3. Data Selection เป็นขั้นตอนการดึงข้อมูลสำหรับการวิเคราะห์จากแหล่งที่บันทึกไว้
4. Data Transformation เป็นขั้นตอนการแปลงข้อมูลให้เหมาะสมสำหรับการใช้งาน
5. Data Mining เป็นขั้นตอนการค้นหารูปแบบที่เป็นประโยชน์จากข้อมูลที่มีอยู่
6. Pattern Evaluation เป็นขั้นตอนการประเมินรูปแบบที่ได้จากการทำเหมืองข้อมูล
7. Knowledge Representation เป็นขั้นตอนการนำเสนอความรู้ที่ค้นพบ โดยใช้เทคนิค

ในการนำเสนอเพื่อให้เข้าใจ (อดุลย์ ยิ้มงาม : 2557 )

### 2.1.8 เทคนิคการทำเหมืองข้อมูล

การทำเหมืองข้อมูล มีเทคนิคต่าง ๆ ที่ใช้ในการแก้ปัญหาอยู่หลายเทคนิค ซึ่งจะไม่มีเทคนิคใดเลยที่สามารถแก้ปัญหาของเหมืองข้อมูลได้ทุกปัญหา ดังนั้นความหลากหลายของเทคนิคเป็นสิ่งที่จำเป็นที่จะนำไปสู่วิธีการแก้ปัญหาที่ดีที่สุด ของการทำเหมืองข้อมูล

1. การวิเคราะห์ทางด้านสถิติ (statistical analysis) เป็นวิธีการที่จะค้นหาความรู้เกี่ยวกับผู้เยี่ยมชมเว็บไซต์ โดยการวิเคราะห์ส่วนของไฟล์ เป็นวิธีหนึ่งที่สามารถบอกชนิดความแตกต่างเกี่ยวกับการวิเคราะห์ทางสถิติบนตัวแปร เช่น หน้าที่ดู เวลาการดูและความยาวของเส้นทางเว็บไซต์ เครื่องมือการวิเคราะห์การจราจร เว็บส่วนมากจะรายงานข้อมูลเกี่ยวกับสถิติ เช่น ความถี่ของหน้าที่เข้าถึงบ่อยๆ เวลาดูเฉลี่ยของหน้าหรือความยาวเฉลี่ยของเส้นทางเว็บไซต์ รายงานนี้อาจจะรวมถึงการวิเคราะห์ข้อผิดพลาดที่จำกัด เช่น การสืบทายการที่ไม่ให้สิทธิเข้าถึง หรือการค้นหายูอาร์แอลที่ไม่สมบูรณ์ รูปแบบของความรู้สามารถเป็นประโยชน์อย่างมากสำหรับการปรับปรุงระบบ การเพิ่มความปลอดภัยของระบบ ทำให้สะดวกในการปรับปรุงเว็บไซต์ และการสนับสนุนการตัดสินใจทางการตลาด

2. กฎความสัมพันธ์หรือความเชื่อมโยง (association rule) คือ การค้นหาความสัมพันธ์ของข้อมูล โดยค้นหาความสัมพันธ์ของข้อมูลทั้งสองชุดหรือมากกว่าสองชุดขึ้นไปไว้ด้วยกัน ความสำคัญของกฎทำการวัดโดยใช้ข้อมูลสองตัวด้วยกันคือ ค่าสนับสนุน (Support) ซึ่งเป็นเปอร์เซ็นต์ของการดำเนินการที่กฎสามารถนำไปใช้ หรือเป็นเปอร์เซ็นต์ของการดำเนินการที่กฎที่ใช้มีความถูกต้อง และข้อมูลตัวที่สองที่นำมาใช้วัดคือ ค่าความมั่นใจ (Confidence) ซึ่งเป็นจำนวนของกรณีที่ถูกถูกต้องโดยสัมพันธ์กับจำนวนของกรณีที่ถูกสามารถนำไปใช้ได้ ในการหากฎความสัมพันธ์นั้นจะมีขั้นตอนวิธีการหาหลายวิธีด้วยกัน แต่ขั้นตอนวิธีที่เป็นที่รู้จักและใช้อย่างแพร่หลายคือ อัลกอริทึม Apriori และ FP-Growth

3. คลัสเตอร์ริง (clustering) คือ การจัดกลุ่มข้อมูลซึ่งมีลักษณะคล้ายกับการแบ่งประเภทแต่จะไม่เหมือนกัน โดยการแบ่งประเภทจะวิเคราะห์ข้อมูลตามต้นแบบ แต่สำหรับการแบ่งกลุ่มเป็นการวิเคราะห์โดยไม่พิจารณาจัดกลุ่มตามประเภทที่มีหรือที่รู้จัก แต่จะใช้ขั้นตอนวิธีการจัดกลุ่มเพื่อค้นหากลุ่ม

ที่สามารถยอมรับได้เพื่อจัดเข้ากลุ่ม กล่าวคือ กลุ่มของวัตถุมีการสร้างขึ้นโดยเปรียบเทียบวัตถุที่มีความเหมือนกันจัดเข้ากลุ่มเดียวกัน

4. การแบ่งประเภทของข้อมูล (classification) คือ การจัดแบ่งประเภทของข้อมูล โดยหาชุดต้นแบบหรือชุดของการทำงานที่อธิบายและแบ่งประเภทข้อมูล วัตถุประสงค์ เพื่อให้สามารถใช้เป็นต้นแบบทำนายประเภทของวัตถุหรือข้อมูลที่ไม่มีการระบุประเภทหรือชนิดของข้อมูล ซึ่งต้นแบบสร้างจากการวิเคราะห์ชุดของข้อมูลฝึกสอน (Training Data) โดยอาจจะเป็นกลุ่มข้อมูลที่มีการระบุประเภทหรือกลุ่มเรียบร้อยแล้ว รูปแบบของต้นแบบแสดงได้หลายแบบ เช่น Classification Rules, Decision Trees หรือ Neural Networks เป็นต้น

5. เทคนิคการสืบค้นรูปแบบลำดับเหตุการณ์ (sequential pattern) เป็นการหาความสัมพันธ์ของข้อมูลระหว่างทรานแซคชัน ซึ่งทำให้มีเวลาเข้ามาเกี่ยวข้อง ด้วยรูปแบบลำดับที่น่าสนใจจะแสดงว่าถ้าเกิดเหตุการณ์นี้แล้ว หรือพบกลุ่มของข้อมูลชุดนี้แล้ว มีแนวโน้มที่จะเกิดเหตุการณ์อะไรหรือพบกลุ่มของข้อมูลใดตามมามากภายหลัง

6. แบบจำลองการพึ่งพา (dependency modeling) เป็นการค้นพบรูปแบบอื่นๆ ที่มีประโยชน์ในการทำเหมืองข้อมูลเว็บ เป้าหมายคือการพัฒนาต้นแบบที่สามารถทำได้ของตัวแทนที่สำคัญของการพึ่งพา ท่ามกลางตัวแปรต่างๆ ในโดเมนเว็บ เช่น อาจจะใช้วิธีการสร้างแบบจำลองสถานที่ที่แตกต่างกัน ให้ผู้เยี่ยมชมได้เห็นในขณะที่เลือกซื้อสินค้าทางออนไลน์ มีหลายเทคนิคการเรียนรู้ที่เป็นไปได้สามารถถูกใช้เป็นแบบของการค้นหาพฤติกรรมของผู้ใช้ เช่น ตัวแบบฮิดเดนมาร์คอฟ (Hidden Markov Model) และแนวคิดข่ายงานความเชื่อเบย์ (Bayesian Belief Network)

### 2.1.9 อัลกอริทึม FP-Growth

อัลกอริทึม FP-Growth ย่อมาจาก Frequent Pattern growth เป็นขั้นตอนวิธีที่พัฒนาขึ้นเพื่อลดจำนวนครั้งของการอ่านข้อมูลจากฐานข้อมูล โดยนำเสนอโครงสร้างข้อมูลที่เรียกว่า FP-Tree (Frequent Pattern Tree) โดยจะมีการอ่านข้อมูลจากฐานข้อมูลเพียง 2 ครั้ง ทำให้การประมวลผลมีความเร็วมากขึ้น โดยรอบแรก จะเริ่มจากการอ่านข้อมูล เพื่อหาค่าความถี่ เรียงลำดับข้อมูลตามความถี่จากมากไปหาน้อย และอ่านข้อมูลในฐานข้อมูลรอบที่สอง โดยนำข้อมูลที่เรียงลำดับตามความถี่มาสร้าง FP-Tree ซึ่งในขั้นตอนนี้จะทำการตัดข้อมูลที่มีความถี่น้อยกว่าที่ต้องการทิ้ง และทำการหา prefix path subtree , condition base FP-Tree และ Frequent itemsets ของแต่ละโหนด

## 2.2 งานวิจัยที่เกี่ยวข้อง

### 2.2.1 การพยากรณ์การซื้อสินค้าประเภทกระเบื้องปูพื้นสำหรับร้านค้าปลีกโดยวิธีภูมิความสัมพันธ์

เป็นงานวิจัยของ ชีระวัฒน์ แสนปัญญา เป็นการประยุกต์ใช้แนวคิดเรื่องการทำเหมืองข้อมูลมาวิเคราะห์พฤติกรรมของผู้บริโภคสินค้าประเภทกระเบื้องปูพื้น เพื่อหาภูมิความสัมพันธ์ของการเลือกซื้อ

สินค้าเพื่อนำไปจำหน่ายสำหรับธุรกิจค้าปลีก ซึ่งประกอบไปด้วย รายการใบสั่งซื้อสินค้า รายการสินค้า ซึ่งผลที่ได้จะมีประโยชน์ต่อการวางแผนการผลิตสินค้าให้เพียงพอกับความต้องการและพยากรณ์พฤติกรรมการเลือกซื้อสินค้า ฐานข้อมูลที่น่ามาวิเคราะห์ถูกจัดเก็บไว้ในโปรแกรมจัดการฐานข้อมูล Microsoft SQL Server 2008 ของปี พ.ศ. 2553 จำนวน 22,507 รายการ และโปรแกรม Weka Version 3.6 โดยใช้ อัลกอริทึม Apriori เพื่อหาความสัมพันธ์ พบกฎทั้งสิ้น 4 กฎ โดยมีความเชื่อมั่น 35, 33, 31, 31 เปอร์เซนต์ตามลำดับ

### 2.2.2 การประยุกต์ใช้ทฤษฎีเหมืองข้อมูลในบริการยืมคืนของห้องสมุด กรณีศึกษา สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์

เป็นงานวิจัยของ ดุลยรัตน์ กรณธ์แสง เป็นการประยุกต์ใช้ทฤษฎีเหมืองข้อมูลในบริการยืมคืนของห้องสมุด กรณีศึกษา สำนักหอสมุด มหาวิทยาลัยธรรมศาสตร์ โดยในปัจจุบันห้องสมุดมีทรัพยากรเป็นจำนวนมาก ทำให้ผู้ยืมหนังสือไม่สามารถใช้ทรัพยากรของห้องสมุดได้อย่างมีประสิทธิภาพ ดังนั้นจึงจำเป็นต้องค้นหาความสัมพันธ์ (Association Rule) และการจัดกลุ่ม (Cluster) ของข้อมูลการยืมคืนหนังสือ ซึ่งเป็นการเพิ่มศักยภาพของห้องสมุดให้สามารถใช้ทรัพยากรที่มีอยู่ได้อย่างมีประสิทธิภาพมากขึ้น

### 2.2.3 การพัฒนาระบบสารสนเทศเพื่อการพยากรณ์ผู้เข้าศึกษา โดยผ่านเครือข่ายอินเทอร์เน็ต

เป็นงานวิจัยของ สุมิตรา นวลมีศรี เป็นการพัฒนาระบบสารสนเทศเพื่อการพยากรณ์ ผู้เข้าศึกษาโดยผ่านเครือข่ายอินเทอร์เน็ต ด้วยการสร้างและทดสอบตัวแบบการพยากรณ์ผู้เข้าศึกษา โดยใช้กฎการจำแนกเทคนิคต้นไม้ตัดสินใจด้วยวิธีการตรวจสอบไขว้ จำนวน 3 ตัวแบบ วิธีการแบ่งข้อมูลแบบสุ่ม ด้วยการแบ่งร้อยละ จำนวน 3 ตัวแบบ และวิธีการแบ่งข้อมูลสร้างตัวแบบและทดสอบออกจากกัน จำนวน 1 ตัวแบบ รวมทั้งสิ้น 7 ตัวแบบการพยากรณ์ ผลจากการศึกษาพบว่า วิธีการแบ่งข้อมูลสร้างตัวแบบ และทดสอบออกจากกันได้ค่าประสิทธิภาพต่าง ๆ สูงกว่าตัวแบบอื่น ๆ โดยมีค่าความถูกต้องเท่ากับ 94% ค่าความแม่นยำเท่ากับ 94.30% ค่าความระลึกเท่ากับ 94% และค่าความถ่วงดุลเท่ากับ 93.70% แสดงว่าตัวแบบที่สร้างด้วยวิธีการแบ่งข้อมูลสร้างตัวแบบและทดสอบออกจากกัน มีความถูกต้องและแม่นยำในการพยากรณ์ผู้เข้าศึกษา และได้ทำการประเมินประสิทธิภาพของระบบด้วยค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน กล่าวได้ว่า ระบบสารสนเทศเพื่อการพยากรณ์ผู้เข้าศึกษา โดยผ่านเครือข่ายอินเทอร์เน็ตมีประสิทธิภาพ อยู่ในเกณฑ์ดีและสามารถนำไปประยุกต์ เพื่อใช้ในการพยากรณ์ผู้เข้าศึกษาได้อย่างมีประสิทธิภาพ

### 2.2.4 การพยากรณ์ราคาข้าวเปลือกโดยใช้เทคนิคการทำเหมืองข้อมูล

เป็นงานวิจัยของ อัมรินทร์ ก้อนแพง เป็นการออกแบบและพัฒนาโมเดลการพยากรณ์ โดยใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) ซึ่งเป็นส่วนหนึ่งของระบบสนับสนุนการตัดสินใจ ช่วยใน



การพัฒนาการพยากรณ์ราคาข้าวเปลือกล่วงหน้า โดยใช้เทคนิคโครงข่ายประสาทเทียม ซึ่งคล้ายกระบวนการตัดสินใจของสมองมนุษย์ โดยขั้นตอนในการวิเคราะห์ข้อมูล ใช้หลักการของกระบวนการหาความรู้แบบ CRISP-DM เป็นมาตรฐานในการดำเนินงานการพัฒนาการพยากรณ์ราคาข้าวเปลือกทั้ง 4 ชนิด

จากการวัดประสิทธิภาพของโมเดลของระบบพยากรณ์กับข้อมูลจริง ค่าความถูกต้องในการพยากรณ์ราคาข้าวเปลือก คือ ข้าวเหนียว เป็น 99.91% ข้าวเปลือกเจ้าหน้าปี เป็น 99.61% ข้าวหอมมะลิ เป็น 89.80% และ ข้าวเจ้าหน้าปีและนาปรังความชื้น 14-15% เป็น 77.01% ทำให้โมเดลที่สร้างขึ้นสำหรับพยากรณ์ราคาข้าวเปลือกนี้ สามารถนำไปใช้ประโยชน์ ต่อการพยากรณ์ราคาข้าวเปลือกแก่เกษตรกรผู้ปลูกข้าวได้

2.2.5 การศึกษาพฤติกรรมการใช้บริการสนามของนักกอล์ฟ โดยใช้อัลกอริทึมของกฎความสัมพันธ์ กรณีศึกษา สนามกอล์ฟคริสตัล เบย์ กอล์ฟคลับ

เป็นงานวิจัยของ ไชยยศ เอื้อวิริยะไชยกุล และสมชาย ปราการเจริญ เป็นการนำเทคนิคการทำเหมืองข้อมูล มาใช้กับข้อมูลการใช้บริการสนามกอล์ฟ สารสนเทศการบริการ ประกอบด้วย การเช่ารถกอล์ฟ (Golf cart) การใช้บริการห้องอาหาร (Club house restaurant) การใช้บริการซุ้มน้ำ (Kiosks) การใช้บริการร้านโปรซอฟ และการใช้บริการสนามซ้อม (Driving range) โดยศึกษาความสัมพันธ์ของนักกอล์ฟจำนวน 140,263 คน ที่มาใช้บริการของสนาม ตั้งแต่ปี 2549 – 2553 ผลของกฎความสัมพันธ์ โดยใช้อัลกอริทึม Apriori พบว่ามีอยู่ 8 กฎ และมีค่า Minimum support 0.2 ค่า Maximum confidence 56% Minimum confidence 30% จากนั้นนำไปวิเคราะห์การถดถอย (Regression analysis) ซึ่งใช้เป็นแนวทาง การพยากรณ์ การจัดสรรงบประมาณรายจ่ายประจำปี การจัดใช้ทรัพยากรต่างๆ ของสนามในปีต่อไป

2.2.6 ระบบเหมืองข้อความจากคลังความรู้สมุนไพรไทย

เป็นงานวิจัยของ จารุวัฒน์ ไพไธล และ ผศ.ดร.รัชฎา คงคะจันทร์ เป็นการนำเทคนิคการทำเหมืองข้อมูลมาประยุกต์กับข้อมูลคลังความรู้สมุนไพรไทย โดยใช้เทคนิคการประมวลผลภาษาธรรมชาติ เช่น การตัดคำระดับประโยค การตัดคำระดับคำ การลดรูปให้อยู่ในลักษณะรากคำ การกำกับคำหรือการพิจารณาความสัมพันธ์ของคำที่อยู่ติดกัน และมีความเกี่ยวข้องกันด้วยวิธีหรือประโยค นิพจน์ระบุนาม การระบุความสัมพันธ์ ผลลัพธ์จะได้ความสัมพันธ์ระหว่างสมุนไพร กับประสิทธิภาพการใช้งานของสมุนไพร นั้นๆ

2.2.7 การสร้างโมเดลการจัดการระบบนักศึกษาสัมพันธ์ โดยใช้เทคนิคเหมืองข้อมูล

เป็นงานวิจัยของ ปราณี มณีรัตน์ เป็นการนำเทคนิคการทำเหมืองข้อมูลมาประยุกต์ในการวิเคราะห์หาปัจจัยที่มีผลต่อความพึงพอใจของนักศึกษา ในการสร้างระบบบริหารความสัมพันธ์นักศึกษา

โดยการแบ่งกลุ่มข้อมูลด้วยอัลกอริทึม K-Mean และเทคนิค JRip หาความสัมพันธ์ระหว่างปัจจัยร่วมกัน โดยจะสอบถามความพึงพอใจของนักศึกษาในด้านต่างๆ คือ ภาพรวมเนื้อหาหลักสูตร คุณสมบัติอาจารย์ และสิ่งอำนวยความสะดวกของแต่ละมหาวิทยาลัย และนำผลการวิเคราะห์มาปรับปรุงการสร้างความพึงพอใจ ให้กับนักศึกษาให้ตรงตามความต้องการของกลุ่มนักศึกษาแต่ละกลุ่มให้มากที่สุด

2.2.8 การศึกษาการหาพฤติกรรมการยกเลิกการใช้บริการธุรกรรมทางการเงินผ่านโทรศัพท์เคลื่อนที่ จากรายการบันทึกการทำธุรกรรมโดยการทำให้เหมือนข้อมูล

เป็นงานวิจัยของ เบญญา จันทพันธ์ เป็นการนำเทคนิคการทำเหมืองข้อมูลมาประยุกต์ในการพัฒนาแบบจำลอง ที่สามารถสนับสนุนกิจกรรมทางการตลาด โดยหาพฤติกรรมของกลุ่มที่มีแนวโน้มยกเลิกการใช้บริการ จากข้อมูลรายการธุรกรรม จากผู้ให้บริการโทรศัพท์เคลื่อนที่แห่งหนึ่ง ผลที่ได้จากการทำให้เหมือนข้อมูลพบว่า ความสัมพันธ์ของความถี่ในการใช้บริการ ค่าบริการและสถานะแพ็คเกจบริการเสริม จากชุดข้อมูลเรียนรู้และข้อมูลทดสอบ มีผลต่อแนวโน้มการยกเลิกการใช้บริการและผลการวัดประสิทธิภาพของตัวแบบ เมื่อดำเนินการโดยใช้กลุ่มประชากรทั้งหมด พบว่า แบบจำลองมีประสิทธิภาพในการวิเคราะห์ข้อมูลถูกต้องแม่นยำ สามารถนำไปประยุกต์ใช้งานเพื่อสนับสนุนกิจกรรมทางการตลาดได้อย่างมีประสิทธิภาพ

2.2.9 การใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษาวิทยาลัยราชพฤกษ์

เป็นงานวิจัยของ เรวดี ศักดิ์คุณธรรม เป็นการนำเทคนิคการทำเหมืองข้อมูลมาประยุกต์ในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผล ทางการเรียนของนักศึกษา จากข้อมูลการลงทะเบียนของนักศึกษา มาทำการสร้างโมเดล 2 โมเดล คือ โมเดลวิเคราะห์หาพฤติกรรมทางการเรียนของนักศึกษาในแต่ละสาขาวิชา ซึ่งโมเดลบอกได้ว่า ผลการเรียนกลุ่มวิชาแกนการเงิน มีผลต่อผลการเรียน ในกลุ่มวิชาเอกบังคับของสาขาวิชาคอมพิวเตอร์ธุรกิจ และสาขาวิชาการจัดการการโรงแรมและท่องเที่ยวมากที่สุด และผลการเรียนกลุ่มวิชาแกนธุรกิจ มีผลต่อผลการเรียนในกลุ่มวิชาเอกบังคับของสาขาวิชาการตลาดและการจัดการมากที่สุด ซึ่งผลการทำนายที่ได้ เป็นเพียงข้อมูลที่ช่วยประกอบการตัดสินใจในการให้คำปรึกษาของอาจารย์ที่ปรึกษาเท่านั้น ในความเป็นจริง ควรจะต้องอาศัยปัจจัยอีกหลายอย่างประกอบการตัดสินใจด้วย เช่น การเอาใจใส่ในการเรียน การมาเรียน การส่งงาน เป็นต้น

2.2.10 การประยุกต์ใช้เหมืองข้อมูลเพื่อการทำนายสถานภาพนักศึกษา

เป็นงานวิจัยของ มานิตา สองสี ไชยยันต์ ปาละมาณและวีรพงศ์ วุฒิศักดิ์ เป็นการนำเทคนิคการทำเหมืองข้อมูล มาประยุกต์ในการการวิเคราะห์เพื่อค้นหารูปแบบและความสัมพันธ์ หรือความรู้ที่ซ่อนอยู่ในข้อมูล เพื่อพัฒนาระบบสารสนเทศ ทำนายสถานภาพของนักศึกษา โดยศึกษารูปแบบหรือคุณลักษณะและปัจจัยที่มีผลต่อการศึกษาของนักศึกษาที่มีแนวโน้มที่จะพ้นสภาพ นักศึกษาที่จะ

ลาออก นักศึกษาที่สำเร็จการศึกษา ภายในระยะเวลาของหลักสูตร และนักศึกษาที่สำเร็จการศึกษาเกินระยะเวลาของหลักสูตร ในการวิเคราะห์ข้อมูล จะใช้ข้อมูลนักศึกษาระหว่างปีการศึกษา 2544-2551 จำนวน 5,229 คน แต่ในจำนวนข้อมูลดังกล่าว มีข้อมูลที่สมบูรณ์สามารถนำมาวิเคราะห์ได้เพียงจำนวน 1,808 คน เลือกใช้เทคนิคการทำเหมืองข้อมูลแบบ Classification โดยเลือกการทำนายข้อมูลด้วยวิธี Decision Tree และใช้อัลกอริทึมชนิด J48 เพื่อให้ได้ตัวแบบที่มีค่าความถูกต้องในการทำนายที่น่าเชื่อถือ ผลการนำระบบสารสนเทศ เพื่อการทำนายสถานภาพนักศึกษาไปให้อาจารย์ผู้สอนและอาจารย์ที่ปรึกษา ใช้ พบว่ามีระดับความพึงพอใจอยู่ที่ 3.35 จาก 5.00 ผลการวิจัยครั้งนี้ชี้ให้เห็นว่า ค่าความถูกต้องของการทำนายข้อมูลค่อนข้างต่ำ ทำให้ทำนายได้ไม่แม่นยำเท่าที่ควร จึงทำให้ระดับความพึงพอใจในการนำระบบสารสนเทศเพื่อการทำนายสถานภาพนักศึกษาอยู่ในระดับไม่สูงนัก