

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการจัดโครงการนักศึกษาเรื่องระบบทำนายการฟื้นสภาพของนักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์มหาวิทยาลัยราชภัฏบุรีรัมย์ ด้วยเทคนิคการทำเหมืองข้อมูลได้มีการศึกษาทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้องต่อการจัดทำโครงการนักศึกษาในหัวข้อดังต่อไปนี้

- 2.1 บริบทของคณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์
 - 2.2 ความรู้เกี่ยวกับปัจจัยการสถานภาพนักศึกษาและการฟื้นสภาพของนักศึกษา
 - 2.3 การทำเหมืองข้อมูล (Data Mining)
 - 2.4 กระบวนการวิเคราะห์ข้อมูลด้วยCRISP-DM
 - 2.5 การนำ Data Mining ไปประยุกต์ใช้กับงานด้านต่าง ๆ
 - 2.6 Web Application
 - 2.7 งานวิจัยที่เกี่ยวข้อง
- โดยมีรายละเอียดแต่ละหัวข้อ ดังนี้

2.1 บริบทของคณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

คณะวิทยาศาสตร์ ได้เริ่มก่อตั้งขึ้นในปี พ.ศ. 2519 ตามพระราชบัญญัติวิทยาลัยครู พ.ศ. 2518 โดยเปลี่ยนจากหมวดวิชาวิทยาศาสตร์มาเป็น คณะวิชาวิทยาศาสตร์และเทคโนโลยี ปีการศึกษา 2521 ได้เปิดสอนระดับปริญญาตรีในหลักสูตรวิทยาศาสตร์ทั่วไป (ค.บ. 2 ปีหลัง)

ปีการศึกษา 2524 เปิดสอนหลักสูตร ค.บ. 4 ปี วิชาเอกวิทยาศาสตร์ทั่วไป

ปีการศึกษา 2527 เปิดสอนเทคนิคการอาชีพระดับ ป.กศ. ชั้นสูง 3สาขา คือ สาขาการอาหาร สาขาการก่อสร้าง และสาขากสิกรรม และได้มีการแก้ไขพระราชบัญญัติวิทยาลัยครู พ.ศ. 2518 เป็นผลให้คณะวิชาวิทยาศาสตร์และเทคโนโลยี สามารถเปิดสอนสาขาวิชาวิทยาศาสตร์ได้ โดยผู้สำเร็จการศึกษาระดับปริญญาตรีจะได้รับคุณวุฒิเป็นวิทยาศาสตรบัณฑิต คณะวิชาวิทยาศาสตร์และเทคโนโลยีจึงสามารถผลิตบัณฑิตทั้งสาขาครุศาสตรบัณฑิต (วิชาเอกด้านวิทยาศาสตร์) และวิทยาศาสตรบัณฑิต โครงสร้างการบริหารงานของคณะประกอบด้วย 9 ภาควิชา คือ 1) ภาควิชาเคมี 2) ภาควิชาคณิตศาสตร์ และสถิติประยุกต์ 3) ภาควิชาชีววิทยา 4) ภาควิชาฟิสิกส์และภาควิชาวิทยาศาสตร์ทั่วไป 5) ภาควิชาอุตสาหกรรมศิลป์ 6) ภาควิชาสุขศึกษา 7) ภาควิชาคหกรรมศาสตร์ 8) ภาควิชาคอมพิวเตอร์ และ 9) ภาควิชาเกษตรศาสตร์

ในปี พ.ศ. 2538 มีการประกาศใช้พระราชบัญญัติสถาบันราชภัฏ พ.ศ. 2535 แทนพระราชบัญญัติวิทยาลัยครู วิทยาลัยครูบุรีรัมย์จึงเปลี่ยนเป็นสถาบันราชภัฏบุรีรัมย์ คณะวิชาวิทยาศาสตร์และเทคโนโลยีเปลี่ยนเป็นคณะวิทยาศาสตร์และเทคโนโลยี และหัวหน้าคณะเป็นคณบดี เปลี่ยนโครงสร้างการบริหารงานวิชาการ จากเดิมที่เป็นภาควิชามาเป็นการบริหารงานวิชาการแบบโปรแกรมวิชา

ในปีการศึกษา 2543 คณะวิทยาศาสตร์และเทคโนโลยีมีโปรแกรมวิชาในสังกัดจำนวน 21 โปรแกรมวิชา ต่อมาในปีการศึกษา 2543 มีโปรแกรมวิชาเพิ่มขึ้นเป็น 25 โปรแกรมวิชา

ในปีการศึกษา 2544 สถาบันราชภัฏบุรีรัมย์กำหนดให้นักศึกษาสาขาครุศาสตร์ที่เริ่มเข้าศึกษาในปี พ.ศ. 2544 ทั้งหมดสังกัดคณะครุศาสตร์ทุกสาขาวิชา

ในปี พ.ศ. 2548 สถาบันราชภัฏบุรีรัมย์ได้เปลี่ยนเป็น มหาวิทยาลัยราชภัฏบุรีรัมย์ คณะวิทยาศาสตร์และเทคโนโลยีเปลี่ยนเป็นคณะวิทยาศาสตร์ มีการจัดการบริหารภายใน (ระหว่างรอการอนุมัติจากกระทรวงศึกษาธิการ) ดังนี้ สาขาวิชาในกลุ่มของอุตสาหกรรมศิลป์ ได้แก่ ก่อสร้าง อิเล็กทรอนิกส์ เทคโนโลยีอุตสาหกรรม แยกออกเป็นโครงการจัดตั้งคณะเทคโนโลยีอุตสาหกรรม ส่วนคณะวิทยาศาสตร์เปลี่ยนแปลงโครงสร้างการบริหารจากโปรแกรมวิชาเป็นภาควิชา ประกอบไปด้วย 1) ภาควิชาวิทยาศาสตร์ 2) ภาควิชาวิทยาศาสตร์ประยุกต์ และ 3) ภาควิชาคณิตศาสตร์และสถิติประยุกต์ ภาคเรียนที่ 2 ปีการศึกษา 2549 คณะวิทยาศาสตร์ได้รับอนุมัติจากกระทรวงศึกษาธิการให้จัดแบ่งส่วนราชการเป็น 3 หน่วยงาน คือ 1) สำนักงานคณบดี 2) ภาควิชาวิทยาศาสตร์ และ 3) ภาควิชาวิทยาศาสตร์ประยุกต์ (รวมภาควิชาวิทยาศาสตร์ประยุกต์เข้ากับภาควิชาคณิตศาสตร์และสถิติประยุกต์) และมีการเพิ่มศูนย์วิทยาศาสตร์และวิทยาศาสตร์ประยุกต์อีก 1 หน่วยงาน ซึ่งเป็นการจัดแบ่งเพิ่มเติมเป็นการภายในตามประกาศของมหาวิทยาลัย (ประวัติและที่มาของคณะวิทยาศาสตร์, 2549 สืบค้นจาก <http://science.bru.ac.th/ประวัติความเป็นมา>)

ปรัชญา/วิสัยทัศน์/พันธกิจ

ปรัชญา “ผลิตบัณฑิตทางวิทยาศาสตร์และเทคโนโลยีมีความรู้คู่คุณธรรมเพื่อพัฒนาท้องถิ่น”

วิสัยทัศน์ (Vision) คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ เป็นองค์กรแห่งการเรียนรู้ที่หลากหลายทางด้านวิทยาศาสตร์และเทคโนโลยีของภาคตะวันออกเฉียงเหนือและแข่งขันได้ในระดับสากล

พันธกิจ (Mission)

- ผลิตบัณฑิตด้านวิทยาศาสตร์และเทคโนโลยีที่มีความรู้ มีคุณธรรมและมีคุณภาพตามมาตรฐานสากล

- ให้บริการ ถ่ายทอดเทคโนโลยีและเป็นแหล่งเรียนรู้ทางวิทยาศาสตร์และเทคโนโลยีแก่ท้องถิ่น

- ทำการวิจัยเพื่อสร้างองค์ความรู้สู่การพัฒนา แก้ปัญหาท้องถิ่นและตอบสนองความต้องการของสังคม

- ทะนุบำรุงศิลปวัฒนธรรม อนุรักษ์ทรัพยากรธรรมชาติและสิ่งแวดล้อม

- สืบสานโครงการอันเนื่องมาจากแนวพระราชดำริ

ประเด็นยุทธศาสตร์

ยุทธศาสตร์ที่ 1 สร้างโอกาสทางการศึกษาและมุ่งผลิตที่มีคุณภาพ เพิ่มทักษะความสามารถทางวิชาการและวิชาชีพให้มีความรู้ คู่คุณธรรม นำชุมชนพัฒนาและเป็นที่ยอมรับในระดับสากล

ยุทธศาสตร์ที่ 2 พัฒนาหลักสูตร ส่งเสริมทักษะและสมรรถนะนักศึกษา คณาจารย์และบุคลากรให้มีคุณภาพทุกมิติที่มุ่งสู่ประชาคมอาเซียน

ยุทธศาสตร์ที่ 3 มุ่งการผลิตงานวิจัยและบูรณาการองค์ความรู้สู่การบริการทางวิชาการที่นำไปใช้ประโยชน์ทั้งการศึกษาและเชิงพาณิชย์ เพื่อการพัฒนาท้องถิ่น และชุมชนอย่างยั่งยืนตามหลักปรัชญาเศรษฐกิจพอเพียง

ยุทธศาสตร์ที่ 4 พัฒนาสมรรถนะของระบบเทคโนโลยีสารสนเทศให้ทันสมัยเป็นที่ยอมรับในระดับสากล

ยุทธศาสตร์ที่ 5 อนุรักษ์มรดกทางวัฒนธรรม สร้างองค์ความรู้ สืบสานภูมิปัญญาท้องถิ่นอีสานใต้เพื่อเผยแพร่สู่สากล

ยุทธศาสตร์ที่ 6 พัฒนาระบบบริหารจัดการอย่างมีคุณภาพตามหลักธรรมาภิบาลสู่การเป็นมหาวิทยาลัยชั้นนำ

ยุทธศาสตร์ที่ 7 บริหารจัดการทรัพยากร และทรัพย์สินอย่างมีประสิทธิภาพและคุ้มค่าเพื่อเพิ่มรายได้

2.2 ความรู้เกี่ยวปัจจัยการสถานภาพนักศึกษาและการพัฒนาของนักศึกษา

โดยที่เป็นการสมควรให้มีข้อบังคับมหาวิทยาลัยราชภัฏบุรีรัมย์ ว่าด้วยการประเมินผลการศึกษาสำหรับนักศึกษาภาคปกติ ภาคพิเศษเพื่อใช้ในมหาวิทยาลัยราชภัฏบุรีรัมย์ ฉะนั้นอาศัยอำนาจตามความในมาตรา 18(2) แห่งพระราชบัญญัติมหาวิทยาลัยราชภัฏ พ.ศ. 2547 สภามหาวิทยาลัยราชภัฏบุรีรัมย์จึงวางข้อบังคับว่าด้วยการประเมินผลการศึกษาของมหาวิทยาลัยราชภัฏบุรีรัมย์ไว้ดังต่อไปนี้ (ระเบียบมหาวิทยาลัยราชภัฏบุรีรัมย์ ว่าด้วยการศึกษาระดับปริญญาตรี พ.ศ. 2550, 2547)

2.2.1 การวัดและประเมินผลการเรียน

2.2.1.1 นักศึกษาจะมีสิทธิ์ในการสอบปลายภาคต้องอยู่ในเกณฑ์ต่อไปนี้

- 1) มีเวลาเรียนในรายวิชานั้น ๆ ไม่น้อยกว่าร้อยละ 80 ของเวลาเรียนทั้งหมด

2) ในกรณีที่มึเวลาเรียนในรายวิชาใดน้่อยกว่าร้อยละ 80 แต่ไม่น้่อยกว่าร้อยละ 60 ให้อยูในดุลยพินิจของอาจารย์ผู้สอนและได้รับการอนุมัติจากผู้อำนวยการสำนักสงเสริมวิชาการและงานทะเบียน

3) ในกรณีที่มึเวลาเรียนในรายวิชาใดน้่อยกว่าร้อยละ 60 จะไม่มีสิทธิ์สอบในวิชานั้น

4) ผูไม่มีสิทธิ์สอบปลายภาคตามข้อ 2 และ 3 จะได้รับการพิจารณาผลการเรียนเป็น "F" ยกเว้นกรณีที่ไม่มีสิทธิ์สอบดังกล่าวเนื่องจากเหตุจำเป็นและผู้อำนวยการสำนักสงเสริมวิชาการและงานทะเบียนตามข้อ 2 เห็นควรให้ออนรายวิชา

2.2.1.2 ให้มึการประเมินผลการเรียนในรายวิชาต่า ง ๆ ตามหลักสูตรเป็น 2 แบบ ดังนี้

1) ระบบค่าระดับคะแนน แบ่งเป็น 8 ระดับ

ระดับคะแนน	ความหมายของผลการเรียน	ค่าระดับคะแนน
A	ดีเยี่ยม	4.00
B+	ดีมาก	3.5
B	ดี	3.0
C+	ดีพอใช้	2.5
C	พอใช้	2.0
D+	อ่อน	0.4
D	อ่อนมาก	0.0
F	ตก	0

ระบบนี้ใช้สำหรับการประเมินรายวิชาที่เรียนตามหลักสูตรค่าระดับคะแนนที่ถือว่าสอบได้ต่ากว่า "D" ถ้านักศึกษาได้ค่าระดับคะแนนในรายวิชาได้ต่ากว่า "D" ต่องลงทะเบียนและเรียนใหม่จนกว่าจะสอบได้ กรณีวิชาเลือกถ้าได้ค่าระดับคะแนนต่ากว่า "D" สามารถเปลี่ยนไปเลือกเรียนรายวิชาอื่นได้ และให้นำหน่วยกิตรายวิชาที่ติด "F" เป็นตัวหารในการคิดค่าคะแนนเฉลยสะสม ส่วนการประเมินรายวิชาเตรียมฝึกประสบการณ์วิชาชีพและรายวิชาการฝึกประสบการณ์วิชาชีพ ถ้าได้ค่าระดับคะแนนต่ากว่า "C" ถือว่าสอบตก นักศึกษาจะต่องลงทะเบียนและเรียนใหม่ ถ้าได้รับการประเมินต่ากว่า "C" เป็น ครั้งที่ 2 ถือว่าหมดสภาพการเป็นนักศึกษา

2) ระบบไม่มีค่าระดับคะแนน กำหนดสัญลักษณ์การประเมินดังนี้

ระดับการประเมิน ผลการศึกษา

S (Satisfactory) พอใจหรือผ่าน

U (Unsatisfactory) ยังไม่พอใจหรือไม่ผ่าน

ระบบนี้ใช้สำหรับการประเมินรายวิชาที่หลักสูตรบังคับให้เรียนเพิ่มตามข้อกำหนดเฉพาะและรายวิชาที่สภามหาวิทยาลัยกำหนดให้เรียนเพิ่มรายวิชาที่ได้ได้ผลประเมิน U นักศึกษาต้องลงทะเบียนและเรียนใหม่จนกว่าจะสอบได้

2.2.1.3 สัญลักษณ์อื่น มีดังนี้

- Au (Audit) ใช้สำหรับการลงทะเบียนเพื่อร่วมเรียนโดยไม่นับหน่วยกิต
- W (Withdraw) ใช้สำหรับการบันทึกหลังจากได้รับอนุมัติให้ออนรายวิชานั้น ก่อนกำหนดสอบปลายภาคไม่น้อยกว่า 1 สัปดาห์ หรือตามที่มหาวิทยาลัยกำหนดหรือได้รับอนุมัติให้ออนวิชาเรียนในกรณีที่นักศึกษาลาพักการศึกษาหรือถูกสั่งให้พักการศึกษา หลังจากลงทะเบียนเรียนในภาคเรียนนั้นแล้ว

- I (Incomplete) ใช้สำหรับบันทึกการประเมินที่ไม่สมบูรณ์ในรายวิชาที่นักศึกษายังทำงาน ไม่เสร็จเมื่อสิ้นภาคเรียนหรือขาดสอบนักศึกษาที่ได้ "I" ต้องดำเนินการขอรับการประเมินเพื่อเปลี่ยนระดับคะแนนให้เสร็จสิ้นในภาคเรียนถัดไป ดังนี้

- กรณีนักศึกษายังทำงานไม่เสร็จ ผู้สอนพิจารณาผลงานที่ค้างอยู่เป็น 0 (ศูนย์) และประเมินผลการเรียนจากคะแนนที่มีอยู่แล้ว

- กรณีนักศึกษาไม่ดำเนินการขอรับการประเมินเพื่อเปลี่ยนระดับคะแนนให้เสร็จสิ้นภายในภาคเรียนถัดไป นายทะเบียนจะเปลี่ยนผลการเรียนเป็น "F"

2.2.1.4 นักศึกษาที่เข้าศึกษาในหลักสูตรปริญญาตรี (หลังอนุปริญญา) จะลงทะเบียนเรียนรายวิชาซ้ำกับรายวิชาที่ศึกษาแล้วในระดับอนุปริญญาไม่ได้ หากลงทะเบียนซ้ำให้เว้นการนับหน่วยกิตเพื่อพิจารณาวิชาเรียนตามหลักสูตรที่กำลังศึกษาอยู่

2.2.1.5 การหาค่าระดับคะแนนเฉลี่ย

ค่าระดับคะแนนเฉลี่ยประจำภาคและค่าระดับคะแนนเฉลี่ยสะสมให้คิดเป็นเลขทศนิยม 2 ตำแหน่งโดยไม่ปัดเศษ สำหรับรายวิชาที่ยังมีผลการเรียน "I" ไม่นำหน่วยกิตมารวมในการหาค่าระดับคะแนนเฉลี่ย กรณีที่สอบตกและต้องเรียนซ้ำให้นำรวมทั้งหน่วยกิตที่สอบตกและเรียนซ้ำเพื่อใช้เป็นตัวหาร

2.2.1.6 ผู้ที่สำเร็จการศึกษาตามหลักสูตรต้องมีคุณสมบัติครบถ้วนทุกข้อ ดังนี้

- มีความประพฤติดี
- สอบได้ในรายวิชาต่างครบตามหลักสูตร รวมทั้งรายวิชาที่สภามหาวิทยาลัยกำหนด
- ได้ค่าระดับคะแนนเฉลี่ยสะสมไม่ต่ำกว่า 2.00

2.2.1.7 ระยะเวลาการศึกษา

- หลักสูตรปริญญาตรี 4 ปี มีเวลาเรียนไม่เกิน 8 ภาคการศึกษาปกติใช้เวลาศึกษาไม่เกิน 10 ปีการศึกษาสำหรับนักศึกษาภาคปกติและมีเวลาเรียนไม่เกิน 17 ภาคการศึกษาปกติใช้เวลาศึกษาไม่เกิน 15 ปีการศึกษาสำหรับนักศึกษาภาคพิเศษ

- หลักสูตรปริญญาตรี 4 ปี มีเวลาเรียนไม่เกิน 6 ภาคการศึกษาปกติใช้เวลาศึกษาไม่เกิน 8 ปีการศึกษาสำหรับนักศึกษาภาคปกติ และมีเวลาเรียนไม่ก่อน 15 ภาคการศึกษาปกติใช้เวลาศึกษาไม่เกิน 12 ปีการศึกษาสำหรับนักศึกษาภาคพิเศษ

- หลักสูตรปริญญาตรี 2 ปี มีเวลาเรียนไม่เกิน 4 ภาคการศึกษาปกติใช้เวลาศึกษาไม่เกิน 4 ปีการศึกษาสำหรับนักศึกษาภาคปกติและมีเวลาเรียนไม่เกิน 8 ภาคการศึกษาปกติใช้เวลาศึกษาไม่เกิน 6 ปีการศึกษาสำหรับนักศึกษาภาคพิเศษ

2.2.1.8 นักศึกษาที่ทุจริตหรือร่วมทุจริตในการสอบรายวิชาใด ให้นักศึกษาผู้นั้นได้รับผลการเรียนมีค่าระดับ "F" ในรายวิชานั้นและมหาวิทยาลัยพิจารณาโทษตามควรแก่กรณี

2.2.1.9 มหาวิทยาลัยจะประเมินผลโดยการคิดคะแนนเฉลี่ยสะสม (Cumulative Grade Point Average = Cumulative G.P.A) เมื่อสิ้นภาคการศึกษาปกติทุกภาค

2.2.1.10 การคิดคะแนนเฉลี่ยสะสม

- คิดจากคะแนนทุกรายวิชาที่นักศึกษาลงทะเบียนเรียนที่มีค่าคะแนน ระบุไว้ในข้อ 2.1 ทั้งรายวิชาที่สอบได้และสอบตก

- การคิดคะแนนเฉลี่ยสะสมให้คิดจากค่าคะแนนทุกรายวิชา โดยเอาผลรวมทั้งหมดของผลคูณระหว่างค่าคะแนนที่ได้กับจำนวนหน่วยกิตของแต่ละรายวิชาและหารด้วยจำนวนหน่วยกิตสะสม ค่าระดับคะแนนเฉลี่ยสะสมให้คิดเป็นเลขทศนิยม 2 ตำแหน่งโดยไม่ปัดเศษ

2.2.2 การพ้นสภาพการเป็นนักศึกษา (ระเบียบมหาวิทยาลัยราชภัฏบุรีรัมย์ ว่าด้วยการศึกษา ระดับปริญญาตรี พ.ศ.2550, 2547) การพ้นสภาพการเป็นนักศึกษานักศึกษาภาคปกติจะพ้นสภาพการเป็นนักศึกษาเมื่ออยู่ในเกณฑ์ ข้อใดข้อหนึ่งดังต่อไปนี้

2.2.2.1 ผลการประเมินได้คะแนนเฉลี่ยสะสมต่ำกว่า 0.60 เมื่อสิ้นภาคการศึกษาปกติที่ 2 นับตั้งแต่เริ่มเข้าเรียน

2.2.2.2 ผลการประเมินได้คะแนนเฉลี่ยสะสมต่ำกว่า 0.80 ในภาคการศึกษาปกติที่ 4 ภาคการศึกษาปกติที่ 6 ภาคการศึกษาปกติที่ 8 ภาคการศึกษาปกติที่ 10 ภาคการศึกษาปกติที่ 12 ภาคการศึกษาปกติที่ 14 ภาคการศึกษาปกติที่ 16 และภาคการศึกษาปกติที่ 18 นับตั้งแต่เริ่มเข้าเรียน

2.2.2.3 นักศึกษาลงทะเบียนเรียนครบตามหลักสูตรกำหนดแต่ยังได้ค่าระดับคะแนนเฉลี่ยสะสมต่ำกว่า 0.80

2.2.2.4 มีสภาพเป็นนักศึกษาครบ 10 ปีการศึกษาในกรณีที่เรียนหลักสูตรปริญญาตรี 5 ปี หรือครบ 4 ปีการศึกษาในกรณีที่เรียนหลักสูตรปริญญาตรี 4 ปี หรือครบ 4 ปี การศึกษาในกรณีที่เรียนหลักสูตรปริญญาตรี 2 ปี

2.2.2.5 นักศึกษาไม่ผ่านรายวิชาการเตรียมลកประสบการณ์วิชาชีพและการฝึกประสบการณ์วิชาชีพเป็นครั้งที่ 2

สำหรับนักศึกษาภาคพิเศษ การพัฒนาการเป็นนักศึกษาเมื่อผลการประเมินได้คะแนนเฉลี่ยสะสมต่ำกว่า 0.80 เมื่อสิ้นภาคการศึกษาที่ 4 กรณีเรียนหลักสูตรปริญญาตรี 2 ปี และเมื่อสิ้นภาคการศึกษาที่ 7 กรณีเรียนหลักสูตรปริญญาตรี 5 ปีและ 5 ปี หรือใช้เวลาศึกษาครบตามที่หลักสูตรกำหนดและขาดคุณสมบัติตาม 6.2 และ 6.3 ในการเป็นผู้สำเร็จการศึกษาตามหลักสูตรหรือนักศึกษาลงทะเบียนเรียนครบตามที่หลักสูตรกำหนดแต่ยังได้ค่าระดับคะแนนเฉลี่ยสะสมต่ำกว่า 0.80 หรือนักศึกษาไม่ผ่านรายวิชาการเตรียมฝึกประสบการณ์วิชาชีพและการออกประสบการณ์วิชาชีพเป็นครั้งที่ 2

เมื่อนักศึกษาเรียนได้จำนวนหน่วยกิตครบตามที่กำหนดไว้ในหลักสูตรแล้วและได้ค่าระดับคะแนนเฉลี่ยสะสมตั้งแต่ 0.80 แต่ไม่ถึง 2.00 ให้เลือกเรียนรายวิชาเพิ่มเติมเพื่อทำค่าระดับคะแนนเฉลี่ยสะสมให้ถึง 2.00

2.3 การทำเหมืองข้อมูล (Data Mining)

เหมืองข้อมูล คือ การวิเคราะห์ข้อมูลที่ออกแบบมาเพื่อสนับสนุนการตัดสินใจ เป็นเครื่องมือที่ช่วยค้นหาและแยกแยะสาระที่เป็นประโยชน์ที่ถูกจัดเก็บในฐานข้อมูลขนาดใหญ่ออกมาเป็นการเพิ่มคุณค่าให้กับฐานข้อมูลที่มีอยู่ ลักษณะของสาระจะต้องมีความชัดเจน มีรูปแบบมีเหตุผลใหม่ ใช้ประโยชน์ได้และเข้าใจได้ (Data Mining) เป็นขั้นตอนหนึ่งในการทำกระบวนการค้นหาความรู้ซึ่งเหมืองข้อมูลยังเป็นกระบวนการของการกลั่นกรองสารสนเทศที่ซ่อนอยู่ในฐานข้อมูลใหญ่ จัดได้ว่าเป็นเทคโนโลยีใหม่ที่เข้ามาช่วยให้การวิเคราะห์ข้อมูลทำได้โดยอัตโนมัติและมีประสิทธิภาพสูง จึงได้รับความสนใจนำไปใช้อย่างแพร่หลายในทุกวงการเพื่อทำนายแนวโน้มและพฤติกรรม โดยอาศัยข้อมูลในอดีตและเพื่อใช้สารสนเทศเหล่านี้ในการสนับสนุนการตัดสินใจทางธุรกิจและการประมวผล ข้อมูลเกี่ยวกับลักษณะการใช้ช่วงระยะเวลาและเชื่อมโยงข้อมูลที่มีผู้เข้าใช้บริการจากแหล่งข้อมูลต่าง ๆ เช่น ผู้ใช้บริการเว็บ เพื่อนำมาพิจารณาปรับปรุงการให้บริการโดยอาจจะเพิ่มหรือลดการให้บริการบางชนิดให้เหมาะสมกับกลุ่มผู้ใช้แต่ละสภาพแวดล้อม ซึ่งอาจจะมีความสนใจแตกต่างกันไป ซึ่งอาจจะจะมีหนึ่งเทคนิคหรือมากกว่ามาประมวผลเพื่อดึงความรู้หรือสิ่งที่ต้องการจากฐานข้อมูลที่ได้ผ่านขั้นตอนการจัดเตรียมข้อมูล (Data Cleaning) เพื่อนำข้อมูลเข้าสู่กระบวนการประมวผลโดยใช้เทคนิคในการทำงานเหมืองข้อมูลต่อไป

การทำเหมืองข้อมูล คือกระบวนการค้นหาสารสนเทศหรือข้อความรู้ที่อยู่ในฐานข้อมูลขนาดใหญ่ที่ซับซ้อนเพื่อนำข้อความรู้ที่ได้ไปใช้ประโยชน์ในการตัดสินใจ สารสนเทศที่ได้สามารถนำมาสร้างการพยากรณ์หรือสร้างตัวแบบสำหรับการจำแนกหน่วยหรือกลุ่มหรือแสดงความสัมพันธ์ระหว่างหน่วยต่าง ๆ หรือให้ข้อสรุปของสาระในฐานข้อมูล การทำเหมืองข้อมูลประกอบขึ้นด้วยการนำกระบวนการทางสถิติและการเรียนรู้ผ่านระบบคอมพิวเตอร์เพื่อสร้างตัวแบบ กฎเกณฑ์รูปแบบการพยากรณ์และข้อความรู้จากฐานข้อมูลขนาดใหญ่ โดยการนำข้อความรู้มาขึ้นตอนการดำเนินงานหลายขั้นตอนซึ่งต้องอาศัยเทคนิคหรือวิธีการต่าง ๆ เช่น วิธีการจัดกลุ่มการค้นหาความสัมพันธ์ การพยากรณ์ เป็นต้น การดำเนินงานมักอยู่ในลักษณะของการสร้างตัวแบบ (modeling) ที่อธิบายความเป็นไปหรือสภาพการณ์หนึ่งที่เกิดขึ้นแล้วหรือที่เราทราบคำตอบ แล้วนำตัวแบบนี้มาใช้อธิบายสถานการณ์ที่ยังไม่เกิดขึ้นหรือไม่ทราบคำตอบ ตัวแบบเหล่านี้อาจเป็นตัวแบบที่เรียบง่ายไปจนถึงตัวแบบที่ยุ่งยากซับซ้อนและอาจใช้การผสมผสานแนวคิดหรือเครื่องมือต่าง ๆ เข้าด้วยกันเพื่อที่จะสามารถสกัดข้อความรู้ที่อยู่ในข้อมูลขนาดใหญ่ได้โดยใช้เทคโนโลยีคลังข้อมูล (Data Warehouse) เข้ามาช่วยในการจัดการข้อมูลเพื่อเพิ่มประสิทธิภาพของการทำเหมืองข้อมูล ดังนั้น ถ้ามีฐานข้อมูลขนาดใหญ่ที่มีข้อมูลคุณภาพดีเทคโนโลยีการทำเหมืองข้อมูลจะช่วยให้การค้นหาหรือแสวงหาโอกาสทางธุรกิจใหม่ โดยการนำข้อความรู้จากข้อมูลจะก่อให้เกิดกระบวนการอัตโนมัติในการค้นพบสารสนเทศหรือข้อความรู้ในฐานข้อมูลขนาดใหญ่ ด้วยการใช่วิธีการ เช่น การพยากรณ์แนวโน้มและพฤติกรรม การบริโภคแบบอัตโนมัติ หรือเกิดกระบวนการอัตโนมัติในการค้นพบรูปแบบที่ไม่เคยรู้จักมาก่อนด้วยการใช้วิธีการค้นหาเข้าไปในรายละเอียดของฐานข้อมูลเพื่อหารูปแบบที่ซ่อนอยู่ในฐานข้อมูลนั้น (สัมมนาหนึ่งทศวรรษไอทีจุฬาฯ, 2545)

2.3.1 ประเภทข้อมูลที่สามารถทำ Data Mining

2.3.1.1 ฐานข้อมูลเชิงสัมพันธ์ (Relation Database) จะมีการจัดเก็บข้อมูลในลักษณะที่เป็นกลุ่มของข้อมูลที่มีความสัมพันธ์กันในฐานข้อมูลหนึ่ง ๆ สามารถที่จะมีตารางตั้งแต่ 1 ตารางเป็นต้นไปและในแต่ละตารางนั้นก็อาจมีได้หลายคอลัมน์ (Column) หลายแถว (Row)

2.3.1.2 คลังข้อมูล (Data Warehouse) เป็นการเก็บรวบรวมข้อมูลจากหลายแหล่งมาเก็บไว้ในรูปแบบเดียวกันและรวบรวมไว้ในที่เดียวกัน

2.3.1.3 ฐานข้อมูลการเปลี่ยนแปลงรายการ (Transactional Database) เกิดจากแฟ้มหลักที่มีการจัดเก็บข้อมูลที่ทำไว้ก่อนรูปแบบการเก็บแบบถาวรและมีข้อมูลเก็บไว้อย่างสมบูรณ์และเมื่อมีการแก้ไขหรือปรับปรุงแฟ้มข้อมูลหลักนี้ ระบบจะใช่วิธีสร้างแฟ้มข้อมูลขึ้นมาใหม่ซึ่งจะเปลี่ยนแปลงข้อมูลเฉพาะบางรายการ เรียกว่า " ฐานข้อมูลการเปลี่ยนแปลงรายการ " (Transaction Database) แล้วเขียนโปรแกรมให้คอมพิวเตอร์อ่านแฟ้มข้อมูลใหม่นี้เข้าไปจัดการแก้ไขปรับปรุงข้อมูลในแฟ้มหลักให้คู่ Transaction Database เปรียบเทียบ

2.3.1.4 ฐานข้อมูลขั้นสูง (Advanced Database) เป็นฐานข้อมูลที่จัดเก็บข้อมูลในรูปแบบอื่น ๆ เช่น ข้อมูลแบบ Object-oriented ข้อมูลที่เป็น text file ข้อมูลมัลติมีเดีย ข้อมูลในรูปแบบของ web งานของเหมืองข้อมูล (Task of data mining)

ในการวิเคราะห์เหมืองข้อมูลจะเลือกใช้เทคนิคการทำงานอย่างไรอย่างหนึ่งไม่ได้เพราะในแต่ละขั้นตอนวิธีก็จะมีเทคนิคของเหมืองข้อมูลที่แตกต่างกันไปขึ้นอยู่กับงานวิจัยและวัตถุประสงค์ของงาน รูปแบบของข้อมูลที่ผู้วิจัยมีอยู่ในฐานข้อมูล ซึ่งจะสามารถจัดรูปแบบของขั้นตอนวิธีของการทำเหมืองข้อมูลทั้งหมด 6 ขั้นตอนวิธีได้ ดังนี้

1) การจัดหมวดหมู่ (Classification) การจัดหมวดหมู่ของเหมืองข้อมูลเพราะการทำความเข้าใจและการติดต่อสื่อสารต่าง ๆ ก็เกี่ยวข้องกับการแบ่งเป็นหมวดหมู่การจัดแยกประเภทและการแบ่งแยกชนิดโดยการจัดหมวดหมู่ประกอบด้วยการสำรวจจุดเด่นของวัตถุที่ปรากฏออกมาและทำการกำหนดจุดเด่นนั้น ๆ เป็นตัวที่ใช้แบ่งหมวดหมู่งานในการแบ่งหมวดหมู่และให้ข้อมูลเรียนรู้รูปแบบข้อมูล (Training Set) วิธีการนี้เป็นกระบวนการหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence) ของตัวอย่างในแต่ละหมวดหมู่ ซึ่งมีภาระหน้าที่ในการสร้างโมเดลของบางชนิดที่ไม่สามารถจะจัดหมวดหมู่ของข้อมูลได้ให้สามารถจัดเป็นหมวดหมู่ได้ ตัวอย่างของการจัดหมวดหมู่ เช่น การจัดหมวดหมู่ของผู้ยื่นขอเครดิต (Credits) เป็นระดับต่ำ ระดับกลางและระดับสูงของความเสี่ยงที่จะได้รับ เป็นต้น

2) การประเมินค่า (Estimation) การประเมินค่าทางธุรกิจอย่างต่อเนื่องจะก่อให้เกิดผลลัพธ์ที่มีประโยชน์กับธุรกิจการป้อนข้อมูลที่เรามีอยู่เข้าไป เพื่อใช้ในการประเมินสิ่งต่าง ๆ ที่จะทำให้เกิดประโยชน์หรือสำหรับตัวแปรที่เราไม่รู้ค่าแน่นอน เช่น รายได้จากการค้าขาย จุดสูงสุดทางธุรกิจ หรือคุณภาพของบัตรเครดิต ในทางปฏิบัติการประเมินค่าจะถูกใช้ในการทำงานการจัดหมวดหมู่ ตัวอย่างของการประเมินค่า เช่น การประเมินรายได้รวมของครอบครัว หรือการประเมินจำนวนบุตรในครอบครัว

3) การทำนายล่วงหน้า (Prediction) การทำนายล่วงหน้าก็เป็นงานที่มีลักษณะคล้ายกับการจัดหมวดหมู่หรือการประเมินค่า จะใช้สถิติการบันทึกของการจัดหมวดหมู่ในการทำนายอนาคตของพฤติกรรมข้อมูลหรือการประเมินค่าที่จะเกิดขึ้นในอนาคต ตัวอย่างของงานการทำนายล่วงหน้า เช่น การทำนายการเปลี่ยนแปลงพฤติกรรมของตลาด การทำนายจำนวนลูกค้าที่จะออกจากธุรกิจของเราใน 6 เดือนข้างหน้าหรือการทำนายอนาคตว่าจะเกิดโรคระบาดอะไรขึ้นและกระทรวงสาธารณสุขจะมีแผนการแก้ไขปัญหาอย่างไร เป็นต้น

4) การจัดกลุ่มโดยอาศัยความใกล้ชิดกันหรือการวิเคราะห์ของตลาด (Affinity Group) งานในการจัดกลุ่มหรือการวิเคราะห์ตลาด คือการตัดสินใจรวมสิ่งที่สามารถไปด้วยกันเข้าไว้ในกลุ่ม

เดียวกันตัวอย่างของการจัดกลุ่มโดยอาศัยความใกล้ชิดกัน หรือการวิเคราะห์ตลาด เช่น การตัดสินใจว่าสิ่งใดบ้างที่จะไปอยู่ด้วยกันอย่างสม่ำเสมอในรถเข็นในซูเปอร์มาร์เกต

5) การรวมตัว (Clustering) การรวมตัว คืองานที่ทำการรวมส่วนต่าง ๆ ในแต่ละส่วนที่ต่างชนิดกันให้อยู่รวมกันเป็นกลุ่มย่อยหรือคลัสเตอร์ (Clusters) โดยในแต่ละคลัสเตอร์อาจจะประกอบด้วยส่วนต่าง ๆ ที่ต่างชนิดกัน ซึ่งความแตกต่างของการรวมตัวจากการจัดหมวดหมู่ คือการรวมตัวจะไม่พึ่งพาอาศัยการกำหนดหมวดหมู่ล่วงหน้า และไม่ใช้ตัวอย่างข้อมูลจะรวมตัวกันบนพื้นฐานของความคล้ายในตัวเอง

6) การบรรยาย (Description) ในบางครั้งวัตถุประสงค์ของเหมืองข้อมูล คือ ต้องการอธิบายความสับสนของฐานข้อมูลในทางที่จะเพิ่มความเข้าใจในส่วนของประชากร ผลิตภัณฑ์หรือขบวนการให้มากขึ้นเทคนิคเหมืองข้อมูลส่วนใหญ่ต้องการข้อมูลเรียนรู้รูปแบบข้อมูล (Training Set) ข้อมูลจำนวนมากที่ประกอบด้วยการวิเคราะห์

2.4 กระบวนการวิเคราะห์ข้อมูลด้วยCRISP-DM

CRISP_DM Model: The New Blueprint for Data Mining ได้กล่าวถึง กระบวนการทำ Data mining ว่ามี 6 ขั้นตอนด้วยกันคือ (ทวีรัตน์ นวลช่วย, 2551)

2.4.1 ขั้นตอนที่ 1 การทำความเข้าใจในเรื่องธุรกิจที่จะทำ

ขั้นตอนแรกจัดได้ว่าเป็นขั้นตอนที่มีความสำคัญมากที่สุด ถือว่าเป็นการเริ่มต้นการทำความเข้าใจต่าง ๆ ให้ถ่องแท้ก่อนจะทำขั้นตอนอื่น ๆ ต่อไป ในขั้นตอนนี้จะเน้นที่เราต้องทราบวัตถุประสงค์ของการทำโปรเจกต์เหมืองข้อมูลขององค์กรของเราเสียก่อน จากนั้นจะเกิดองค์ความรู้ขึ้นมาเอง แล้วจึงแปลงองค์ความรู้เพื่อกำหนดค่านิยามปัญหาในการทำเหมืองข้อมูล ถัดจากนั้นจึงเริ่มวางแผนที่ดำเนินการในขั้นต้น วางแผนขั้นตอนวิเคราะห์ ขั้นตอนออกแบบและหาวิธีการแก้ปัญหา ในช่วงสุดท้าย ในขั้นตอนการทำความเข้าใจธุรกิจนั้น มีขั้นตอนย่อย ๆ อีก ได้แก่ การกำหนดวัตถุประสงค์ทางธุรกิจ แน่นนอนว่าวัตถุประสงค์หลักทางธุรกิจ คือ ผลกำไร อันนี้ทุกคนทราบดีแต่การได้มาซึ่งผลกำไรนั้นมาจากอะไรบ้าง เช่น

- มาจากการมีลูกค้ามาก ๆ และการรักษาลูกค้าเอาไว้ไม่ให้ไปหาคู่แข่งของเราจะทำอย่างไร
- มาจากการเพิ่มยอดขายให้กับสินค้าโดยให้ลูกค้าซื้อเพิ่มขึ้น เช่น ยาสีฟันที่มีขายเป็นแพ็คเกจ การโฆษณาของนมให้กินวันละ 2 กล่อง
- มาจากการซื้อของที่ใช่ต่อเนื่อง หรือคู่กัน เช่น ช๊อไซ่ ร่วมกับน้ำมันพืช

2.4.2 ขั้นตอนที่ 2 การทำความเข้าใจกับข้อมูลที่จัดเก็บ

ในขั้นตอนนี้ เราต้องทำอะไรบ้างกับข้อมูลแน่นอนว่าองค์กรที่จะทำ Data mining นั้นจะต้องมีฐานข้อมูลเป็นของตนเอง ได้แก่ ฐานข้อมูลจากการขายของระบบจุดขายหน้าร้าน (point of sale: POS) และควรมีระบบบัตรสมาชิกจึงจะทำให้ data mining มีประสิทธิภาพสูงสุดในขณะเดียวกัน อาจนำข้อมูลจากการสำรวจพฤติกรรมผู้บริโภคในแหล่งที่ตั้งเดียวกัน ข้อมูลจากงานวิจัยภายนอกองค์กร ข้อมูลเหล่านี้มันจะมีแอตทริบิวต์ไม่ตรงกับฐานข้อมูลของเรา เช่น เพศของเราอาจเก็บ M กับ F แต่จากการสำรวจแบบสอบถามเขาอาจเก็บ ชาย หญิง ดังนั้นสิ่งที่เราจะต้องทำให้ขั้นตอนนี้ ได้แก่

- รวบรวมข้อมูลที่เกี่ยวข้อง
- ข้อมูลต้องถูกต้อง เชื่อถือได้
- ปริมาณของข้อมูลต้องมากพอต่อการนำไปใช้ในการวิเคราะห์

2.4.3 ขั้นตอนที่ 3 การเตรียมข้อมูล

ดังได้กล่าวมาแล้วในขั้นตอนที่ 2 ว่าข้อมูลอาจจะมาหลายแหล่งเอามารวบรวมอยู่ในโครงสร้างเดียวกัน ดังนั้นกระบวนการเตรียมข้อมูลเพื่อให้ถูกต้องเชื่อถือได้จึงต้องใช้เวลาในการทำนานมากและใช้เวลามากที่สุดในกระบวนการทำ data mining ทั้ง 6 ขั้นตอน ดังนั้นบริษัทที่ฉลาด ๆ จึงทำเครื่องมือหรือซอฟต์แวร์สำหรับงานนี้ขึ้นมาโดยเฉพาะ เช่น RapidMiner, Weka, R SPSS เป็นต้น ในขั้นตอนนี้ 3 นี้มีขั้นตอนย่อยออกเป็น 3 ขั้นตอนย่อยอีก นั่นคือ

2.4.3.1 การคัดเลือกข้อมูล ข้อมูลตามแอตทริบิวต์ในฐานข้อมูลในระบบการขายมีการเก็บไว้มากมาย เราจะเลือกเอาข้อมูลที่เป็นต้องใช้กับการวิเคราะห์เท่านั้น ประเด็นนี้ก็ต้องสอดคล้องกับขั้นตอนที่ 1 และ 2 เช่น เราตั้งข้อสมมติฐานว่า เพศ อายุ ส่วนสูง น้ำหนักมีผลต่อการซื้อสินค้าเป้าหมายของเราหรือไม่ เราจำเป็นต้องเลือกแอตทริบิวต์เหล่านี้เข้าไปด้วย รวมทั้งสินค้าที่เขาซื้อไป สรุปสั้นในกระบวนการนี้ ได้แก่ 1) กำหนดเป้าหมายว่าเราจะวิเคราะห์หอะไร 2) เลือกใช้เฉพาะข้อมูลที่เกี่ยวข้องกับสิ่งที่เราจะทำการวิเคราะห์เท่านั้น เพื่อไม่ให้เสียเวลากับการเตรียมข้อมูลและการประมวลผลของเครื่องคอมพิวเตอร์

2.4.3.2 การทำให้ข้อมูลคลีน (clean) คลีน บางตำราแปลคำว่า "คลีน" ได้ว่า กลั่นกรอง เขาคลีนข้อมูลกันอย่างไร แล้วข้อมูลที่ไม่คลีนเป็นข้อมูลแบบไหน เช่น 1) ข้อมูลที่ซ้ำกันคลีนด้วยการลบให้เหลือหนึ่งเดียว 2) ข้อมูลที่ผิดพลาดแก้ไขโดยการทำให้ถูก เช่น วันที่เก็บผิดรูปแบบจาก 31 ธ.ค. 2559 หรือ 31/12/2016 เป็น 2016-12-31 3) ข้อมูลที่หายไปก็เพิ่มเข้าไป เช่น จำนวนบุตร ถ้าไม่มีบุตรบางข้อมูลจะเว้นว่างเอาไว้ แก้โดยการใส่ 0 4) ข้อมูลที่ไม่เข้าพวกแปลกแยกออกไปแก้โดยการใส่ข้อมูลค่าเฉลี่ยหรือค่าปริยาย (default)

2.4.3.3 แปลงข้อมูล เป็นขั้นตอนการแปลงข้อมูลในรูปแบบตารางฐานข้อมูลให้อยู่ในรูป item set เพื่อใช้สำหรับการวิเคราะห์ด้วยวิธีการของ data mining เช่น กฎความสัมพันธ์หรือจำนวนความถี่ เป็นต้น

2.4.4 ขั้นตอนที่ 4 การสร้างโมเดล

เป็นขั้นตอนการวิเคราะห์การวิเคราะห์ข้อมูลด้วยเทคนิค data mining กันเลย ซึ่งมีอยู่ด้วยกันหลายเทคนิค ได้แก่

2.4.4.1 Classification การจำแนก แจกแจงเป็นโมเดลที่ใช้สำหรับนำข้อมูลที่มีอยู่มาทำนายอนาคต เช่น การพยากรณ์การขายสินค้า A ในฤดูฝนที่จะถึงหรือการพยากรณ์อากาศ เป็นต้น

2.4.4.2 Clustering เป็นการแบ่งกลุ่มหรือจัดกลุ่มในพวกเดียวกันคล้ายคลึงกัน เช่น การแบ่งสินค้าตามรูปร่าง แบ่งสินค้าตามราคาหรือแบ่งสินค้าตามชนิดการใช้งาน

2.4.4.3 Association rules เป็นการหาความสัมพันธ์ของข้อมูลที่เกิดร่วมกัน เช่น ค้นหาสินค้าที่มีการซื้อร่วมกันบ่อย ๆ แล้วนำเสนอให้ลูกค้ารายใหม่ที่ใช้ใน amazon.com เป็นต้น

2.4.5 ขั้นตอนที่ 5 การประเมินผล

เป็นขั้นตอนการประเมินผลหรือวัดหาประสิทธิภาพของการใช้โมเดลในขั้นตอนที่ 4 ว่าโมเดลที่เลือกใช้งานใน data mining มีความถูกต้องเชื่อมั่นได้เพียงไร ในขั้นตอนนี้จะใช้โมเดล decision tree หรือ neural network เป็นเครื่องมือในการประเมิน

2.4.6 ขั้นตอนที่ 6 การนำไปใช้งาน

เมื่อได้ผลลัพธ์ที่ถูกต้อง ตามวัตถุประสงค์ทุกประการแล้ว ในขั้นตอนนี้จะเป็นการนำไปใช้งานจริง ตัวอย่างที่เราเห็นกัน ได้แก่ โฆษณาใน google, amazon, lazada, youtube ระบบเล่นอัตโนมัติ เป็นต้น

2.5 การนำ Data Mining ไปประยุกต์ใช้กับงานด้านต่าง ๆ

เทคนิคการทำดาต้าไมนิง (Data Mining) นั้นสามารถนำไปวิเคราะห์ข้อมูลเพื่อช่วยประกอบการตัดสินใจและแก้ไขปัญหาในธุรกิจต่าง ๆ ได้เป็นจำนวนมาก ดังนี้ (นิพนธ์ อยู่จุ, 2555)

2.5.1 ทางด้านการตลาด สามารถนำเทคนิคการทำดาต้าไมนิงไปใช้ในการทำนายผลการตอบสนองกับการเปิดตัวสินค้าใหม่ การทำนายยอดขายเมื่อมีการลดราคาสินค้า และการทำนายกลุ่มลูกค้าที่น่าจะใช้สินค้าของเรา เป็นต้น ซึ่งกระบวนการทำนายเหล่านี้จะต้องอาศัยข้อมูลเกี่ยวกับพฤติกรรมการใช้สินค้าของลูกค้าในอดีต

2.5.2 ทางด้านการเงินการธนาคาร สามารถนำเทคนิคในการทำดาต้าไมนิงไปช่วยในการคาดการณ์ถึงโอกาสในการชำระหนี้ของลูกค้า ค้นหาลูกค้าที่ขาดคุณภาพ เพื่อหลีกเลี่ยงความเสี่ยง

ในการปล่อยกู้หรือค้นหาลูกค้าชั้นดีเพื่อเสนอการปล่อยกู้ นอกจากนี้ยังสามารถใช้ในการทำนายแนวโน้มของพฤติกรรมการใช้บัตรเครดิตของลูกค้าได้โดยการเก็บพฤติกรรมการใช้บัตรเครดิตของลูกค้าในอดีตหรือทำการเก็บลักษณะข้อมูลของลูกค้าในแต่ละกลุ่มทั้งกลุ่มลูกค้าชั้นดีและกลุ่มลูกค้าที่ขาดคุณภาพ เพื่อนำมาเป็นพื้นฐานในการจัดกลุ่มให้กับข้อมูลที่เข้ามาใหม่ ซึ่งกระบวนการต่าง ๆ เหล่านี้ส่วนมาก จะใช้เทคนิค Classification และเทคนิค Association เป็นต้น

2.5.3 ทางด้านสถานีโทรทัศน์หรือวิทยุสามารถนำเทคนิคการทำดาต้าไมนนิ่งไปใช้ในการค้นหารายการที่ดี และเหมาะสมต่อช่วงเวลามากที่สุดเพื่อทำการวางผังรายการในแต่ละเดือน โดยจะต้องอาศัยข้อมูลพฤติกรรมชมรายการของลูกค้ามาวิเคราะห์ว่ากลุ่มลูกค้ากลุ่มใดมักจะเลือกดูรายการในช่วงไหนมากที่สุด

2.5.4 ทางด้านการค้าขาย สามารถนำเทคนิคในการทำดาต้าไมนนิ่งไปช่วยในการทำนายผลกำไรเมื่อทำการลงทุนซื้อสินค้ามาเพื่อขายและสามารถนำมาใช้ในการค้นหาจุดคุ้มทุน เช่น มีการนำรายได้ที่ได้จากการขายในอดีตมาทำนายยอดขายในอนาคตเพื่อทำการวางแผนการผลิตและการขายให้มีความเหมาะสมมากยิ่งขึ้น

2.5.5 ทางด้านโรงพยาบาล สามารถนำเทคนิคดาต้าไมนนิ่งมาใช้ในการทำนายอาการของโรคได้อย่างแม่นยำยิ่งขึ้น เพื่อกำหนดวิธีการรักษาได้อย่างเหมาะสมที่สุด เช่น นำข้อมูลอาการของผู้ป่วยในอดีตมาหาความสัมพันธ์โดยใช้เทคนิค Association เพื่อหาว่าส่วนมากเมื่อผู้ป่วยมีอาการอย่างใดอย่างหนึ่งแล้วมักจะเกิดอาการใดตามมา ส่งผลให้มีการวางแผนในการรักษาโรสดังกล่าวได้ดียิ่งขึ้น

2.6 เว็บแอปพลิเคชัน (Web Application)

เว็บแอปพลิเคชัน (Web Application) คือ เว็บแอปพลิเคชัน (Web Application) ที่ถูกเขียนขึ้นมาเพื่อเป็นเบราว์เซอร์ (Browser) สำหรับการใช้งานเว็บเพจ (Webpage) ต่าง ๆ ซึ่งถูกปรับแต่งให้แสดงผลแต่ส่วนที่จำเป็น เพื่อเป็นการลดทรัพยากรในการประมวลผลของตัวเครื่องสมาร์ทโฟนหรือแท็บเล็ต ทำให้โหลดหน้าเว็บไซต์ได้เร็วขึ้น อีกทั้งผู้ใช้งานยังสามารถใช้งานผ่าน Internet (อินเทอร์เน็ต) และ Intranet (อินทราเน็ต) ในความเร็วต่ำได้

ข้อดีของ เว็บแอปพลิเคชัน (Web Application) คือ ในส่วนของการใช้งานที่สามารถใช้งานได้ง่าย สะดวกทุกที่ทุกเวลา ถ้าหากไม่มีเครื่องคอมพิวเตอร์แต่ต้องการใช้ เว็บแอปพลิเคชัน (Web Application) ก็สามารถใช้แอปพลิเคชันประเภทนี้ได้รวมถึงมีการอัปเดต แก้ไขข้อผิดพลาดต่าง ๆ อยู่ตลอดเวลาและใช้งานได้ทุกแพลตฟอร์ม

ตารางที่ 2.1 การพัฒนา Web Application

เทคโนโลยี	ภาษาที่ใช้พัฒนา
CGI (Common Gateway Interface)	C, C++, Shell Script, Perl, Python, Tcl
ASP (Active Server Page)	VBScript, JScript
PHP (Personal Home Page > Professional Home Page > PHP Hypertext Proprocessor)	PHP Script
JSP (Java Server Page)	JavaScript, Java Applet, Java Application
PSP (Python Server Page)	Python Script
ASP.NET (Active Server Page.NET)	VB.NET, C#, J#

HTML (Hyper Text Markup Language) เป็นภาษามาตรฐานสากลที่ใช้นำเสนอข้อมูลแบบผสมผสานในการสื่อสารแบบ World-Wide-Web :WWW (Web) ซึ่งเป็นการเชื่อมต่อเครือข่ายของเครื่องคอมพิวเตอร์ทั่วโลก (Internet) รูปแบบหนึ่ง ข้อมูลในรูปแบบต่าง ๆ ไม่ว่าจะเป็นข้อความ รูปภาพ เสียง ภาพเคลื่อนไหวหรืออื่น ๆ จะถูกเชื่อมโยงเข้าหากัน ด้วยชุดคำสั่งต่าง ๆ เพื่อให้แสดงผลออกมาคล้ายกับสิ่งพิมพ์ สไลด์ หรือ แบบมัลติมีเดีย

HTML มีโครงสร้างการเขียนโดยอาศัยตัวกำกับ (Tag) ควบคุมการแสดงผลข้อความ รูปภาพ หรือวัตถุอื่น ๆ ผ่านโปรแกรมเบราว์เซอร์ แต่ละ Tag อาจจะมีส่วนขยายที่เรียกว่า Attribute สำหรับระบุหรือควบคุมการแสดงผล ของเว็บได้ด้วย HTML เป็นภาษาที่ถูกพัฒนาโดย World Wide Web Consortium (W3C) จากแม่แบบของภาษา SGML (Standard Generalized Markup Language) โดยตัดความสามารถบางส่วนออกไป เพื่อให้สามารถทำความเข้าใจและเรียนรู้ได้ง่าย และด้วยประเด็นดังกล่าว ทำให้บริการ WWW เติบโตขยายตัวอย่างกว้างขวางตามไปด้วย Tag

โครงสร้างของภาษา HTML ไฟล์เอกสาร HTML ประกอบด้วยส่วนประกอบสองส่วนคือ Head กับ Body โดยสามารถเปรียบเทียบได้ง่าย ๆ ก็คือ ส่วน Head จะคล้ายกับส่วนที่เป็น Header ของหน้าเอกสารทั่วไป หรือบรรทัด Title ของหน้าต่างการทำงานในระบบ Windows สำหรับส่วน Body จะเป็นส่วนเนื้อหาของเอกสารนั้น ๆ โดยทั้งสองส่วนจะอยู่ภายใน Tag <HTML>...</HTML>

```

<html>
  <head>
    <tag คำสั่ง>
  </head>
  <body>
    <tag คำสั่ง>
  </body>
</html>

```

ภาพที่ 2.1 Tag <HTML>...</HTML>

2.6.1 CSS (Cascading Style Sheet)

CSS ย่อมาจาก Cascading Style Sheet มักเรียกโดยย่อว่า "สไตลชีต" คือภาษาที่ใช้เป็นส่วนของการจัดรูปแบบการแสดงผลเอกสาร HTML โดยที่ CSS กำหนดกฎเกณฑ์ในการระบุรูปแบบ (หรือ Style) ของเนื้อหาในเอกสาร อันได้แก่ สีของข้อความ สีพื้นหลัง ประเภทตัวอักษร และการจัดวางข้อความ ซึ่งการกำหนดรูปแบบ หรือ Style นี้ใช้หลักการของการแยกเนื้อหาเอกสาร HTML ออกจากคำสั่งที่ใช้ในการจัดรูปแบบการแสดงผล กำหนดให้รูปแบบของการแสดงผลเอกสารไม่ขึ้นอยู่กับเนื้อหาของเอกสาร เพื่อให้ง่ายต่อการจัดรูปแบบการแสดงผลล์พ์ของเอกสาร HTML โดยเฉพาะในกรณีที่มีการเปลี่ยนแปลงเนื้อหาเอกสารบ่อยครั้ง หรือต้องการควบคุมให้รูปแบบการแสดงผลเอกสาร HTML มีลักษณะของความสม่ำเสมอทั่วกันทุกหน้าเอกสารภายในเว็บไซต์เดียวกัน โดยกฎเกณฑ์ในการกำหนดรูปแบบ (Style) เอกสาร HTML ถูกเพิ่มเข้ามาครั้งแรกใน HTML 4.0 เมื่อปี พ.ศ. 2539 ในรูปแบบของ CSS level 1 Recommendations ที่กำหนดโดยองค์กร World Wide Web Consortium หรือ W3C

2.6.2 ภาษา PHP

พีเอชพี (PHP) คือ ภาษาคอมพิวเตอร์ในลักษณะเซิร์ฟเวอร์-ไซด์ สคริปต์ โดยลิขสิทธิ์อยู่ในลักษณะโอเพนซอร์ส ภาษาพีเอชพีใช้สำหรับจัดทำเว็บไซต์ และแสดงผลออกมาในรูปแบบ HTML โดยมีรากฐานโครงสร้างคำสั่งมาจากภาษาซี ภาษาจาวา และภาษาเพิร์ล ซึ่งภาษาพีเอชพี นั้นง่ายต่อการเรียนรู้ ซึ่งเป้าหมายหลักของภาษานี้ คือ ให้นักพัฒนาเว็บไซต์สามารถเขียนเว็บเพจที่มีความตอบโต้ได้อย่างรวดเร็ว

ในช่วงแรกภาษาที่นิยมใช้ในการทำงานบนระบบ Network คือ HTML (Hypertext Markup Language) แต่ภาษา HTML เป็น Static Language ก็คือ ภาษาที่ใช้สร้างข้อมูลประเภทตัวอักษร ภาพ หรือ Object อื่น ๆ ที่ไม่สามารถเปลี่ยนแปลงได้ ด้วยตัวของมันเองหรือเรียกง่าย ๆ

ว่าข้อมูลที่คงที่นั่นเอง ก็ทำให้ไม่ยืดหยุ่นต่อความต้องการของมนุษย์ไม่มีวันเพียงพอจึงได้มีการพัฒนาภาษาที่เป็น Dynamic Language ขึ้นมา ก็คือ ภาษาที่มีข้อมูลถูกเปลี่ยนแปลง Auto ตามเงื่อนไขต่าง ๆ ที่ผู้เขียนกำหนดไว้ มีการประกาศตัวแปรได้ ข้อมูลที่เปลี่ยนแปลงก็เปลี่ยนแปลงตามตัวแปรนี้เป็นที่มาของภาษา PHP, CGI, ASP เป็นต้น โดยเฉพาะภาษาประเภท Scripts ที่สามารถติดต่อกับผู้ใช้ได้ และหนึ่งในภาษาเหล่านั้นก็คือ PHP ซึ่งเป็นภาษาที่ได้รับความนิยมอย่างมากในปัจจุบัน ถูกสร้างขึ้นเมื่อ ปี 1994 ก่อน Windows 95 ปีเดียวเอง โดยนาย Rasmus Lerdorf

PHP เป็นภาษาจำพวก scripting language คำสั่งต่าง ๆ จะเก็บอยู่ในไฟล์ที่เรียกว่า สคริปต์ (script) และเวลาใช้งานต้องอาศัยตัวแปลชุดคำสั่ง ตัวอย่างของภาษาสคริปก็เช่น JavaScript, Perl เป็นต้น ลักษณะของ PHP ที่แตกต่างจากภาษาสคริปต์แบบอื่น ๆ คือ PHP ได้รับการพัฒนาและออกแบบมาเพื่อใช้งานในการสร้างเอกสารแบบ HTML โดยสามารถสอดแทรกหรือแก้ไขเนื้อหาได้โดยอัตโนมัติ ดังนั้นจึงกล่าวว่า PHP เป็นภาษาที่เรียกว่า server-side หรือ HTML-embedded scripting language เป็นเครื่องมือที่สำคัญชนิดหนึ่ง ที่ช่วยให้เราสามารถสร้างเอกสารแบบ Dynamic HTML ได้อย่างมีประสิทธิภาพและมีลูกเล่นมากขึ้น (บัญชา ปะสีละเตสัง, 2558)

2.6.3 MySQL

MySQL คือ โปรแกรมระบบจัดการฐานข้อมูลที่พัฒนาโดยบริษัท MySQLAB มีหน้าที่เก็บข้อมูลอย่างเป็นระบบรองรับคำสั่ง SQL เป็นเครื่องมือสำหรับเก็บข้อมูลที่ต้องใช้ร่วมกับเครื่องมือหรือโปรแกรมอื่นอย่างบูรณาการ เพื่อให้ได้ระบบงานที่รองรับความต้องการของผู้ใช้ เช่นทำงานร่วมกับเครื่องบริการเว็บ (Web Server) เพื่อให้บริการแก่ภาษาสคริปต์ที่ทำงานฝั่งเครื่องบริการ (Server-Side Script) เช่น ภาษา php ภาษา asp.net หรือภาษาเจเอสพี เป็นต้น หรือทำงานร่วมกับโปรแกรมประยุกต์ (Application Program) เช่น ภาษาวิซวลเบสิกดอทเน็ต ภาษาจาวา หรือภาษาซีชาร์ป เป็นต้น โปรแกรมถูกออกแบบให้สามารถทำงานได้บนระบบปฏิบัติการที่หลากหลายและเป็นระบบฐานข้อมูลโอเพนซอร์ซ (Open Source) ที่ถูกนำไปใช้งานมากที่สุด (บัญชา ปะสีละเตสัง, 2558)

2.7 งานวิจัยที่เกี่ยวข้อง

ขวัญฤทัย นกแก้ว (2555) วิจัยเรื่อง การวิเคราะห์ปัจจัยที่ส่งผลต่อการฟื้นสภาพของนักศึกษา ระดับปริญญาตรี โดยใช้เทคนิคเหมืองข้อมูล งานวิจัยนี้เป็นการนำเสนอการวิเคราะห์ปัจจัยที่ส่งผลต่อการฟื้นสภาพของนักศึกษาระดับปริญญาตรี โดยใช้การค้นหากฎความสัมพันธ์ (Association Rule Discovery) ซึ่งเป็นเทคนิคหนึ่งของเหมืองข้อมูล (Data Mining) เทคนิคการจำแนกประเภทข้อมูล (Data Classification) เป็นเทคนิคหนึ่งในการจำแนกประเภทข้อมูล ให้ความแม่นยำในการทำนายสูง เนื่องจากมีการนำเทคนิคการหากฎความสัมพันธ์เข้ามารวมกับเทคนิคการจำแนกประเภทข้อมูลเพื่อ

เพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูล ในงานวิจัยนี้ใช้ชุดข้อมูลของนักศึกษามหาวิทยาลัยราชภัฏยะลา ที่เข้าศึกษาระหว่างปี พ.ศ. 2550-2553 ซึ่งพันสภาพการเป็นนักศึกษาจำนวน 1804 คน มีจำนวน 14 ปัจจัย ผู้วิจัยได้ทำการวิเคราะห์ปัจจัยที่ส่งผลต่อการพันสภาพของนักศึกษา ผลจากการวิเคราะห์พบว่าปัจจัยสำคัญ 2 ปัจจัยคือ สาขาที่นักศึกษาเลือกเรียนในระดับปริญญาตรี และวุฒิการศึกษาเดิมในระดับมัธยมปลาย

จิระนันต์ เจริญรัตน์, วีระศักดิ์ เจริญรัตน์ (2559) วิจัยเรื่อง การวิเคราะห์ปัจจัยที่ส่งผลต่อการพันสภาพของนักศึกษาที่มีผลการเรียนปกติโดยใช้ต้นไม้ตัดสินใจ งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อการพันสภาพของนักศึกษาระดับปริญญาตรี ที่มีผลการเรียนของเกรดเฉลี่ยสะสมมากกว่า 2.00 ขึ้นไป ในการวิเคราะห์ข้อมูลจะใช้ข้อมูลนักศึกษาระดับปริญญาตรี คณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏสกลนคร ที่เข้าศึกษาระหว่างปี พ.ศ. 2553-2557 มีจำนวน 3,385 ชุดข้อมูล เลือกใช้เทคนิคการทำเหมืองข้อมูลแบบ Classification เลือกการทำนายข้อมูลด้วยวิธี Decision Tree และใช้อัลกอริธึมชนิด J48 การทดสอบโมเดลที่ได้จะทำการทดสอบแบบ 10-fold cross validation โดยใช้โปรแกรม WEKA ผลการวิเคราะห์พบว่าปัจจัยที่มีความสำคัญที่จะส่งผลต่อการพันสภาพของนักศึกษาที่มีผลการเรียนปกติ แบ่งเป็น 4 กลุ่ม ตามกลุ่มเกรดเฉลี่ยสะสม ดังนี้ กลุ่ม Best (GPA>3.50 ขึ้นไป) ปัจจัยคือ วุฒิการศึกษาเดิม กลุ่ม Excellent (GPA=3.00-3.50) ปัจจัยคืออาชีพมารดา และสาขาวิชาที่เรียน กลุ่ม Good (GPA=2.50-2.99) ปัจจัยคือ ทุนกู้ยืมเพื่อการศึกษา สถานภาพของครอบครัว รายได้ของบิดารายได้ของมารดาและจังหวัด กลุ่ม Medium (GPA=2.00-2.49) ปัจจัยคือ ทุนกู้ยืมเพื่อการศึกษา สถานภาพครอบครัวและรายได้ของมารดา จากผลการวิจัยนี้สามารถนำกฎการจำแนกข้อมูลที่ได้มาใช้ในการพัฒนาระบบพยากรณ์การพันสภาพของนักศึกษาต่อไปและผู้บริหารสามารถใช้เป็นแนวทางในการบริหารจัดการหรืออาจารย์ที่ปรึกษาสามารถวางแผนการเรียน ดูแลการลงทะเบียนเรียนของนักศึกษาให้ความช่วยเหลือและส่งเสริมนักศึกษาได้อย่างเหมาะสม

ฐานิตา ลอยวิรัตน์ (2558) วิจัยเรื่อง การศึกษาสาเหตุและปัจจัยที่ส่งผลต่อการพันสภาพการเป็นนักศึกษาระดับปริญญาตรี มหาวิทยาลัยสงขลานครินทร์ งานวิจัยนี้มีจุดมุ่งหมายเพื่อการศึกษาและเปรียบเทียบข้อมูลนักศึกษาระดับปริญญาตรี ที่ยื่นคำร้องขอลาออกจากการเป็นนักศึกษา ตั้งแต่ปีการศึกษา 2554-2556 โดยรวบรวมข้อมูลของนักศึกษาที่ยื่นคำร้องขอลาออกจากการเป็นนักศึกษา จำนวน 14 คณะ ของมหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่ คือ คณะการแพทย์แผนไทย คณะทรัพยากรธรรมชาติ คณะทันตแพทยศาสตร์ คณะเทคนิคการแพทย์ คณะนิติศาสตร์ คณะวิทยาศาสตร์ คณะพยาบาลศาสตร์ คณะแพทยศาสตร์ คณะเภสัชศาสตร์ คณะวิทยาการจัดการ คณะวิศวกรรมศาสตร์ คณะศิลปศาสตร์ คณะเศรษฐศาสตร์ และคณะอุตสาหกรรมเกษตร มีนักศึกษามายื่นคำร้องขทั้งหมด จำนวน 600 คน ผลการศึกษาพบว่า

นักศึกษาคณะวิทยาศาสตร์ได้ยื่นคำร้องขอลาออกจากการเป็นนักศึกษามากที่สุด (ร้อยละ 27.67) รองลงมา คือ นักศึกษาคณะวิศวกรรมศาสตร์ (ร้อยละ 19.50) และจากนักศึกษาทั้งหมด พบว่า เป็นนักศึกษาชั้นปีที่ 1 มากที่สุด (ร้อยละ 73.00) รองลงมา คือ นักศึกษาชั้นปีที่ 2 (ร้อยละ 21.67) สำหรับเหตุผลของนักศึกษาที่ยื่นคำร้องขอลาออกจากการเป็นนักศึกษา คือ ต้องการศึกษาต่อสถาบันอื่นมากที่สุด (ร้อยละ 57.12) รองลงมา คือ ไม่ถนัดสาขาวิชาที่เรียน (ร้อยละ 28.79)

พฤตพิพงษ์ เพ็งศิริ และคณะ (2556) วิจัยเรื่อง การวิเคราะห์ความสัมพันธ์ของปัจจัยข้อมูล นักศึกษาที่มีผลต่อระดับผลการเรียนด้วยเทคนิคต้นไม้ตัดสินใจ กรณีศึกษานักศึกษามหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ บทความนี้เป็นนำเสนอการวิเคราะห์กระบวนการตัดสินใจด้วยเทคนิคต้นไม้ตัดสินใจเป็นเทคนิคอย่างหนึ่งในการทำเหมืองข้อมูล ซึ่งอาศัยความสัมพันธ์ของปัจจัยข้อมูล นักศึกษาเป็นการบ่งชี้ถึงระดับผลการเรียนของนักศึกษามหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ซึ่งผลลัพธ์ที่ได้นี้จะส่งผลทำให้เอื้อประโยชน์กับแนวทางในการปรับปรุงบริหารจัดการและส่งเสริมการวางแผนของหลักสูตรการศึกษาที่เกี่ยวข้อง โดยนำปัจจัยข้อมูลของนักศึกษา ได้แก่ ข้อมูลทั่วไป ข้อมูลส่วนตัว ข้อมูลการศึกษา ข้อมูลระดับผลการเรียนของนักศึกษากลุ่มที่ศึกษาในชั้นปีสุดท้ายที่สำเร็จการศึกษา มาหาความสัมพันธ์กับผลการเรียนที่ได้โดยวิเคราะห์จากระดับผลการเรียน (เกรดเฉลี่ย) ของการเรียนในชั้นปีที่ 1 ผลการทดลองพบว่าปัจจัยข้อมูลของนักศึกษาที่เหมาะสมในการเรียนมีทั้งหมด 7 ตัวแปร จากข้อมูลนำเข้าทั้งหมด 12 ตัวแปร ทั้งนี้ 7 ตัวแปรมาจากต้นไม้ตัดสินใจที่ได้มาสามารถสรุปเลือกเฉพาะกิ่งที่มีผลมากที่สุด โดยวัดค่าความแม่นยำ (Accuracy) ได้ค่าสูงถึง 84.78% ซึ่งถือว่าอยู่ในเกณฑ์สูง ดังนั้น ตัวแปรที่น่าจะเป็นปัจจัยที่ส่งผลกับระดับผลการเรียนที่จบการศึกษาสูงสุด คือ ความสม่ำเสมอการเข้าเรียนในการเรียน

เยาวภา ภารสำเร็จ และคณะ (2555) วิจัยเรื่อง การเปรียบเทียบอัลกอริทึมเหมืองข้อมูลเพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อระดับผลการเรียนของนักศึกษา งานวิจัยนี้เป็นการนำเสนอการเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูลของอัลกอริทึมเหมืองข้อมูล 3 แบบ คือ C4.5, Naïve Bayes และ k-Nearest Neighbor อัลกอริทึมที่ให้ค่าประสิทธิภาพสูงที่สุดจะถูกนำมาใช้ในการค้นหาปัจจัยที่ส่งผลต่อระดับผลการเรียนของนักศึกษาโดยการลดการนำเข้าที่ละตัวแปร ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูล นักศึกษาจากมหาวิทยาลัยเทคโนโลยีราชมงคลอีสานวิทยาเขตสกลนคร ระหว่างปีการศึกษา 2553-2555 จำนวน 4,591 ชุดข้อมูล 17 แอทริบิวต์ ผลการเปรียบเทียบพบว่า แบบ C4.5 ให้ค่าประสิทธิภาพสูงที่สุด ร้อยละ 73.55 ซึ่งมากกว่าแบบ k-Nearest Neighbor และแบบ Naïve Bayes ซึ่งมีค่าร้อยละ 66.63 และร้อยละ 49 ตามลำดับ ผลการค้นหาปัจจัยที่ส่งผลต่อระดับผลการเรียนของนักศึกษาพบตัวแปรที่มีความสำคัญเรียงจากสำคัญมากไปหาน้อย คือ ชั้นปี จำนวน พี่น้องที่กำลังศึกษา อายุ จำนวนพี่น้องทั้งหมดและสาขาวิชา นอกจากการค้นหาปัจจัยแล้วสามารถนำกฎการจำแนกข้อมูลที่ได้มาใช้ในการพัฒนาระบบพยากรณ์ระดับผลการเรียนของนักศึกษาได้อีกด้วย