ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เรื่อง ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

โดย นางสาวบุษบา สังข์วรรณะ

ได้รับอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปรัชญาดุษฎีบัณฑิต สาขาวิชาแทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษา

_____คณบดีบัณฑิตวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ จันทร์วิพัฒน์)

คณะกรรมการสอบวิทยานิพนธ์

_____ประธานกรรมการ
(ศาสตราจารย์ ดร.ปิยะ โควินท์ทวีวัฒน์)

_____กรรมการ
(รองศาสตราจารย์ ดร.พัลลภ พิริยะสุรวงศ์)

_____กรรมการ
(ศาสตราจารย์ ดร.ปรัชญนันท์ นิลสุข)

_____กรรมการ
(รองศาสตราจารย์ ดร.ปณิตา วรรณพิรุณ)

_____กรรมการ
(อาจารย์ ดร.ณรงค์เดช รัตนานนท์เสถียร)

_____กรรมการ
(รองศาสตราจารย์ ดร.ณมน จีรังสุวรรณ)

ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

นางสาวบุษบา สังข์วรรณะ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปรัชญาดุษฎีบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษา คณะครุศาสตร์อุตสาหกรรม
บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2563
ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ชื่อ : นางสาวบุษบา สังข์วรรณะ
ชื่อวิทยานิพนธ์ : ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี
สาขาวิชา : เทคโนโลยีสารสนเทศและสื่อสารเพื่อการศึกษา มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก : รองศาสตราจารย์ ดร.พัลลภ พิริยะสุรวงศ์
ปีการศึกษา : 2563

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) วิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี 2) พัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี 3) ออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี 4) พัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี และ 5) ศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี กลุ่มตัวอย่าง ได้แก่ นักศึกษาภาษาอังกฤษที่เข้าปีการศึกษา 2555-2557 ซึ่งสำเร็จการศึกษาแล้ว จำนวน 141 คน สำหรับพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีและนักศึกษาภาษาอังกฤษที่ลงทะเบียนเรียนปีการศึกษา 2/2563 สำหรับศึกษาผลการใช้งานระบบ โดยวิธีเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling) เทคนิคที่ใช้ ได้แก่ Information Gain เทคนิค Naïve Bayes สร้างตัวแบบและเปรียบเทียบตัวแบบด้วยการทดสอบประสิทธิภาพแบบ 10-fold Cross Validation ผลการวิจัย พบว่า 1) ปัจจัยการพยากรณ์ผลการเรียน มี 14 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอมที่ 1-5 และผลการเรียนในรายวิชา จำนวน 9 วิชา 2) ตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ใช้เทคนิค Naïve Bayes ให้ค่าความถูกต้องร้อยละ 84.33 3) ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีประกอบด้วยส่วน 4 ส่วน ได้แก่ ส่วนของผู้ใช้สำหรับตรวจสอบข้อมูลผลการเรียนส่วนของผู้ดูแลระบบ ส่วนของการพยากรณ์ผลการเรียน และส่วนของการเสนอแผนการเรียนในอนาคต 4) ผลการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ซึ่งผ่านการประเมินประสิทธิภาพจากผู้เชี่ยวชาญ พบว่า ประสิทธิภาพโดยรวมมีประสิทธิภาพอยู่ในระดับดี ผลการประเมินประสิทธิภาพจากผู้ใช้อยู่ในระดับดีมาก และ 5) ผลของการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีสามารถพยากรณ์ผลการเรียนและเสนอแผนการเรียนในอนาคตให้กับนักศึกษากลุ่มอ่อนเพื่อเป็นแนวทางในการเพิ่มผลการเรียนของนักศึกษาและลดการออกกลางคันของนักศึกษาได้

(วิทยานิพนธ์มีจำนวนทั้งสิ้น 196 หน้า)

คำสำคัญ : พยากรณ์ผลการเรียน ระบบอัจฉริยะ ออกกลางคัน

_____อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

Name : Miss Budsaba Sungwanna
Thesis Title : A Intelligent Systems for Learning Result Prediction to Reduce Drop-out of Undergraduate Students
Major Field : Information Technology and Communication for Education King Mongkut's University of Technology North Bangkok
Thesis Advisor : Associate Professor Dr.Pallop Piriyasurawong
Academic Year : 2020

Abstract

The objectives of this research were to; 1) analyze the factors of an intelligent systems for learning results prediction to reduce drop-out of undergraduate students 2) develop the model of an intelligent systems for learning results prediction to reduce drop-out of undergraduate students 3) design an intelligent systems for learning results prediction to reduce drop-out of undergraduate students 4) develop an intelligent systems for learning results prediction to reduce drop-out of ndergraduate students and 5) study the results by using of an intelligent systems for learning results prediction to reduce drop-out of undergraduate students. The samples were 141 undergraduate students who study English Education program in Academic year 2012-2014 at Kanchanaburi Rajabhat University by purposive sampling. The using Information Gain technique and Naive Bayes to construct the model of learning results prediction and compare with the efficiency of 10-fold Cross Validation. The research results were as follows 1) the learning results prediction factors had 14 factors, for example, mean of GPA from semester 1 to 5 and learning results about 9 subjects, 2) the model of an intelligent systems for learning results prediction to reduce drop-out of undergraduate students using Naïve Bayes was 84.33 % for accuracy, 3) an intelligent system for learning results prediction to reduce drop-out of undergraduate students was consisted of 4 parts; part one was the user for determine the learning results, part two was admin, part three was learning results prediction, and the last part was presentation the future of learning plan, 4) the results of an intelligent system development for learning results prediction to reduce drop-out of undergraduate students was passed from the experts showed that the overall efficiency at good level, the results from the user at excellent level, and 5) the results of using of an intelligent system for learning results prediction to reduce drop-out of undergraduate students could predict the learning results and

presenting the future of learning plan to guideline the increasing learning results for the weak group and reduce for drop-out of undergraduate student.

(Total 196 pages)

Keywords : Learning Result Prediction, Intelligent Systems, Drop-out

P. Piriyakul Advisor

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงอย่างสมบูรณ์ได้เนื่องด้วยความกรุณาอย่างยิ่งจากอาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก รองศาสตราจารย์ ดร.พัลลภ พิริยะสุรวงศ์ ที่ได้กรุณาให้คำปรึกษา แนวคิด ข้อชี้แนะ ตลอดจนข้อคิดเห็นอันเป็นประโยชน์อย่างยิ่งแก่งานวิจัยผู้วิจัยรู้สึกซาบซึ้งและขอกราบขอบพระคุณท่านด้วยความเคารพอย่างสูงมา ณ ที่นี้

ขอกราบขอบพระคุณ ศาสตราจารย์ ดร.ปิยะ โควินท์ทวีวัฒน์ ที่ให้ความเมตตาและกรุณาให้เกียรติเป็นประธานสอบป้องกันวิทยานิพนธ์ ศาสตราจารย์ ดร.ปรัชญนันท์ นิลสุข รองศาสตราจารย์ ดร.ปณิตา วรรณพิรุณ รองศาสตราจารย์ ดร.ณมน จีรังสุวรรณ และ ดร.ณรงค์เดช รัตนานนท์เสถียร คณะกรรมการสอบป้องกันวิทยานิพนธ์ที่ให้คำแนะนำที่มีคุณค่าต่อการปรับปรุงงานวิจัย ให้มีความสมบูรณ์มากยิ่งขึ้น ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณมหาวิทยาลัยราชภัฏกาญจนบุรี และผู้เชี่ยวชาญที่ได้สละเวลาให้ข้อเสนอแนะที่เป็นประโยชน์แก่งานวิจัยที่สำเร็จลุล่วงอย่างสมบูรณ์ รวมทั้งสาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษาที่ให้การสนับสนุนและอำนวยความสะดวกในการทำวิทยานิพนธ์ในครั้งนี้

ท้ายนี้ผู้วิจัยใคร่ขอกราบขอบพระคุณบุพการี คุณพ่อประทีป สังข์วรรณะ และคุณแม่ฉลวย สังข์วรรณะ ตลอดจนบุคคลที่เกี่ยวข้องที่ยังไม่ได้กล่าวนามไว้ ณ ที่นี้ที่สนับสนุนและให้กำลังใจ ความห่วงใย จนเป็นแรงผลักดันให้ผู้วิจัยเสมอมาจนสำเร็จการศึกษา

บุษบา สังข์วรรณะ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ง
กิตติกรรมประกาศ	ฉ
สารบัญตาราง	ซ
สารบัญภาพ	ฎ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์การวิจัย	4
1.3 ขอบเขตการวิจัย	4
1.4 นิยามศัพท์เฉพาะ	5
1.5 ประโยชน์ที่คาดว่าจะได้รับ	5
1.6 กรอบแนวคิดที่ใช้ในการวิจัย	6
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	7
2.1 เกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ. 2548	8
2.2 เหมืองข้อมูล (Data Mining)	10
2.3 ระบบสนับสนุนการตัดสินใจ	35
2.4 ขั้นตอนการพัฒนาระบบ	37
2.5 การวิเคราะห์ข้อมูลด้วย RapidMiner Studio	40
2.6 การสร้างแอพพลิเคชั่นด้วย Appsheet	42
2.7 การสร้างแดชบอร์ดด้วย Google Data Studio	43
2.8 การจัดการฐานข้อมูลด้วย Google Sheet	48
2.9 งานวิจัยที่เกี่ยวข้อง	49
บทที่ 3 วิธีการดำเนินการวิจัย	55
3.1 ระยะที่ 1 การวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	55
3.2 ระยะที่ 2 พัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	84
3.3 ระยะที่ 3 ออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	110
3.4 ระยะที่ 4 พัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	114
3.5 ระยะที่ 5 ศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	117

สารบัญ (ต่อ)

	หน้า
บทที่ 4 ผลการวิจัย	121
4.1 ระยะที่ 1 ผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	121
4.2 ระยะที่ 2 ผลการพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	123
4.3 ระยะที่ 3 ผลการออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	128
4.4 ระยะที่ 4 ผลการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	131
4.5 ระยะที่ 5 ผลการศึกษาประสิทธิภาพการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	133
บทที่ 5 ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	137
5.1 บทนำ	137
5.2 ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	141
บทที่ 6 สรุป อภิปรายผล ข้อเสนอแนะ	145
6.1 สรุปผลการวิจัย	145
6.2 อภิปรายผล	147
6.3 ข้อเสนอแนะ	150
บรรณานุกรม	153
ภาคผนวก ก	159
รายนามผู้เชี่ยวชาญ	160
ภาคผนวก ข	161
เครื่องมือการใช้งานระบบ	162
ภาคผนวก ค	167
การวิเคราะห์ข้อมูล	198
ภาคผนวก ง	191
บทความวิจัยที่ได้รับการตอบรับตีพิมพ์ในวารสารวิชาการระดับนานาชาติ	192
ประวัติผู้วิจัย	195

สารบัญตาราง

ตารางที่		หน้า
2-1	คะแนนของนักเรียนในชั้นเรียนเพื่อพิจารณาการแบ่งกลุ่ม	20
2-2	ค่าที่น้อยที่สุดระหว่างของคลัสเตอร์ (1, 2)	20
2-3	ผลลัพธ์การปรับปรุงตารางชุดข้อมูลเมื่อรวมคลัสเตอร์ (1, 2)	21
2-4	ปรับปรุงค่าภายในเมตริกซ์เมื่อรวมคลัสเตอร์ (1, 2)	21
2-5	ผลลัพธ์การปรับปรุงตารางชุดข้อมูลเมื่อรวมคลัสเตอร์ (3, 5)	22
2-6	ปรับปรุงค่าภายในเมตริกซ์เมื่อรวมคลัสเตอร์ (3, 5)	22
2-7	ผลลัพธ์การปรับปรุงตารางชุดข้อมูลเมื่อรวมคลัสเตอร์ (1, 2, 4)	22
2-8	ปรับปรุงค่าภายในเมตริกซ์เมื่อรวมคลัสเตอร์ (1, 2, 4)	22
2-9	ผลลัพธ์การปรับปรุงตารางชุดข้อมูลเมื่อรวมคลัสเตอร์ (1, 2, 3, 4, 5)	23
2-10	Transaction Database ที่จะใช้หากฎความสัมพันธ์	24
2-11	การคำนวณค่า Support ของสินค้าแต่ละชนิด	25
2-12	ทำการตัด Diapers เนื่องจากมีค่า Support ต่ำกว่าค่า Minimum Support	25
2-13	ค่า Support ของ Itemset ที่มีความยาว 2	16
2-14	ค่า Frequent Itemset ที่มีความยาว 2	16
2-15	ค่า Itemset ที่มีความยาวเท่ากับ 3	27
2-16	ค่า Frequent Itemset ทั้งหมด	27
2-17	กฎความสัมพันธ์ทั้งหมดที่สร้างได้พร้อมทั้งค่า Confidence และ Lift	28
3-1	การสังเคราะห์การคัดเลือกปัจจัยการเรียนรู้	57
3-2	แอททริบิวต์สำหรับวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษาระดับปริญญาตรี	60
3-3	อัตราค่าน้ำหนักของปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา	61
3-4	รายละเอียดข้อมูลแอททริบิวต์สำหรับการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียน	63
3-5	ข้อมูลผลการเรียนที่จัดกลุ่มด้วยเทคนิค K-mean	66
3-6	การจัดกลุ่มผลการเรียน (GPAX)	66
3-7	ค่าเฉลี่ยของผลการเรียน (GPAX)	68
3-8	การจัดกลุ่มข้อมูลผลการเรียนนักศึกษาที่สำเร็จการศึกษา	68
3-9	ค่าน้ำหนักของคุณลักษณะจากการใช้เทคนิค Information Gain (IG)	73
3-10	ผลการเปรียบเทียบค่าความถูกต้องของการพยากรณ์ผลการเรียนด้วยเทคนิค Decision Tree จากการลดปัจจัยครั้งละ 1 ปัจจัย	78
3-11	ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Decision Tree	79
3-12	ผลการเปรียบเทียบค่าความถูกต้องของการพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes จากการลดปัจจัยครั้งละ 1 ปัจจัย	81

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
3-13	ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Naïve Bayes	82
3-14	การเปรียบเทียบประสิทธิภาพการพยากรณ์ผลการเรียนด้วยเทคนิค Decision Tree และเทคนิค Naïve Bayes	84
3-15	รายการข้อมูลสำหรับพัฒนาตัวแบบพยากรณ์ผลการเรียนของนักศึกษา	86
3-16	ความน่าจะเป็น P(C) ของคลาส GPAX	93
3-17	โมเดล Naïve Bayes	93
3-18	ข้อมูลนักศึกษาที่ยังไม่ทราบคลาสคำตอบ	97
3-19	ความน่าจะเป็น P(C) ของคลาส GPAX หลังปรับ Laplace Smoothing	98
3-20	โมเดล Naïve Bayes จากการปรับ Laplace Smoothing	98
3-21	โมเดลพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.1	102
3-22	ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.1	104
3-23	โมเดลพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.2	104
3-24	ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.2	106
3-25	โมเดลพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.3	107
3-26	ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.3	109
3-27	ผลการพยากรณ์ผลการเรียน (GPAX) ในอนาคตเมื่อจบการศึกษา	109
3-28	ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา	110
3-29	ข้อมูลนักศึกษา (Student)	113
3-30	ข้อมูลรายวิชา (Subject)	113
3-31	ข้อมูลรายวิชา (Subject)	113
3-32	ข้อมูลสาขาวิชา (Major)	114
3-33	ข้อมูลผลการเรียนแต่ละเทอม (Grade_Term)	114
3-34	ข้อมูลสถานะ (Status)	114
3-35	ข้อมูลอาจารย์ที่ปรึกษา (Advisor)	114
4-1	ปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา	122
4-2	รายการปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี	123
4-3	ความน่าจะเป็น P(C) ของคลาส GPAX	124
4-4	ตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes	124
4-5	การพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาของนักศึกษาปัจจุบัน	128
4-6	ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ	131

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
4-7	ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญด้านความสามารถของระบบตรงความต้องการ	131
4-8	ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญด้านการประเมินประสิทธิภาพการใช้งานของระบบ	132
4-9	ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญด้านผลลัพธ์ที่ได้จากระบบ	132
4-10	ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญด้านความปลอดภัย	133
4-11	ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้ใช้งาน	133
4-12	การพยากรณ์ผลการเรียนของนักศึกษาที่มีระดับผลการเรียนอยู่ในกลุ่มต่ำ	134
4-13	ผลการพยากรณ์ผลการเรียนของนักศึกษากลุ่มตัวอย่าง	135
ค-1	ผลการเรียนของนักศึกษาสำหรับการสร้างตัวแบบพยากรณ์	198
ค-2	ข้อมูลผลการเรียนเฉลี่ย (GPAX) ของนักศึกษาแต่ละคลาส	170
ค-3	ข้อมูลผลการเรียนรายวิชาและผลการเรียนเฉลี่ยแต่ละเทอม	170
ค-4	ความน่าจะเป็น P(C) ของคลาส GPAX	187
ค-5	ตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes	187

สารบัญภาพ

ภาพที่		หน้า
1-1	กรอบแนวคิดการวิจัย	6
2-1	ขั้นตอนกระบวนการทำเหมืองข้อมูล CRISP-DM	11
2-2	การเตรียมข้อมูล	12
2-3	เทคนิคการวิเคราะห์ข้อมูลด้วยดาต้า ไมน์นิง	12
2-4	การคัดเลือกแอททริบิวต์หรือคุณสมบัติที่สำคัญ	15
2-5	การจัดกลุ่มข้อมูลด้วยวิธี K-means	19
2-6	จำนวนคลัสเตอร์เริ่มต้น	20
2-7	ทำการรวมคลัสเตอร์ (1, 2)	21
2-8	ทำการรวมคลัสเตอร์ (3, 5)	21
2-9	การรวมคลัสเตอร์ (1, 2, 4)	22
2-10	ทำการรวมคลัสเตอร์ (1, 2, 3, 4, 5)	23
2-11	สรุปขั้นตอนการรวมคลัสเตอร์ในแต่ละขั้นตอน	23
2-12	กราฟเด็นโดรแกรมที่ได้จากวิธีการ Agglomerative Hierarchical Cluster	23
2-13	ต้นไม้ตัดสินใจ (Decision Tree)	30
2-14	การวัดประสิทธิภาพด้วยเทคนิค Self-consistency Test	33
2-15	การวัดประสิทธิภาพด้วยเทคนิค Split Test	34
2-16	การแบ่งกลุ่มข้อมูลทดสอบ (10-fold Cross-validation)	35
2-17	กระบวนการตัดสินใจและแก้ปัญหา	37
2-18	วงจรพัฒนาระบบ (System Development Life Cycle : SDLC)	38
2-19	ส่วนประกอบหน้าต่าง Design ใน RapidMiner	40
2-20	Operators	41
2-21	เครื่องมือพัฒนาแอพพลิเคชันด้วย Appsheet	42
2-22	โปรแกรม Google Data Studio	44
2-23	การเชื่อมต่อ Data Source	44
2-24	เลือก Google Analytic	45
2-25	เชื่อมต่อ Google Analytics	45
2-26	ผลการเชื่อมต่อ Google Analytics	46
2-27	Data Studio	46
2-28	การสร้างรายงาน	47
2-29	การเพิ่ม Data Source	47
2-30	การเพิ่มรูปแบบของกราฟในรูปแบบต่าง ๆ	48
2-31	รูปแบบของกราฟของ Data Studio	48
3-1	การวิเคราะห์ข้อมูลตามกระบวนการของ CRISP-DM	59

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
3-2	ปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา	62
3-3	แสดงรายละเอียดและลักษณะของข้อมูลแต่ละคุณลักษณะ	65
3-4	จุดศูนย์กลางของผลการเรียนแต่ละกลุ่ม	66
3-5	Clustering (k-Mean)	67
3-6	กราฟผลการเรียน	67
3-7	ขั้นตอนการคัดเลือกคุณลักษณะ	73
3-8	การคัดเลือกคุณลักษณะที่สำคัญด้วยเทคนิค Information Gain	73
3-9	การวัดประสิทธิภาพ แบบ 10-fold Cross Validation	75
3-10	การจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree)	76
3-11	ประสิทธิภาพตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Decision Tree จากปัจจัย 33 ปัจจัย	76
3-12	การจำแนกข้อมูลด้วยเทคนิค Naïve Bayes	77
3-13	ประสิทธิภาพตัวแบบพยากรณ์ผลการเรียนด้วย Naïve Bayes จากปัจจัย 33 ปัจจัย	77
3-14	การลดปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาครั้งละ 1 ปัจจัย	78
3-15	ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Decision Tree	80
3-16	ปัจจัยการพยากาารณ์ผลการเรียน จำนวน 6 ปัจจัย	80
3-17	ประสิทธิภาพการพยากรณ์ผลการเรียนด้วยเทคนิคต้นไม้ตัดสินใจ จากปัจจัย 6 ปัจจัย	81
3-18	ประสิทธิภาพการพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes จากปัจจัย 14 ปัจจัย	83
3-19	ขั้นตอนการสร้างตัวแบบพยากรณ์ผลการเรียน	85
3-20	การแบ่งกลุ่มข้อมูลทดสอบ (10-fold cross-validation)	87
3-21	การพัฒนาตัวแบบพยากรณ์ผลการเรียนโดยใช้เทคนิค Decision Tree	87
3-22	การพัฒนาตัวแบบพยากรณ์ผลการเรียนโดยใช้เทคนิค Naïve Bayes	88
3-23	แอททริบิวต์พยากรณ์ผลการเรียนของนักศึกษา	89
3-24	แผนภาพ Use Case Diagram ของระบบ	112
3-25	แผนผังแสดงความสัมพันธ์ของข้อมูล	112
3-26	แผนภาพการทำงานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา	116
4-1	หน้าจอระบบพยากรณ์ผลการเรียน	129
4-2	เมนูหลักการพยากรณ์ผลการเรียน	130
4-3	รายงานผลการพยากรณ์ผลการเรียน	130

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
5-1	แผนภาพการทำงานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา	141
5-2	หน้าจอระบบพยากรณ์ผลการเรียน	412
5-3	เมนูหลักการพยากรณ์ผลการเรียน	143
5-4	สรุปรายงานผลการพยากรณ์ผลการเรียน	143

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การศึกษาไทยแบ่งการจัดการศึกษาออกเป็น 2 ระดับ คือ ระดับการศึกษาขั้นพื้นฐาน ซึ่งต้องจัดอย่าง 12 ปี และระดับการศึกษาอุดมศึกษา หรือหลังการศึกษาขั้นพื้นฐานโดยจะแบ่งออกเป็นระดับต่ำกว่าปริญญา และปริญญา ซึ่งการจัดการศึกษาระดับอุดมศึกษา มีจุดประสงค์เพื่อพัฒนาผู้เรียนให้มีลักษณะของความเป็นมนุษย์ที่สมบูรณ์ สามารถดำรงตนอยู่ในสังคมพหุวัฒนธรรมภายใต้กระแสโลกาภิวัฒน์ ที่มีการสื่อสารแบบไร้พรมแดน มีศักยภาพในการเรียนรู้ตลอดชีวิต มีความสามารถในการปฏิบัติงานได้ ตามกรอบมาตรฐานและจรรยาบรรณที่กำหนด สามารถสร้างสรรค์งานที่เกิดประโยชน์ต่อตนเอง และสังคม ทั้งในระดับท้องถิ่นและสากล

การจัดการศึกษาในระดับอุดมศึกษา ตามพระราชบัญญัติการอุดมศึกษา พ.ศ. 2562 กำหนดวัตถุประสงค์ไว้ดังนี้ 1) พัฒนาทรัพยากรบุคคลให้มีความเชี่ยวชาญตามสาขาวิชาหรือวิชาชีพที่ตนถนัด สามารถตอบสนองต่อความต้องการของประเทศและสร้างขีดความสามารถในการแข่งขันในระดับโลกได้ 2) พัฒนาบุคคลให้เป็นมนุษย์ที่สมบูรณ์ทั้งร่างกาย จิตใจ สติปัญญา ความรู้ และทักษะที่จำเป็น เป็นคนดี มีวิจัย ภูมิใจในชาติ เข้าใจสังคมและวัฒนธรรมในการดำรงชีวิต สามารถปรับเปลี่ยนตนเองเพื่อรองรับสังคโลกที่จะเปลี่ยนแปลงในอนาคตมีความรับผิดชอบต่อครอบครัว ชุมชน สังคม และประเทศชาติ ร่วมกันแก้ปัญหาสังคม และสามารถอยู่ร่วมกับผู้อื่นได้อย่างมีความสุข และ 3) ตอบสนองยุทธศาสตร์ชาติ แผนแม่บท แผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ แผนการศึกษาแห่งชาติและแผนด้านการอุดมศึกษา และต้องเชื่อมโยงกับการศึกษาในระสดับที่ต่ำกว่า เพื่อเตรียมความพร้อมให้กับทรายากรบุคคลของประเทศในการเข้ารับการศึกษาในระดับอุดมศึกษา รวมทั้งส่งเสริมให้มีการศึกษาอบรมเพื่อเสริมสร้างทักษะในการประกอบอาชีพของบุคคลและการศึกษาตลอดชีวิต

สถาบันอุดมศึกษาเป็นสถาบันทางสังคมซึ่งมีบทบาทโดยตรงต่อการพัฒนาทรัพยากรมนุษย์ การจัดศึกษาในมหาวิทยาลัยมุ่งพัฒนานักศึกษาให้มีความเจริญก้าวหน้าทางสติปัญญาและความคิดอันนำไปสู่ความก้าวหน้าทางวิชาการ ที่จะสร้างสรรค์กำลังคนในระดับวิชาชีพชั้นสูงเพื่อพัฒนาสังคมและมุ่งเน้นพัฒนาคนให้มีความสมบูรณ์ด้านคุณธรรม จริยธรรม เพื่อให้สามารถดำเนินชีวิตอันมีคุณค่าต่อตนเอง สังคมและประเทศชาติ

ตามข้อกำหนดพระราชบัญญัติการศึกษาแห่งชาติกำหนดให้เด็กต้องรับการศึกษาขั้นพื้นฐานและเข้าศึกษาภาคบังคับ จำนวน 9 ปี และเมื่อเข้าศึกษาแล้วจะต้องพยายามให้เด็กศึกษาจนจบภาคการศึกษาภาคบังคับและได้ศึกษาในระดับที่สูงขึ้นต่อไป แต่ปัญหาหลักคือนักเรียนศึกษาเล่าเรียนไม่จบการศึกษาตามเกณฑ์ที่กำหนดและได้ทำการลาออกกลางคัน แต่สาเหตุของการลาออกกลางคันมาจากเหตุและปัจจัยหลายอย่างขึ้นอยู่แต่ละสถานที่แตกต่างกันไป เพื่อเป็นการแก้ปัญหาทำให้เกิดผลเสียต่อประเทศหลายประกาศทั้งความสูญเสียงบประมาณทางการศึกษาที่รัฐจัดให้มีการศึกษา

ทำให้ทรัพยากรบุคคลที่ควรจะมีคุณค่ากลายเป็นทรัพยากรที่มีปัญหา และต้องเสียงบประมาณในการแก้ปัญหาที่ถือเป็นการสูญเปล่าทางการศึกษาเมื่อเทียบกับการลงทุนรวมถึงด้านต่าง ๆ อีกมากมาย ดังนั้นการวิเคราะห์สาเหตุและปัจจัยที่ส่งผลให้นักศึกษาลาออกกลางคันจะสามารถทำให้โรงเรียนนหรือสถาบันระดับอุดมศึกษานำมาเป็นแนวทางในการแก้ปัญหาและลดจำนวนการออกกลางคันของนักศึกษาได้เป็นอย่างดี

มหาวิทยาลัยราชภัฏกาญจนบุรี จัดการเรียนการสอนในระดับอุดมศึกษา เปิดรับนักศึกษารุ่นแรกในปีการศึกษา 2519 ในการบริหารจัดการสิ่งที่เป็นความเสี่ยงของมหาวิทยาลัยด้านหนึ่งคือจำนวนนักศึกษา โดยมหาวิทยาลัยได้จัดทำข้อบังคับมหาวิทยาลัยราชภัฏกาญจนบุรีว่าด้วยการศึกษาระดับอนุปริญญา ปริญญาตรี และปริญญาตรี (ต่อเนื่อง) พ.ศ. 2551 แบ่งออกเป็น 8 หมวด โดยหมวดที่ 6 ได้กล่าวถึงสถานภาพนักศึกษาในข้อที่ 30 สถานภาพนักศึกษา ซึ่งมีข้อย่อยดังนี้ ข้อ 30.1 สถานภาพนักศึกษาสิ้นสุดลงด้วย 3.1.1 ตาย 30.1.2 ลาออก 30.1.3 ขาดคุณสมบัติการเข้าเป็นนักศึกษา 30.1.4 ไม่ลงทะเบียนเรียน ในภาคการศึกษาใดภาคการศึกษาหนึ่ง และไม่ลาพักการเรียน 30.1.5 นักศึกษาภาคปกติได้ค่าระดับคะแนนเฉลี่ยสะสมต่ำกว่า 1.60 เมื่อลงทะเบียนเรียนครบหลักสูตรหรือมีผลการเรียนแล้ว 2 ภาคการศึกษาปกติ หรือได้ค่าระดับคะแนนเฉลี่ยสะสมต่ำกว่า 1.80 เมื่อลงทะเบียนเรียนครบหลักสูตรหรือมีผลการเรียนแล้ว 4 ภาคการศึกษาปกตินับแต่วันเข้าเรียนและทุก ๆ สองภาคการศึกษาปกติถัดไป สำหรับนักศึกษาภาคพิเศษ จะพ้นสภาพการเป็นนักศึกษาเมื่อผลการประเมินได้คะแนนเฉลี่ยต่ำกว่า 1.80 เมื่อสิ้นภาคการศึกษาที่ 4 นับตั้งแต่เริ่มเข้าเรียนกรณีหลักสูตรระดับอนุปริญญาและระดับปริญญาตรี (ต่อเนื่อง) และเมื่อสิ้นภาคการศึกษาที่ 7 นับตั้งแต่เริ่มเข้าเรียนกรณีหลักสูตรปริญญาตรี 4 ปี และสิ้นภาคการศึกษาที่ 8 กรณีหลักสูตรปริญญาตรี 5 ปี หรือนักศึกษาลงทะเบียนครบตามที่หลักสูตรกำหนดแต่ยังได้ค่าระดับคะแนนเฉลี่ยสะสมต่ำกว่า 1.80 การนับภาคการศึกษา นับรวมทั้งภาคการศึกษาที่มีการลาพักการเรียนด้วย เมื่อนักศึกษาเรียนครบตามโครงสร้างของหลักสูตรแล้ว และได้ค่าระดับคะแนนเฉลี่ยสะสมตั้งแต่ 1.80 แต่ไม่ถึง 2.00 ให้เลือกเรียนรายวิชาเพิ่มเติมเพื่อทำค่าระดับคะแนนเฉลี่ยสะสมให้ถึง 2.00 ทั้งนี้ต้องอยู่ในระยะเวลาที่กำหนด หรือตามระยะเวลาการศึกษาที่กำหนดของการจัดการศึกษาภาคพิเศษนั้น ๆ 30.1.6 เมื่อลงทะเบียนเรียนครบกำหนดระยะเวลาการศึกษา 30.1.7 นักศึกษาไม่ผ่านรายวิชาเตรียมฝึกประสบการณ์วิชาชีพ หรือรายวิชาการฝึกประสบการณ์วิชาชีพเป็นครั้งที่ 2 30.1.8 ให้พ้นสถานภาพการเป็นนักศึกษาตามระเบียบว่าด้วยวินัยนักศึกษา และ 30.1.9 เรียนครบตามหลักสูตรและได้รับอนุมัติให้สำเร็จการศึกษา

ระบบงานทะเบียนและประมวลผลของมหาวิทยาลัยราชภัฏกาญจนบุรี มีหน้าที่ติดตามและประเมินผลการเรียนของนักศึกษาที่เรียนทุกภาคการศึกษาตลอดจนครบตลอดหลักสูตร โดยนักศึกษาจะจบการศึกษาได้จะต้องมีผลการเรียนเฉลี่ย (GPAX) ไม่ต่ำกว่าเกณฑ์ที่กำหนด ซึ่งผลการเรียนเฉลี่ยเกิดจากการเรียนของแต่ละวิชาในหลักสูตร ระบบงานทะเบียนและประมวลผลจะบันทึกผลการเรียนนักศึกษาที่ฐานข้อมูลทะเบียนนักศึกษา ซึ่งระบบเหล่านี้มีการเก็บข้อมูลไว้เป็นจำนวนมากและมีการเก็บข้อมูลอย่างต่อเนื่องตั้งแต่นักศึกษาเริ่มเข้าศึกษาจนกระทั่งจบการศึกษา แต่ข้อมูลเหล่านี้ไม่ถูกนำมาใช้ประโยชน์เท่าที่ควรจากการวิเคราะห์ข้อมูลนักศึกษาที่มีการรายงานตัวเข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 4,372 คน ปรากฏว่ามีข้อมูลนักศึกษามารายงานตัวการเป็นนักศึกษาในสาขาวิชาที่

สมัครแต่ไม่มีการเข้าศึกษาในสาขาดังกล่าว จำนวน 459 คน คิดเป็นร้อยละ 10.49 ซึ่งเป็นปัญหาของมหาวิทยาลัยหลาย ๆ มหาวิทยาลัยที่นักศึกษาสมัครเข้าศึกษาหลายสาขาวิชาหรือหลายสถาบันการศึกษาแต่ไม่มีการเข้าศึกษาจริงในวันเปิดเทอมโดยปัญหาดังกล่าวยากต่อการควบคุม ดังนั้นการวิเคราะห์ข้อมูลนักศึกษาจะทำการวิเคราะห์ข้อมูลนักศึกษาเฉพาะที่มีการเข้ามาศึกษาในมหาวิทยาลัยเท่านั้น ซึ่งได้แก่นักศึกษาจำนวน 3,913 คน โดยผลการวิเคราะห์สถานะนักศึกษาที่มีการเข้าศึกษาในมหาวิทยาลัยฯสามารถสรุปได้ดังนี้ 1) นักศึกษาจบการศึกษา จำนวน 2,379 คน คิดเป็นร้อยละ 60.79 2) นักศึกษาตกค้าง จำนวน 243 คน คิดเป็นร้อยละ 6.21 และ 3) ออกกลางคัน จำนวน 1,291 คน คิดเป็นร้อยละ 32.99 ซึ่งจากข้อมูลดังกล่าวการออกกลางคันของนักศึกษาถือว่าเป็นสิ่งที่สำคัญในการบริหารจัดการศึกษา เพื่อหาสาเหตุและแนวทางในการป้องกันแก้ไขปัญหา ดังนั้นการสืบค้นความรู้จากระบบทะเบียนนักศึกษาจะเป็นประโยชน์ โดยการนำข้อมูลในอดีตมาวิเคราะห์เพื่อสร้างต้นแบบพยากรณ์หรือทำนายแนวโน้มสิ่งที่จะเกิดขึ้นในอนาคต เช่น พยากรณ์สถานภาพนักศึกษา พยากรณ์โอกาสการพ้นสภาพของนักศึกษา พยากรณ์โอกาสสำเร็จการศึกษา หรือการพยากรณ์ผลการเรียนของนักศึกษา โดยการนำเทคนิคเหมืองข้อมูล (Data Mining) มาสืบค้นความรู้ในอดีตเพื่อมาวิเคราะห์และสร้างต้นแบบการพยากรณ์ด้วยเทคนิคที่หลากหลาย อาทิเช่น การใช้เทคนิคกฎความสัมพันธ์ (Association Rule) หรือการจำแนกประเภท (Classification) ซึ่งเป็นเทคนิคหนึ่งของการทำเหมืองข้อมูล (Data Mining) เพื่อหาความสัมพันธ์ของผลการเรียนแต่ละวิชาและใช้เทคนิคฟัซซีกฎความสัมพันธ์ (Fuzzy Association Rule Mining) ในการคาดการณ์ผลการเรียนเฉลี่ยว่าอยู่ระดับใด เพื่อนำผลไปวิเคราะห์ผลการเรียนก่อนลงทะเบียน ถ้าผลการคาดการณ์ได้ผลการเรียนอยู่ในระดับต่ำ ผู้เรียนจะได้ปรับแผนการเรียนก่อนการลงทะเบียน และให้ผลการเรียนเป็นไปตามเกณฑ์วัดผล เทคนิคการจำแนกประเภทข้อมูล (Classification) และเทคนิคการสร้างโมเดลเพื่อใช้ในการทำนายหรือพยากรณ์สิ่งที่จะเกิดขึ้น

การออกกลางคันของนักศึกษาเป็นปัญหาสำคัญที่สถาบันการศึกษาจะต้องทำการวิเคราะห์หาปัจจัย สาเหตุ หรือแนวทางการป้องกันปัญหาทำให้นักศึกษาไม่สามารถสำเร็จการศึกษาได้ตามแผนการบริหารการศึกษาของหลักสูตร ซึ่งการออกกลางคันของนักศึกษาสะท้อนและมีผลกระทบต่อประสิทธิภาพในการจัดการศึกษาและการจัดการงบประมาณขององค์กรด้วย สำหรับการออกกลางคันของนักศึกษามีทั้งนักศึกษาที่มีผลการเรียนไม่เป็นไปตามเกณฑ์ที่กำหนด หรือออกกลางคันด้วยเหตุผลอื่น ๆ ซึ่งไม่ว่าจะเป็นเพราะปัจจัยใดนักศึกษาดังกล่าวถือว่าไม่สำเร็จการศึกษาตามที่หลักสูตรกำหนด ถือว่าเป็นการสูญเสียโอกาสในการผลิตบัณฑิต เสียทรัพยากรในการลงทุนซึ่งการออกกลางคันของนักศึกษาถือว่าเป็นสิ่งที่สำคัญอย่างยิ่งในการบริหารจัดการ เพราะเป็นการสูญเปล่าและเป็นความล้มเหลวในการจัดการศึกษา ดังนั้นการนำฐานข้อมูลขนาดใหญ่ของระบบงานทะเบียนและประมวลผลของนักศึกษาที่เข้าศึกษาปีการศึกษา พ.ศ. 2555-2557 วิเคราะห์ข้อมูลด้วยเทคนิคเหมืองข้อมูล (Data Mining) มาใช้วิเคราะห์ข้อมูลนักศึกษาในอดีตที่จัดเก็บไว้เพื่อใช้ในการพยากรณ์ผลการเรียนของแต่ละรายวิชาที่มีความสัมพันธ์นั้นจะเป็นแนวทางแก้ปัญหาหรือป้องกันการลดจำนวนการออกกลางคันน้อยลง ทั้งนี้เพื่อเป็นแนวทางให้อาจารย์ที่ปรึกษา นักศึกษา และผู้บริหารสามารถให้ความช่วยเหลือในการวางแผนการเรียน ดูแลการลงทะเบียนของนักศึกษาได้ จากปัญหาดังกล่าวผู้วิจัยจึง

เล็งเห็นถึงความสำคัญในพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

1.2.2 เพื่อพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

1.2.3 เพื่อออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

1.2.4 เพื่อพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

1.2.5 เพื่อศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

1.3 ขอบเขตการวิจัย

1.3.1 ประชากรและกลุ่มตัวอย่าง

1.3.1.1 ประชากร

นักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่เข้าศึกษาปีการศึกษา 2555-2557 จำนวนทั้งหมด 3,913 คน ซึ่งสำเร็จการศึกษาแล้ว จำนวน 2,379 คน และนักศึกษาที่ลงทะเบียนเรียนในภาคการศึกษา 2/2563 จำนวน 3,579 คน

1.3.1.2 กลุ่มตัวอย่าง

นักศึกษาสาขาภาษาอังกฤษ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่เข้าศึกษาปีการศึกษา 2555-2557 ซึ่งสำเร็จการศึกษาแล้ว จำนวน 141 คน และนักศึกษาปัจจุบันที่ลงทะเบียนเรียนภาคการศึกษา 2/2563 จำนวน 40 คน โดยวิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling)

1.3.2 ตัวแปรที่จะศึกษา

1.3.2.1 ตัวแปรต้น

ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

1.3.2.2 ตัวแปรตาม

1.3.2.2.1 การพยากรณ์ผลการเรียน

1.3.2.2.2 การเสนอแผนการเรียน

1.3.3 เครื่องมือที่ใช้ในการวิจัย

1.3.3.1 ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

1.3.3.2 แบบประเมินคุณภาพระบบพยากรณ์ผลการเรียนอัจฉริยะจากผู้เชี่ยวชาญ

1.3.3.3 แบบประเมินคุณภาพระบบพยากรณ์ผลการเรียนอัจฉริยะจากผู้ใช้งาน

1.4 นิยามศัพท์เฉพาะ

ผลการเรียน หมายถึง ระดับผลการเรียนแต่ละรายวิชาที่ลงทะเบียนเรียนแต่ละภาคการศึกษา โดยถ้าผลการเรียนไม่ผ่านเกณฑ์จะส่งผลกระทบต่อการลงทะเบียนเรียนในภาคการศึกษาต่อไป และผู้เรียนจะต้องลงทะเบียนในรายวิชาดังกล่าวใหม่

เกณฑ์การวัดผล หมายถึง เกณฑ์การวัดผลเกณฑ์ขั้นต่ำของแต่ละรายวิชาและเกณฑ์การสำเร็จการศึกษาตามหลักสูตร โดยสถาบันอุดมศึกษากำหนดให้ผู้เรียนต้องเรียนครบตามจำนวนหน่วยกิตที่กำหนดไว้ในหลักสูตร และต้องได้ระดับคะแนนเฉลี่ยไม่ต่ำกว่า 2.00 จากระบบระดับคะแนนหรือเทียบเท่าจึงถือว่าจบหลักสูตรปริญญาตรี

นักศึกษา หมายถึง นักศึกษาระดับปริญญาตรีภาคปกติของมหาวิทยาลัยราชภัฏกาญจนบุรี ซึ่งอยู่ภายใต้เงื่อนไขและข้อกำหนดเกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ. 2548 ในเรื่องของโครงสร้างหลักสูตรและเกณฑ์การวัดผลและการสำเร็จการศึกษา โดยเป็นนักศึกษาเข้าศึกษาปีการศึกษา 2555-2557 และนักศึกษาปัจจุบันที่กำลังศึกษาอยู่ในปีการศึกษาที่ 2/2563

การออกกลางคัน หมายถึง นักศึกษาที่มีผลการเรียนไม่เป็นไปตามเงื่อนไขการวัดผลการสำเร็จการศึกษาตามข้อกำหนดเกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ. 2548 ซึ่งต้องออกจากระบบการศึกษาก่อนเรียนจบหลักสูตร

การพยากรณ์ผลการเรียน หมายถึง การคาดการณ์หรือการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษาในอนาคต โดยใช้เทคนิควิธีการจำแนกประเภทข้อมูล (Classification) แบบมีผู้สอน (Supervised Learning)

ประสิทธิภาพระบบพยากรณ์ หมายถึง การตรวจสอบค่าความถูกต้องและความเชื่อมั่นของตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิคเหมืองข้อมูลในอนาคต จากการสร้างชุดข้อมูลที่ใช้ในการฝึกสอน (Training Data) และชุดข้อมูลที่ใช้ในการทดสอบ (Testing Data)

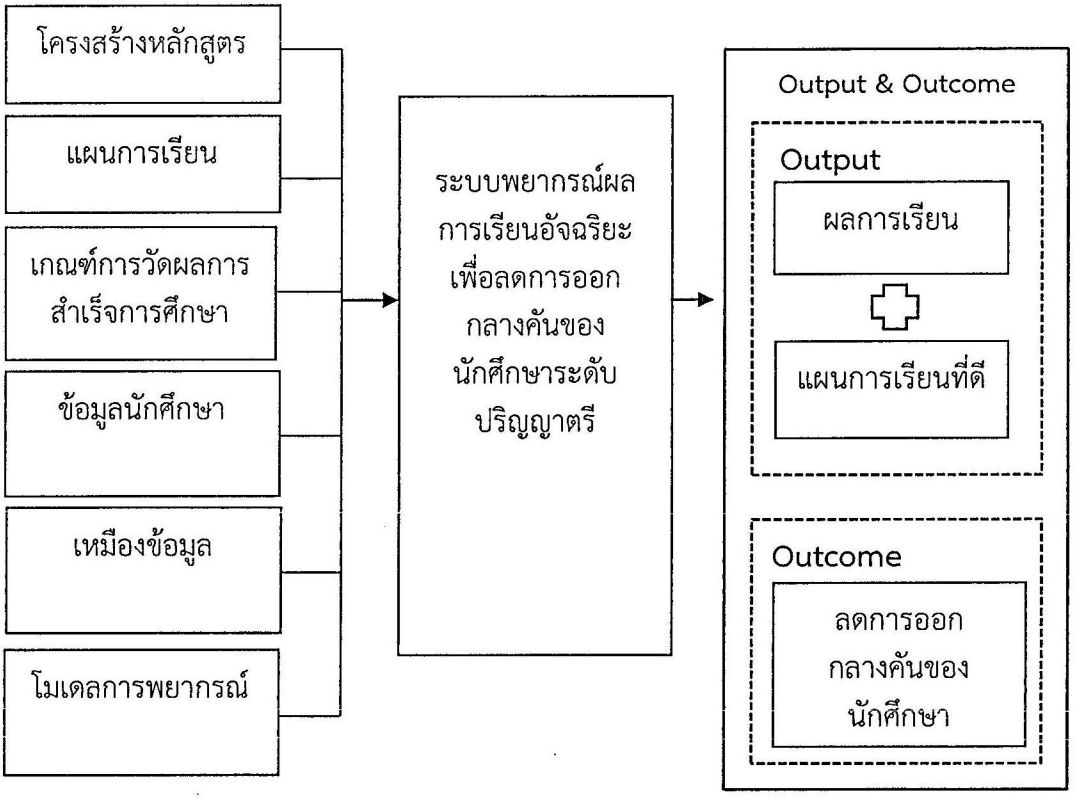
1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ได้ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีที่สามารถลดอัตราการออกคันของนักศึกษาได้

1.5.2 ได้ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีที่สามารถพยากรณ์ผลการเรียนเฉลี่ยในอนาคตของนักศึกษาเมื่อจบการศึกษาได้

1.5.3 ได้ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีที่สามารถนำไปประยุกต์ใช้กับการพยากรณ์ผลการเรียนรายวิชาและแผนการเรียนก่อนลงทะเบียนในปีการศึกษาถัดไปได้

1.6 กรอบแนวคิดการวิจัย (Research Framework)



ภาพที่ 1-1 กรอบแนวคิดการวิจัย

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยเรื่อง ระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ในครั้งนี้ ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องดังต่อไปนี้

2.1 เกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ. 2548
- 2.1.1 โครงสร้างหลักสูตร
- 2.1.2 การลงทะเบียนเรียน
- 2.1.3 เกณฑ์การวัดผลและการสำเร็จการศึกษา

2.2 เหมืองข้อมูล (Data Mining)
- 2.2.1 ความหมายของเหมืองข้อมูล
- 2.2.2 ขั้นตอนการทำเหมืองข้อมูลแบบ CRISP-DM
  - 2.2.2.1 ทำความเข้าใจปัญหา (Business Understanding)
  - 2.2.2.2 การทำความเข้าใจข้อมูล (Data Understanding)
  - 2.2.2.3 การเตรียมข้อมูล (Data Preparation)
  - 2.2.2.4 การสร้างตัวแบบพยากรณ์ (Modeling)
  - 2.2.2.5 การวัดประสิทธิภาพ (Evaluation)
  - 2.2.2.6 การพัฒนาระบบ (Deployment)
- 2.2.3 เทคนิคการวิเคราะห์ข้อมูลด้วยเหมืองข้อมูล
- 2.2.4 การคัดเลือกแอททริบิวต์หรือคุณสมบัติที่สำคัญ (Feature Selection)
  - 2.2.4.1 วิธีฟิลเตอร์ (Filter Approach)
  - 2.2.4.2 วิธีแรปเปอร์ (Wrapper Approach)
  - 2.2.4.3 วิธีฝังตัว (Embed Method)
- 2.2.5 เทคนิคการสร้างตัวแบบ (Modeling)
  - 2.2.5.1 การแบ่งกลุ่มข้อมูล (Clustering)
  - 2.2.5.2 การหาความสัมพันธ์ (Association Rules)
  - 2.2.53 การจำแนกประเภทข้อมูล (Classification)
- 2.2.6 เทคนิคการทดสอบประสิทธิภาพโมเดลแบบการจำแนกประเภทข้อมูล
  - 2.2.6.1 Self-consistency test
  - 2.2.6.2 Split test
  - 2.2.6.3 Cross-validation test

2.3 ระบบสนับสนุนการตัดสินใจ

2.4 ขั้นตอนการพัฒนาระบบ

2.5 โปรแกรม RapidMiner Studio
2.6 การพัฒนาแอพพลิเคชั่นด้วย Appsheet
2.7 การสร้างแดชบอร์ดด้วย Google Data Studio
2.8 ฐานข้อมูลจาก Google Sheet
2.9 งานวิจัยที่เกี่ยวข้อง
2.9.1 งานวิจัยในประเทศ
2.9.2 งานวิจัยต่างประเทศ

2.1 เกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ. 2548

ราชกิจจานุเบกษา เล่ม 122 ตอนพิเศษ 39ง ได้กำหนดเกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ. 2548 เรื่องโครงสร้างหลักสูตร การลงทะเบียนเรียน และเกณฑ์การวัดผลและการสำเร็จการศึกษา ดังนี้

2.1.1 โครงสร้างหลักสูตร (Course Structure)

โครงสร้างหลักสูตร หมายถึง โครงสร้างของหลักสูตร โดยแต่ละสาขาวิชาจะมีความแตกต่างกันออกไป ซึ่งจะมีมาตรฐานการจัดหลักสูตรของแต่ละสาขาตามประกาศของกระทรวงศึกษาธิการว่าด้วย เกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี ประกอบด้วย 3 หมวดหมู่ ได้แก่ หมวดวิชาศึกษาทั่วไป หมวดวิชาเฉพาะ และหมวดวิชาเลือกเสรี โดยมีสัดส่วนจำนวนหน่วยกิตของแต่ละหมวดวิชา ดังนี้

2.1.1.1 หมวดวิชาทั่วไป หมายถึง วิชาที่มุ่งพัฒนาผู้เรียนให้มีความรอบรู้อย่างกว้างขวาง มีโลกทัศน์ที่กว้างไกล มีความเข้าใจธรรมชาติ ตนเอง ผู้อื่น และสังคม เป็นผู้ใฝ่รู้ สามารถคิดอย่างมีเหตุผล สามารถใช้ภาษาในการติดต่อสื่อสารความหมายได้ดี มีคุณธรรม ตระหนักในคุณค่าของศิลปะและวัฒนธรรมทั้งของไทยและของประชาคมนานาชาติ สามารถนำความรู้ไปใช้ในการดำเนินชีวิตและดำรงตนอยู่ในสังคมได้เป็นอย่างดี

สถาบันอุดมศึกษาอาจจัดวิชาศึกษาทั่วไปในลักษณะจำแนกเป็นรายวิชาหรือลักษณะบูรณาการใดก็ได้ โดยผสมผสานเนื้อหาวิชาที่ครอบคลุมสาระของกลุ่มวิชาสังคมศาสตร์ มนุษยศาสตร์ ภาษา และกลุ่มวิทยาศาสตร์กับคณิตศาสตร์ในสัดส่วนที่เหมาะสม เพื่อให้บรรลุวัตถุประสงค์ของรายวิชาศึกษาทั่วไป โดยให้มีจำนวนหน่วยกิตรวมไม่น้อยกว่า 30 หน่วยกิต

อนึ่ง การจัดการศึกษาทั่วไปสำหรับหลักสูตรปริญญาตรี (ต่อเนื่อง) อาจได้รับการยกเว้นรายวิชาที่ได้ศึกษามาแล้วในระดับประกาศนียบัตรวิชาชีพชั้นสูงหรือระดับอนุปริญญา ทั้งนี้จำนวนหน่วยกิตของรายวิชาที่ได้รับการยกเว้นดังกล่าว เมื่อนับรวมกับรายวิชาที่จะศึกษาเพิ่มเติมในหลักสูตรปริญญาตรี (ต่อเนื่อง) ต้องไม่น้อยกว่า 30 หน่วยกิต

2.1.1.2 หมวดวิชาเฉพาะ หมายถึง วิชาแกน วิชาเฉพาะด้าน วิชาพื้นฐานวิชาชีพและวิชาชีพที่มุ่งหมายให้ผู้เรียนมีความรู้ ความเข้าใจและปฏิบัติงานได้ โดยมีจำนวนหน่วยกิตรวม ดังนี้

2.1.1.2.1 หลักสูตรปริญญาตรี (4 ปี) ให้มีจำนวนหน่วยกิตหมวดวิชาเฉพาะรวมไม่น้อยกว่า 84 หน่วยกิต

2.1.1.2.2 หลักสูตรปริญญาตรี (5 ปี) ให้มีจำนวนหน่วยกิตหมวดวิชาเฉพาะรวมไม่น้อยกว่า 114 หน่วยกิต

2.1.1.2.3 หลักสูตรปริญญาตรี (ไม่น้อยกว่า 6 ปี) ให้มีจำนวนหน่วยกิตหมวดวิชาเฉพาะรวมไม่น้อยกว่า 144 หน่วยกิต

2.1.1.2.4 หลักสูตรปริญญาตรี (ต่อเนื่อง) ให้มีจำนวนหน่วยกิตหมวดวิชาเฉพาะรวมไม่น้อยกว่า 42 หน่วยกิต

สถาบันอุดมศึกษาอาจจัดหมวดวิชาเฉพาะในลักษณะวิชาเอกเดี่ยว วิชาเอกคู่ หรือวิชาเอกและวิชาโทก็ได้ โดยวิชาเอกต้องมีจำนวนหน่วยกิตไม่น้อยกว่า 30 หน่วยกิต และวิชาโทต้องมีจำนวนหน่วยกิตไม่น้อยกว่า 15 หน่วยกิต ในกรณีที่จัดหลักสูตรแบบวิชาเอกคู่ต้องเพิ่มจำนวนหน่วยกิตของวิชาเอกอีกไม่น้อยกว่า 30 หน่วยกิต และให้มีจำนวนหน่วยกิตรวมไม่น้อยกว่า 150 หน่วยกิต

2.1.1.3 หมวดวิชาเลือกเสรี หมายถึง วิชาที่มุ่งให้ผู้เรียนมีความรู้ ความเข้าใจ ตามที่ตนถนัดหรือสนใจ โดยเปิดโอกาสให้ผู้เรียนเลือกเรียนรายวิชาใด ๆ ในหลักสูตรระดับปริญญาตรีโดยให้มีจำนวยหน่วยกิตรวมไม่น้อยกว่า 6 หน่วยกิต

สถาบันอุดมศึกษาอาจยกเว้นหรือเทียบโอนหน่วยกิตรายวิชาในหมวดวิชาศึกษาทั่วไป หมวดวิชาเฉพาะ และหมวดวิชาเลือกเสรี ให้กับนักศึกษาที่มีความรู้ความสามารถที่สามารถวัดมาตรฐานได้ ทั้งนี้ นักศึกษาต้องศึกษาให้ครบตามจำนวนหน่วยกิตที่กำหนดไว้ในเกณฑ์มาตรฐานหลักสูตรและเป็นไปตามหลักเกณฑ์การเทียบโอนผลการเรียนระดับปริญญาเข้าสู่การศึกษาในระบบและแนวปฏิบัติที่ดีเกี่ยวกับการเทียบโอน ของสำนักงานคณะกรรมการการอุดมศึกษา

2.1.2 การลงทะเบียนเรียน

การลงทะเบียนเรียน ให้ลงทะเบียนเรียนได้ไม่น้อยกว่า 9 หน่วยกิต และไม่เกิน 12 หน่วยกิต ในแต่ละภาคการศึกษาปกติ สำหรับการลงทะเบียนเรียนเต็มเวลา และให้ลงทะเบียนเรียนได้ไม่เกิน 9 หน่วยกิต ในแต่ละภาคการศึกษาปกติ สำหรับการลงทะเบียนเรียนไม่เต็มเวลา และจะสำเร็จการศึกษาได้ดังนี้

2.1.2.1 หลักสูตรปริญญาตรี (4 ปี) สำเร็จการศึกษาได้ไม่ก่อน 6 ภาคการศึกษาปกติ สำหรับการลงทะเบียนเรียนเต็มเวลา และไม่ก่อน 14 ภาคการศึกษาปกติ สำหรับการลงทะเบียนเรียนไม่เต็มเวลา

2.1.2.2 หลักสูตรปริญญาตรี (5 ปี) สำเร็จการศึกษาได้ไม่ก่อน 8 ภาคการศึกษาปกติ สำหรับการลงทะเบียนเรียนเต็มเวลา และไม่ก่อน 17 ภาคการศึกษาปกติสำหรับการลงทะเบียนเรียนไม่เต็มเวลา

2.1.2.3 หลักสูตรปริญญาตรี (ไม่น้อยกว่า 6 ปี) สำเร็จการศึกษาได้ไม่ก่อน 10 ภาคการศึกษาปกติสำหรับการลงทะเบียนเรียนเต็มเวลา และไม่ก่อน 20 ภาคการศึกษาปกติ สำหรับการลงทะเบียนเรียนไม่เต็มเวลา

2.1.2.4 หลักสูตรปริญญาตรี (ต่อเนื่อง) สำเร็จการศึกษาได้ไม่ก่อน 4 ภาคการศึกษาปกติ สำหรับการลงทะเบียนเรียนเต็มเวลา และไม่ก่อน 8 ภาคการศึกษาปกติสำหรับการลงทะเบียนเรียนไม่เต็มเวลา

สำหรับการลงทะเบียนเรียนในภาคฤดูร้อน ให้ลงทะเบียนเรียนได้ไม่เกิน 9 หน่วยกิต หากสถาบันอุดมศึกษาใดมีเหตุผลและความจำเป็น การลงทะเบียนเรียนที่มีจำนวนหน่วยกิต แตกต่างไปจากเกณฑ์ข้างต้นก็อาจทำได้แต่ทั้งนี้ต้องไม่กระทบกระเทือนต่อมาตรฐานและคุณภาพการศึกษา ทั้งนี้ ต้องเรียนให้ครบตามจำนวนหน่วยกิตตามที่ระบุไว้ในหลักสูตร

2.1.3 เกณฑ์การวัดผลและการสำเร็จการศึกษา

เกณฑ์การวัดผลและการสำเร็จการศึกษา ให้สถาบันอุดมศึกษากำหนดเกณฑ์การวัดผล เกณฑ์ขั้นต่ำของแต่ละรายวิชา และเกณฑ์การสำเร็จการศึกษาตามหลักสูตร โดยต้องเรียนครบตามจำนวนหน่วยกิตที่กำหนดไว้ในหลักสูตร และต้องได้ระดับคะแนนเฉลี่ยไม่ต่ำกว่า 2.00 จากระบบ 4 ระดับคะแนนหรือเทียบเท่า จึงถือว่าเรียนจบหลักสูตรปริญญาตรี

สถาบันอุดมศึกษาที่ใช้ระบบการวัดผลและการสำเร็จการศึกษาที่แตกต่างจากนี้จะต้องกำหนดให้มีคำเทียบเคียงกันได้

2.2 เหมืองข้อมูล (Data Mining)

2.2.1 ความหมายของเหมืองข้อมูล

คำนาย (2557) ได้กล่าวว่า เหมืองข้อมูลหมายถึงซอฟต์แวร์ที่ใช้ในการสร้างคำถามเพื่อให้ได้คำตอบจากข้อมูลที่ซ่อนอยู่ ตามรูปแบบที่กำหนดความสัมพันธ์ระหว่างข้อมูล และกฎเกณฑ์สำหรับการอ้างอิงฐานข้อมูลขนาดใหญ่ สิ่งที่ได้รับคือการคาดการณ์แนวโน้มและพฤติกรรมต่าง ๆ ที่จะเกิดขึ้นในอนาคต ที่มีผลมาจากข้อมูลที่รวบรวมไว้ทำให้ธุรกิจสามารถสร้างความได้เปรียบ ซึ่งสามารถนำความที่ได้รับจากข้อมูลไปใช้เป็นเครื่องมือสนับสนุนการตัดสินใจของผู้บริหารได้ทุกระดับนอกจากนั้นยังสามารถตอบคำถามในทางธุรกิจได้มากมาย สามารถคาดการณ์เหตุการณ์ที่จะเกิดขึ้นในอนาคตได้อย่างแม่นยำ

เอกสิทธิ์ (2557) ได้กล่าวว่า ดาต้า ไมน์นิงคือการค้นพบสิ่งที่มีประโยชน์จากฐานข้อมูลขนาดใหญ่ด้วยการวิเคราะห์ความสัมพันธ์หรือรูปแบบของข้อมูล ซึ่งเทคนิคการวิเคราะห์ข้อมูลด้วยดาต้า ไมน์นิงสามารถแบ่งออกเป็น 2 ประเภทหลัก ๆ ได้แก่ 1) เทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) โดยจะเน้นการพิจารณาข้อมูลเป็นหลัก โดยใช้เทคนิคการค้นหากฎความสัมพันธ์ (Association Rule) หรือการแบ่งกลุ่มข้อมูล (Clustering) และ 2) เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) เป็นการเรียนรู้จากข้อมูลที่มีอยู่ในอดีตเพื่อนำมาสร้างโมเดลสำหรับทำนายหรือคาดการณ์สิ่งที่เกิดขึ้นในอนาคต โดยใช้เทคนิคการจำแนกประเภทข้อมูล (Classification) หรือการประมาณค่าข้อมูล (Regression)

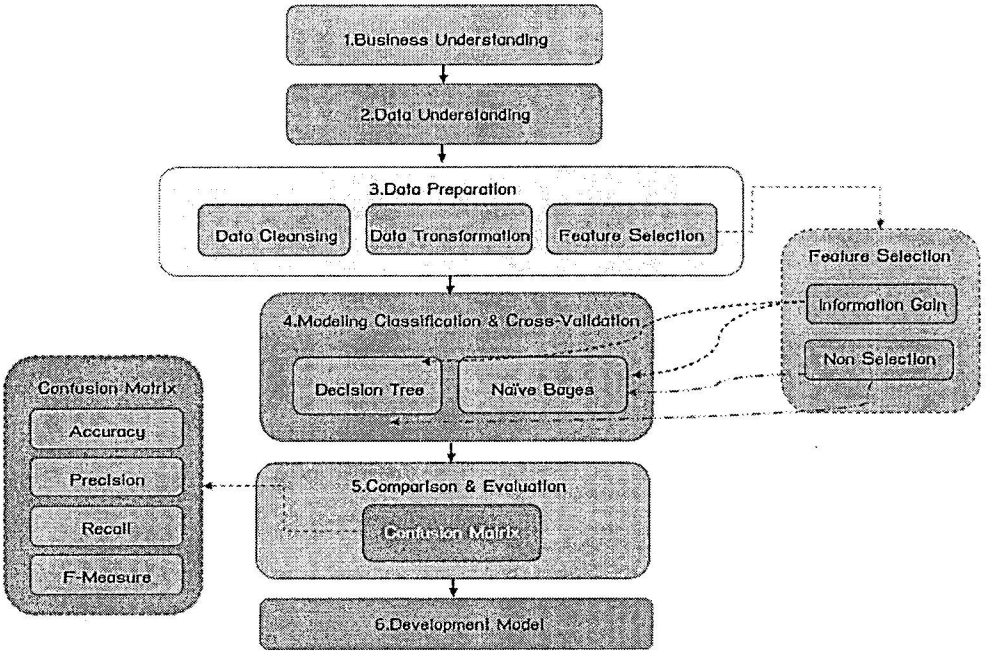
สายชล (2558) ได้กล่าวว่า เหมืองข้อมูลเป็นกระบวนการทำงานที่สกัดข้อมูลจากฐานข้อมูลขนาดใหญ่เพื่อให้ได้สารสนเทศที่มีประโยชน์ที่เรายังไม่ทราบโดยเป็นสารสนเทศที่มีเหตุผลและสามารถนำไปใช้ได้ ซึ่งเป็นสิ่งสำคัญที่จะช่วยการตัดสินใจในการดำเนินงานต่าง ๆ การทำเหมืองข้อมูลสามารถนำมาคาดการณ์การพัฒนาวิวัฒนาการของอนาคตได้

Delavari และคณะ (2008) ได้กล่าวว่า การทำเหมืองข้อมูลเป็นกระบวนการอัตโนมัติที่สกัดความรู้และสารสนเทศที่มีประโยชน์ที่ไม่เคยรู้มาก่อน ซึ่งประกอบด้วยรูปแบบ (Patterns) ความเชื่อมโยง

(Associations) ความเปลี่ยนแปลง (Changes) แนวโน้ม (Trends) ความผิดปรกติ (Anomalies) และ โครงสร้างสำคัญ (Significant Structures) จากเซตข้อมูลที่มีปริมาณมากหรือมีความซับซ้อน

2.2.2 ขั้นตอนการทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) ตามกระบวนการ CRISP-DM ประกอบด้วย 6 ขั้นตอนดังภาพต่อไปนี้



ภาพที่ 2-1 ขั้นตอนกระบวนการทำเหมืองข้อมูล CRISP-DM

ขั้นตอนในกระบวนการ CRISP-DM มีดังนี้

2.2.2.1 ทำความเข้าใจปัญหา (Business Understanding) เป็นขั้นตอนแรกในกระบวนการ CRISP-DM ซึ่งเน้นไปที่การเข้าใจปัญหาและแปลงปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูลทางดาต้า ไมน์นิงพร้อมทั้งวางแผนในการดำเนินการคร่าว ๆ

2.2.2.2 การทำความเข้าใจข้อมูล (Data Understanding) ขั้นตอนนี้เริ่มจากการเก็บรวบรวมข้อมูล หลังจากนั้นจะเป็นการทำความเข้าใจกับข้อมูลที่จะนำมาใช้ความมีลักษณะอย่างไร ตรวจสอบข้อมูลที่ได้ทำการรวบรวมมาได้เพื่อดูความถูกต้องของข้อมูล และพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์

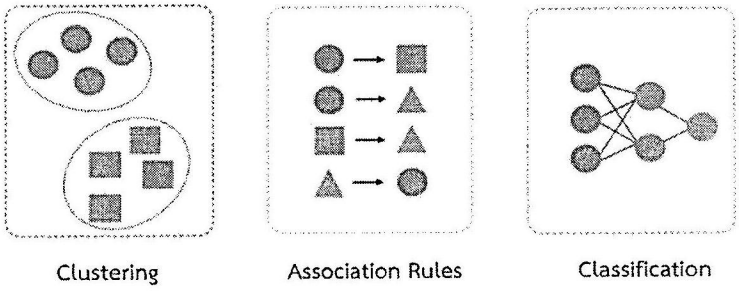
2.2.2.3 การเตรียมข้อมูล (Data Preparation) ขั้นตอนนี้เป็นขั้นตอนที่ทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมา (Raw Data) เพื่อปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่เหมาะสมกลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการแปลงข้อมูลนี้อาจจะต้องมีการทำข้อมูลให้ถูกต้อง (Data Cleaning) เช่น การแปลงข้อมูลให้อยู่ในช่วง (Scale) เดียวกัน หรือการเติมข้อมูลที่ขาดหายไป เป็นต้น โดยขั้นตอนนี้จะเป็นขั้นตอนที่ใช้เวลามากที่สุดของกระบวนการ CRISP-DM โดยการเตรียมข้อมูลที่มีประสิทธิภาพ ควรมีลักษณะสำคัญดังนี้

2.2.2.3.1 ให้ผลลัพธ์ครบถ้วนสมบูรณ์
2.2.2.3.2 ให้ความสำคัญกับนิยามข้อมูล
2.2.2.3.3 จดบันทึกขั้นตอนการเตรียมข้อมูลโดยละเอียด
2.2.2.3.4 ปรับกระบวนการให้เป็นอัตโนมัติมากที่สุด

ภาพที่ 2-2 การเตรียมข้อมูล

2.2.2.4 การสร้างตัวแบบพยากรณ์ (Modeling) ขั้นตอนนี้จะเป็นขั้นตอนการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้า ไมน์นิง เช่น การจำแนกประเภทข้อมูล หรือการแบ่งกลุ่มข้อมูล ซึ่งในขั้นตอนนี้หลายเทคนิคจะถูกนำมาใช้เพื่อให้ได้คำตอบที่ดีที่สุด ดังนั้นในบางครั้งอาจจะต้องมีการย้อนกลับไปที่ขั้นตอนที่ Data Preparation เพื่อแปลงข้อมูลบางส่วนให้เหมาะสมกับแต่ละเทคนิค ตัวอย่างเทคนิคในการวิเคราะห์ข้อมูลต่าง ๆ ดังนี้

2.2.2.4.1 การแบ่งกลุ่มข้อมูล (Clustering)
2.2.2.4.2 การหากฎความสัมพันธ์ (Association Rules)
2.2.2.4.3 การจำแนกประเภทข้อมูล (Classification)



ภาพที่ 2-3 เทคนิคการวิเคราะห์ข้อมูลด้วยดาต้า ไมน์นิง (เอกสิทธิ์, 2560)

2.2.2.5 การวัดประสิทธิภาพ (Evaluation) ในขั้นตอนนี้เราจะได้ผลการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้า ไมน์นิงแล้วก่อนที่จะนำผลลัพธ์ที่ได้ไปใช้งานต่อไปจะต้องมีการวัดประสิทธิภาพของผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ที่ได้ตั้งไว้ในขั้นตอนแรกหรือมีความน่าเชื่อถือมากน้อยเพียงใด ซึ่งอาจจะย้อนกลับไปยังขั้นตอนก่อนหน้าเพื่อเปลี่ยนแปลงแก้ไขเพื่อให้ได้ผลลัพธ์ตามที่ต้องการได้ สำหรับการสร้างโมเดลด้วยเทคนิค Classification มีการทดสอบประสิทธิภาพของโมเดลอยู่ 3 แบบใหญ่ ดังนี้

2.2.2.5.1 Self-consistency test

2.2.2.5.2 Split test

2.2.2.5.3 Cross-validation test

2.2.2.6 การพัฒนาระบบ (Deployment) ในกระบวนการทำงานของ CRISP-DM นั้น ไม่ได้หยุดเพียงแค่ผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้าไมน์นิงเท่านั้น แม้ว่าผลลัพธ์ที่ได้จะแสดงถึงองค์ความรู้ที่มีประโยชน์ แต่จะต้องนำองค์ความรู้ที่ได้เหล่านี้ไปใช้ได้จริงในองค์กรหรือบริษัท ตัวอย่างเช่น การสร้างรายงานเพื่อให้ผู้บริหารหรือนักการตลาดเข้าใจได้ง่ายและสามารถนำไปออกโปรโมชั่นได้ เป็นต้น

2.2.3 เทคนิคการวิเคราะห์ข้อมูลด้วยเหมืองข้อมูล

2.2.3.1 การวิเคราะห์เหมืองข้อมูลตามรูปแบบของการเรียนรู้แบ่งได้เป็น 2 ประเภทดังนี้

2.2.3.1.1 เทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) สามารถแบ่งแยกย่อยได้เป็นเทคนิคการค้นหาความสัมพันธ์ (Association Rule) และการจัดกลุ่มข้อมูล (Clustering)

2.2.3.1.2 เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) เป็นศาสตร์ที่เหมาะกับการสร้างแบบจำลองเพื่อการพยากรณ์ค่าข้อมูล (Predictive Modeling)ในอนาคตซึ่งจะต้องมีการสร้างชุดข้อมูลที่ใช้ในการฝึกฝน (Training Data) และชุดข้อมูลที่ใช้ในการทดสอบ (Test Data) โดย Training Data ต้องมีข้อมูลทั่วไปและ Test Data เป็นคำตอบของ Label หรือ Class ที่กำหนดเพื่อการพยากรณ์ และนำมาใช้ในการวิเคราะห์ Model ในการทำงานของเทคนิคการเรียนรู้แบบมีผู้สอนต้องคำนึงความถูกต้องในการพยากรณ์จึงต้องแบ่งข้อมูลเพื่อ Train กับ Test โดยรูปแบบการวิเคราะห์แบ่งการทำงานเป็นการจำแนกข้อมูล (Classification) และการประมาณค่าข้อมูล (Regression) ซึ่งสามารถสรุปในการแสดงผลลัพธ์แบ่งเป็น 2 วิธี ดังนี้

ก) Regression เป็นตัวเลข

ข) Classification เป็นคลาส Yes No หรือคลาสต่าง ๆ

2.2.3.2 การวิเคราะห์เหมืองข้อมูลตามรูปแบบของ Mahapatra (2001) ซึ่งแบ่งออกเป็น 2 ประเภท คือ เทคนิคด้านคณิตศาสตร์และสถิติ และเทคนิคด้านปัญญาประดิษฐ์ รายละเอียดดังนี้

2.2.3.2.1 เทคนิคด้านคณิตศาสตร์และสถิติ เป็นเทคนิควิธีการทางสถิติที่ใช้การจัดกลุ่มข้อมูล (Cluster Analysis) และสามารถจำแนกข้อมูลออกเป็นกลุ่ม ๆ ได้หรือเทคนิคการวิเคราะห์ถดถอย (Regression) ซึ่งสามารถประยุกต์ใช้กับข้อมูลบางประเภท รวมถึงวิธีการทางสถิติที่ใช้แก้ปัญหาอื่น ๆ เช่น Discriminant Analysis และ Linear Programming เป็นต้น

2.2.3.2.2 เทคนิคด้านปัญญาประดิษฐ์ รูปแบบโครงข่ายประสาทเทียม (Neural Network) เป็นเทคนิคการเรียนรู้ของเครื่องโดยการจดจำรูปแบบ (Pattern Recognition) จากข้อมูลในอดีตและนำรูปแบบที่ได้มาพยากรณ์เพื่อการสนับสนุนการตัดสินใจ โดยการสร้างรูปแบบหรือกฎการตัดสินใจเป็นการกำหนดทางเลือกในการตัดสินใจโดยกำหนดกิ่งก้านสาขาที่ประกอบด้วยปัจจัยที่เป็นไปได้และค้นหาผลที่เป็นไปได้ตามปัจจัยที่เกี่ยวข้อง Case-Based Reasoning เป็นการใช้เหตุการณ์ที่เกิดขึ้นในอดีตเพื่อจดจำลักษณะรูปแบบ (Pattern) Genetic Algorithms เป็นอัลกอริทึม

ทางพันธุ์ศาสตร์ เป็นวิธีการใช้การค้นหาวิธีแก้ปัญหาเพื่อได้จุดที่เหมาะสมที่สุด (Optimum Points) โดยการเลียนแบบขั้นตอนธรรมชาติของการพัฒนาสิ่งมีชีวิต

2.2.4 การคัดเลือกแอททริบิวต์หรือคุณสมบัติที่สำคัญ (Feature Selection)

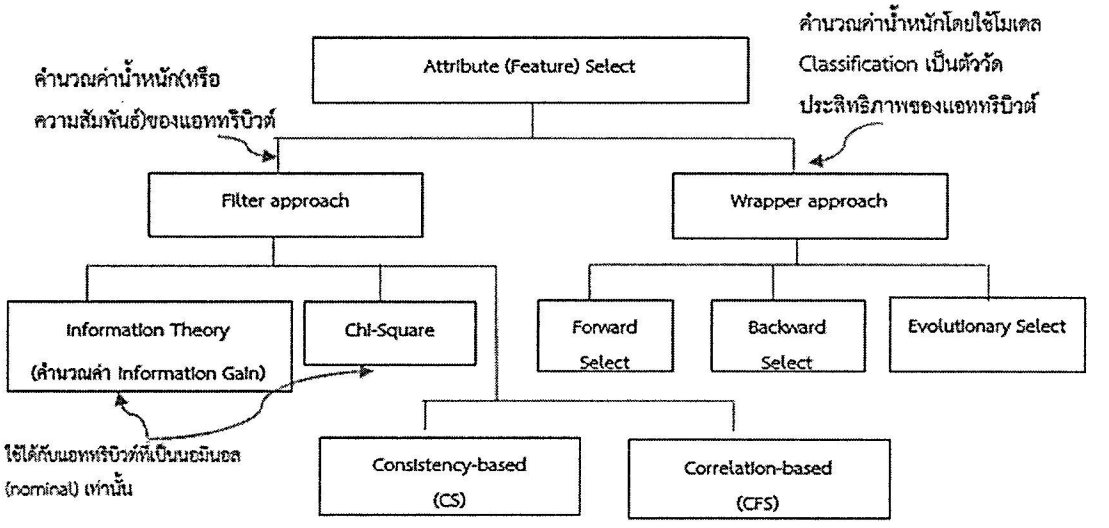
การสร้างตัวแบบพยากรณ์จะมีแอททริบิวต์จำนวนมาก ดังนั้นกระบวนการคัดเลือกแอททริบิวต์หรือคัดเลือกคุณลักษณะเป็นเทคนิคช่วยลดจำนวนแอททริบิวต์หรือตัวแปรที่ใช้ในการสร้างตัวแบบพยากรณ์ ซึ่งอาจเลือกแอททริบิวต์ที่ดีที่สุดเพียงตัวเดียว หรือเลือกกลุ่มของแอททริบิวต์ที่มีความสำคัญต่อการพยากรณ์ โดยการคัดเลือกแอททริบิวต์หรือคุณลักษณะเป็นขั้นตอนที่สำคัญในการเตรียมข้อมูล (Data Preparation) เพราะช่วยลดมิติข้อมูลและทำให้การพยากรณ์มีประสิทธิภาพมากขึ้นโดยเทคนิคการคัดเลือกแอททริบิวต์หรือคุณสมบัติมีหลากหลายเทคนิค ดังรายละเอียดดังนี้

2.2.4.1 กระบวนการเลือกคุณลักษณะแบบวิธีฟิลเตอร์ (Filter Approach) เป็นวิธีที่ใช้เทคนิคการจัดเรียงลำดับของคุณลักษณะหรือเรียกอีกชื่อว่า Feature Ranking technique สำหรับคัดเลือกคุณลักษณะที่เหมาะสม จะเริ่มจากคุณลักษณะทั้งหมดจะถูกวัดค่าความสำคัญ จากนั้นกำหนดระดับ Threshold ขึ้นมาเพื่อกำจัดคุณลักษณะที่มีค่าดัชนีของความสำคัญน้อยกว่าระดับที่อ้างอิงออกไป ตัวอย่างวิธีการเลือกคุณลักษณะ ได้แก่ วิธี Minimal Redundancy and Maximum-Relevance (mRMR) วิธี Relief Feature วิธี Fisher Score Chi-square Test และ Information Gain เป็นต้น ข้อดีของวิธีเลือกคุณลักษณะแบบฟิลเตอร์นี้ คือ เป็นเทคนิคที่คำนวณได้ง่าย รวดเร็ว และหลีกเลี่ยงการเกิดโอเวอร์ฟิตติ้ง (Overfitting) เพราะวิธีนี้ไม่นำผลทดสอบประสิทธิภาพการเรียนรู้ของเครื่องมาพิจารณาร่วมด้วยซึ่งคุณลักษณะที่ถูกคัดเลือกจะไม่ถูกไบแอส สำหรับข้อเสียของวิธีนี้ คือ คุณลักษณะที่ถูกคัดเลือกเป็นคุณลักษณะที่เป็นอิสระต่อกัน เพราะขั้นตอนการคำนวณค่าความสำคัญจะพิจารณาความสัมพันธ์เพียงด้านเดียว คือความสัมพันธ์ระหว่างคุณลักษณะนั้น ๆ กับข้อมูลเอาท์พุตเท่านั้น ไม่ได้คำนึงถึงความสัมพันธ์ระหว่างคุณลักษณะกันเอง ดังนั้นเซตของคุณลักษณะที่ถูกเลือกอาจมีความสัมพันธ์ระหว่างกันหรือไม่มีความสัมพันธ์ระหว่างกันก็ได้ เมื่อนำเซตคุณลักษณะที่ถูกเลือกเหล่านี้มาใช้ในการจำแนกกลุ่มข้อมูล จึงส่งผลทำให้ค่าความถูกต้องของการเรียนรู้ลดลง ดังนั้นการเลือกคุณลักษณะแบบฟิลเตอร์จึงเหมาะสมเฉพาะการวิเคราะห์ข้อมูลที่มีจำนวนมิติไม่สูงมากนัก

2.2.4.2 วิธีเลือกคุณลักษณะแบบวิธีแรปเปอร์ (Wrapper Approach) เป็นวิธีที่ถูกพัฒนามาเพื่อแก้ไขวิธีฟิลเตอร์ ซึ่งวิธีแรปเปอร์มีกระบวนการทำงานที่เพิ่มขึ้นคือ คุณลักษณะทั้งหมดจะถูกจัดให้อยู่ในรูปของเซตคุณลักษณะ หลังจากนั้นดำเนินการค้นหาเซตของคุณลักษณะที่เหมาะสมด้วยการประเมินด้วยฟังก์ชันความเหมาะสม (Fitness Function) เพื่อค้นหาเซตคุณลักษณะที่เหมาะสมที่สุดก่อนเข้าสู่กระบวนการจำแนกข้อมูล วิธีค้นหาเซตคุณลักษณะที่เหมาะสมประกอบด้วย 2 วิธี ได้แก่ 1) วิธีค้นหาเลือกคุณลักษณะโดยลำดับ (Sequential Selection Algorithm) เป็นวิธีการที่เลือกคุณลักษณะด้วยการพิจารณาคุณลักษณะที่เพิ่มทีละตัวตามลำดับ หรือลดคุณลักษณะลงทีละตัวตามลำดับจนกว่าจะได้เซตคุณลักษณะที่เหมาะสม วิธีการนี้เป็นวิธีการที่มีกระบวนการดำเนินงานง่ายแต่มีข้อเสียคือ ใช้เวลาในการประมวลผลในขั้นตอนของการเรียนรู้นานมาก 2) วิธีการค้นหาเซตของคุณลักษณะด้วยวิธีการสุ่มเลือกหรือ วิธีค้นหาคำตอบที่เหมาะสมด้วยอัลกอริทึมของการแก้ไขปัญหาฮิวริสติก (Heuristic Search Algorithm) ได้แก่ Genetic algorithm (GA), Particle Swarm Optimization (PSO) เป็นต้น ซึ่งรายละเอียดของวิธีนี้เราจะอธิบายในหัวข้อการค้นหาคำตอบที่

เหมาะสมด้วยอัลกอริทึม Natural Inspired Algorithms (NIA) กระบวนการค้นหาเซตของคุณลักษณะจะเริ่มต้นด้วย กำหนดจำนวนของเซตของคุณลักษณะและจำนวนของคุณลักษณะที่ถูกเลือก แต่ละเซตของคุณลักษณะจะถูกประเมินด้วยฟังก์ชันความเหมาะสม และค่าฟังก์ชันความเหมาะสมของแต่ละเซตคุณลักษณะถูกนำมาเปรียบเทียบเพื่อค้นหาเซตของคุณลักษณะที่เหมาะสมที่สุด จากนั้นเซตของคุณลักษณะที่เหมาะสมที่สุดจะถูกนำไปใช้กระบวนการจำแนกข้อมูลต่อไป ข้อดีของวิธีการค้นหาแบบฮิวริสติกเมื่อเทียบกับวิธี SSA เป็นวิธีที่ใช้เวลาน้อยกว่า

2.2.4.3 วิธีเลือกคุณลักษณะวิธีฝังตัว (Embed Method) เป็นวิธีถูกออกแบบมาเพื่อแก้ไขข้อเสียของวิธีฟิลเตอร์และวิธีแรปเปอร์ ซึ่งจะใช้เวลาในการประมวลผลน้อยกว่าวิธีแรปเปอร์กระบวนการเลือกคุณลักษณะของวิธีนี้ได้รวมการเลือกคุณลักษณะไว้เป็นส่วนหนึ่งของกระบวนการเรียนรู้ ซึ่งมีข้อดีคือ มีการค้นหาเซตของคุณลักษณะทั้ง Global Space และ Local Space จึงทำให้มีประสิทธิภาพในการค้นหามีประสิทธิภาพที่ดีขึ้น แต่อย่างไรก็ตามวิธีฝังตัวมีข้อเสีย คือ การเลือกเซตคุณลักษณะไม่มีความยืดหยุ่นเนื่องจากขึ้นอยู่กับอัลกอริทึมการจำแนกกลุ่มข้อมูล



ภาพที่ 2-4 การคัดเลือกแอททริบิวต์หรือคุณสมบัติที่สำคัญ

2.2.5 เทคนิคการสร้างตัวแบบ (Modeling)

การสร้างตัวแบบพยากรณ์ (Modeling) เป็นขั้นตอนหนึ่งของการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้าไมน์นิง ซึ่งมีเทคนิคการสร้างตัวแบบหลายรูปแบบอาทิเช่น การหาความสัมพันธ์ การจำแนกประเภทข้อมูล หรือการแบ่งกลุ่มข้อมูล ซึ่งในขั้นตอนนี้หลายเทคนิคจะถูกนำมาใช้เพื่อให้ได้คำตอบที่ดีที่สุด ดังนั้นในบางครั้งอาจจะต้องมีการย้อนกลับไปที่ขั้นตอนที่ Data Preparation เพื่อแปลงข้อมูลบางส่วนให้เหมาะสมกับแต่ละเทคนิคด้วย ตัวอย่างเทคนิคการวิเคราะห์ข้อมูลต่าง ๆ ดังนี้

2.2.5.1 เทคนิคการแบ่งกลุ่มข้อมูล (Clustering Analysis)

เทคนิคนี้การแบ่งกลุ่มข้อมูลเรียกว่า Clustering Analysis จัดอยู่ในประเภทของการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) โดยการแบ่งกลุ่มหน่วยข้อมูล หรือเป็นการแบ่งคน สัตว์ สิ่งของ องค์กร ฯลฯ ออกเป็นกลุ่มย่อยอย่างน้อย 2 กลุ่ม โดยมีหลักเกณฑ์ในการแบ่งดังนี้ “ให้หน่วยที่

อยู่ในกลุ่มเดียวกันมีลักษณะที่สนใจเหมือนกันหรือคล้ายกัน แต่หน่วยที่อยู่ต่างกลุ่มกันจะมีลักษณะที่สนใจต่างกัน" แนวทางของการนำเทคนิคแบ่งกลุ่มข้อมูลไปใช้ คือ จะไม่ใช่การหาผลลัพธ์ที่ต้องการวัดค่าความแม่นยำหากแต่ต้องการหาความสัมพันธ์ของข้อมูลอีกรูปแบบหนึ่ง เช่น การจัดกลุ่มลูกค้าจากพฤติกรรมการซื้อสินค้า การจัดกลุ่มนักท่องเที่ยวจากพฤติกรรมการไปยัง สถานที่ต่าง ๆ หรือการจัดกลุ่มคนไข้ตามอาการหรือความรุนแรงของโรค เป็นต้น เมื่อได้กลุ่มข้อมูลแล้วก็ค่อยเข้าไปดูลักษณะของแต่ละกลุ่มและกำหนดชื่อให้กลุ่มนั้น ๆ การทำ Clustering มีหลากหลายวิธี ซึ่งวิธีการแบ่งกลุ่มที่พบบ่อยมีอยู่ ได้แก่ Centroid-based Clustering พิจารณาความหนาแน่นของข้อมูล (Density-based Clustering) Distribution-based Clustering และการแบ่งข้อมูลแบบลำดับชั้น (Hierarchical Clustering) เป็นต้น วิธีการเลือกใช้การแบ่ง Cluster เมื่อทำการแบ่ง Cluster เรียบร้อยแล้วขั้นตอนต่อไปให้ทำการประเมินคุณภาพของ Cluster และสามารถปรับ Parameters เพื่อให้ดีขึ้นได้

การจัด Case (คน สัตว์ สิ่งของ หรือ องค์กร ฯลฯ) หรือเป็นการจัดตัวแปรออกเป็นกลุ่มย่อย ๆ ตั้งแต่ 2 กลุ่มขึ้นไป Case ที่อยู่ในกลุ่มเดียวกันจะมีลักษณะที่เหมือนกันหรือคล้ายกัน ส่วน Case ที่อยู่ต่างกลุ่มกันจะมีลักษณะที่แตกต่างกัน

ตัวแปรที่อยู่ในกลุ่มเดียวกันจะมีความสัมพันธ์กันมากกว่าตัวแปรที่อยู่ต่างกลุ่มกัน

ตัวแปรที่อยู่ต่างกลุ่มกันจะมีความสัมพันธ์กันน้อยหรือไม่มีความสัมพันธ์กันเลย

เช่น มีวัตถุประสงค์ที่จะจัดสิ่งของ (Object) n สิ่งให้อยู่เป็นกลุ่ม ๆ โดยที่สิ่งของที่อยู่ในกลุ่มเดียวกันมีลักษณะคล้ายกัน (Similarity) หรือใกล้ชิดกัน (Closeness) ดังตัวอย่างผังข้อมูลดังนี้

	ตัวแปรที่ 1	ตัวแปรที่ 2...	ตัวแปรที่ K...	ตัวแปรที่ P
หน่วยที่ 1:	X_{11}	X_{12}	X_{1k}	X_{1p}
หน่วยที่ 2:	X_{21}	X_{22}	X_{2k}	X_{2p}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
หน่วยที่ j:	X_{j1}	X_{j2}	X_{jk}	X_{jp}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
หน่วยที่ n:	X_{n1}	X_{n2}	X_{nk}	X_{np}

โดยที่ X_{jk} หมายถึง การวัดหรือค่าครั้งที่ j ของตัวแปร k

จะเห็นว่าแต่ละหน่วยถูกวัดด้วยตัวแปร p ตัว หน่วย หมายถึง สิ่งของ ในการวิเคราะห์จัดกลุ่มอาจเรียกทับศัพท์ เช่น Item Case หรือ Object แล้วแต่คำไหนสื่อความหมายได้มากกว่ากันการวิเคราะห์จัดกลุ่มนอกจากจะจัดสิ่งของให้อยู่เป็นกลุ่ม ๆ แล้ว ยังสามารถใช้แนวคิดนี้จัดตัวแปร p ตัวให้อยู่เป็นกลุ่ม ๆ ได้ด้วย

2.2.5.1.1 การประยุกต์ใช้การจัดกลุ่มข้อมูลกับงานด้านต่าง ๆ

ก) ด้านการแพทย์

(1) จัดกลุ่มคนไข้ตามอาการหรือความรุนแรงของโรค เพื่อใช้วิธีการรักษาที่แตกต่างกันตามความรุนแรงของโรค

(2) จัดกลุ่มโรงพยาบาลที่มีประสิทธิภาพคล้ายกันไว้ด้วยกัน

(3) จัดกลุ่มประเทศต่าง ๆ ตามความเจริญด้านสาธารณสุข โดยใช้ตัวแปรหรือดัชนีด้านสาธารณสุข เช่น อัตราคนป่วยโรคต่าง ๆ อายุเฉลี่ย ค่ารักษาพยาบาลเฉลี่ยต่อประชากร 1 คน เป็นต้น

ข) ด้านการตลาด

(1) แบ่งผู้บริโภคหรือลูกค้าตามพฤติกรรมการบริโภคสินค้าต่าง ๆ โดยให้ลูกค้าที่มีพฤติกรรมการบริโภคหรือการซื้อสินค้าที่คล้ายกันอยู่ในกลุ่มเดียวกัน ส่วนลูกค้าที่มีพฤติกรรมการบริโภคต่างกันจะอยู่ต่างกลุ่มกัน เมื่อจัดกลุ่มแล้วจะทำให้สามารถวางแผนกลยุทธ์ทางการตลาดสำหรับลูกค้าแต่ละกลุ่มได้อย่างมีประสิทธิภาพ ตัวแปรที่นำมาใช้ในการจัดกลุ่มอาจใช้ตัวแปรด้านพฤติกรรมต่าง ๆ ของลูกค้า

(2) ใช้วางแผนทางด้านการตลาดในพื้นที่ที่แตกต่างกัน โดยเริ่มต้นด้วยการใช้เทคนิคการวิเคราะห์กลุ่มแบ่งพื้นที่ หรือจังหวัดที่ประชากรมีพฤติกรรมการบริโภคคล้ายกัน หรือมีลักษณะประชากรศาสตร์คล้ายกัน เช่น จำนวนประชากร รายได้เฉลี่ย ขนาดพื้นที่ อาชีพ ทัศนคติของคนในพื้นที่ หรือเป็นพื้นที่ที่มีสภาพเศรษฐกิจคล้ายกันไว้ในกลุ่มเดียวกัน

ค) ด้านการศึกษา

(1) จัดกลุ่มนักเรียนตามผลการเรียน (GPAX) ระดับสติปัญญา (IQ) ระดับการศึกษาของผู้ปกครอง เพื่อให้ได้นักเรียนในกลุ่มเดียวกัน ผลการเรียน ระดับสติปัญญาและระดับการศึกษาของผู้ปกครองใกล้เคียงกัน ส่วนนักเรียนที่อยู่ต่างกลุ่มกันจะมีผลการเรียนระดับสติปัญญาและการศึกษาของผู้ปกครองต่างกัน เพื่อให้ครูผู้สอนสามารถวางแผนหรือเลือกเนื้อหาวิธีการสอนตามความเหมาะสมของแต่ละกลุ่ม โดยต่างกลุ่มกันอาจต้องใช้วิธีการสอนที่แตกต่างกันเพื่อทำให้เกิดผลสัมฤทธิ์มากที่สุด

2.2.5.1.2 ขั้นตอนการสร้างโมเดลด้วยวิธีการแบ่งกลุ่มข้อมูล

ขั้นตอนในการสร้างโมเดลประเภทแบบไม่มีผู้สอน (Unsupervised Learning) จะมีความต่างจากการสร้างโมเดลแบบมีผู้สอน (Supervised Learning) คือ การสร้างโมเดลนั้นจะไม่มีการแบ่งชุดข้อมูลออกเป็นชุดสำหรับฝึกสอนและทดสอบ เนื่องจากเป็นชุดข้อมูลที่ไม่มีคลาสคำตอบ จึงทำให้ไม่จำเป็นต้องแบ่งชุดข้อมูลออกเป็น 2 ส่วนเหมือนกับเทคนิคที่ใช้สร้างโมเดลแบบมีผู้สอน ซึ่งมีขั้นตอนของการสร้างโมเดลเป็น 4 ขั้นตอน ดังนี้

ก) กำหนดวิธีวัดความคล้ายหรือความต่างของข้อมูล ตัวอย่างวิธีที่นิยมใช้ เช่น ยูคลิเดียน (Euclidean Distance) โคไซน์ (Cosine) และแมนฮัตตัน (Manhattan Distance) เป็นต้น วิธีการวัดความคล้ายของข้อมูลด้วยวิธีวัดระยะแบบยูคลิเดียน (Euclidean Distance) สามารถคำนวณหาได้ดังสมการที่ (2-1)

$$D = \sqrt{\sum_{i=1}^{d} \left(x_1^i - x_2^i\right)^2} \quad (2\text{-}1)$$

โดยที่

D หมายถึง ระยะห่างระหว่างข้อมูลที่ 1 และ 2 ซึ่งถ้าค่าเท่ากับ 0 แสดงว่าข้อมูลมีความเหมือนกัน

x_1^i หมายถึง ค่าของแอทริบิวต์ที่ i ของข้อมูลที่ 1

x_2^i หมายถึง ค่าของแอทริบิวต์ที่ i ของข้อมูลที่ 2

ข) เลือกอัลกอริทึมที่ใช้ในการจัดกลุ่มข้อมูล ซึ่งหลัก ๆ แบ่งได้เป็น 2 ประเภท ได้แก่ (1) ประเภทที่มีการแบ่งกลุ่มอย่างชัดเจน (Hard Clustering) เป็นเทคนิคที่มีการแบ่งข้อมูลออก จากกันเป็นกลุ่ม ๆ อย่างสิ้นเชิง โดยแต่ละข้อมูลนั้นจะถูกจัดให้อยู่ในกลุ่มใดกลุ่มหนึ่งเท่านั้น เทคนิคประเภทนี้ ได้แก่ เทคนิคการจัดกลุ่มข้อมูลแบบเคมีน (K-means Clustering Algorithm) และ (2) ประเภทที่มีการแบ่งกลุ่มแบบไม่ชัดเจน (Soft Clustering) เป็นเทคนิคการแบ่งที่ข้อมูลสามารถอยู่ในหลาย ๆ กลุ่มได้ โดยขึ้นอยู่กับความน่าจะเป็นของตัวข้อมูล เทคนิคประเภทนี้ ได้แก่ เทคนิคการจัดกลุ่มแบ่งกลุ่มข้อมูลแบบลำดับชั้น (Hierarchical Clustering Methods)

ค) กำหนดจำนวนกลุ่มที่ต้องการ ซึ่งในอัลกอริทึมประเภทที่มีการแบ่งกลุ่มอย่างชัดเจนจำเป็นต้องกำหนดจำนวนกลุ่มที่ต้องการ แต่ถ้าเป็นอัลกอริทึมประเภทที่มีการแบ่งกลุ่มแบบไม่ชัดเจนไม่จำต้องกำหนดจำนวนกลุ่มก็ได้อัลกอริทึมจะทำการหาจำนวนกลุ่มที่เหมาะสมให้กับชุดข้อมูลเอง

ง) ประเมินผลการวิเคราะห์การจัดกลุ่มเนื่องจากเทคนิคการวิเคราะห์จัดกลุ่มเป็นประเภทโมเดลแบบไม่มีผู้สอน ซึ่งจะมีวิธีการวัดผลที่แตกต่างไปจากเทคนิคประเภทแบบมีผู้สอน คือ จะไม่สามารถวัดผลได้จากเปรียบเทียบหาความแม่นยำ เนื่องจากไม่มีผลลัพธ์ตั้งต้นให้เปรียบเทียบ แต่จะสามารถวัดผลได้ด้วยวิธีอื่น เช่น วัดจากความพึงใจในตัวโมเดลที่ได้ หรือวัดผลด้วยวิธีอื่น ๆ เช่น การใช้วิธีการวิเคราะห์จัดกลุ่มข้อมูล เพื่อระบุตำแหน่งในการสร้างศูนย์กระจายสินค้าโดยต้องการให้ได้ศูนย์กระจายสินค้าที่สามารถลดค่าใช้จ่ายในการขนส่งได้มากที่สุด จากตัวอย่างวิธีการวัดผลสามารถทำได้ด้วยการนำโมเดลการจัดกลุ่มที่ได้ ไปสร้างโมเดลในการหาค่าที่เหมาะสมอีกครั้ง (Optimization Model) เพื่อหาจุดที่ให้ค่าขนส่งโดยรวมน้อยที่สุด (อสมา, 2561) เป็นต้น

2.2.5.1.3 การจัดกลุ่มข้อมูลแบบเคมีน (K-mean Clustering Algorithm) เป็นวิธีการที่ถูกนำใช้บ่อยมากที่สุด เนื่องจากมีขั้นตอนการทำงานที่ไม่ซับซ้อนและเข้าใจได้ง่ายโดยมีขั้นตอนการทำงานดังนี้

ขั้นตอนที่ 1 เริ่มต้นจากการกำหนดค่า K หรือจำนวนกลุ่มข้อมูลที่ต้องการ

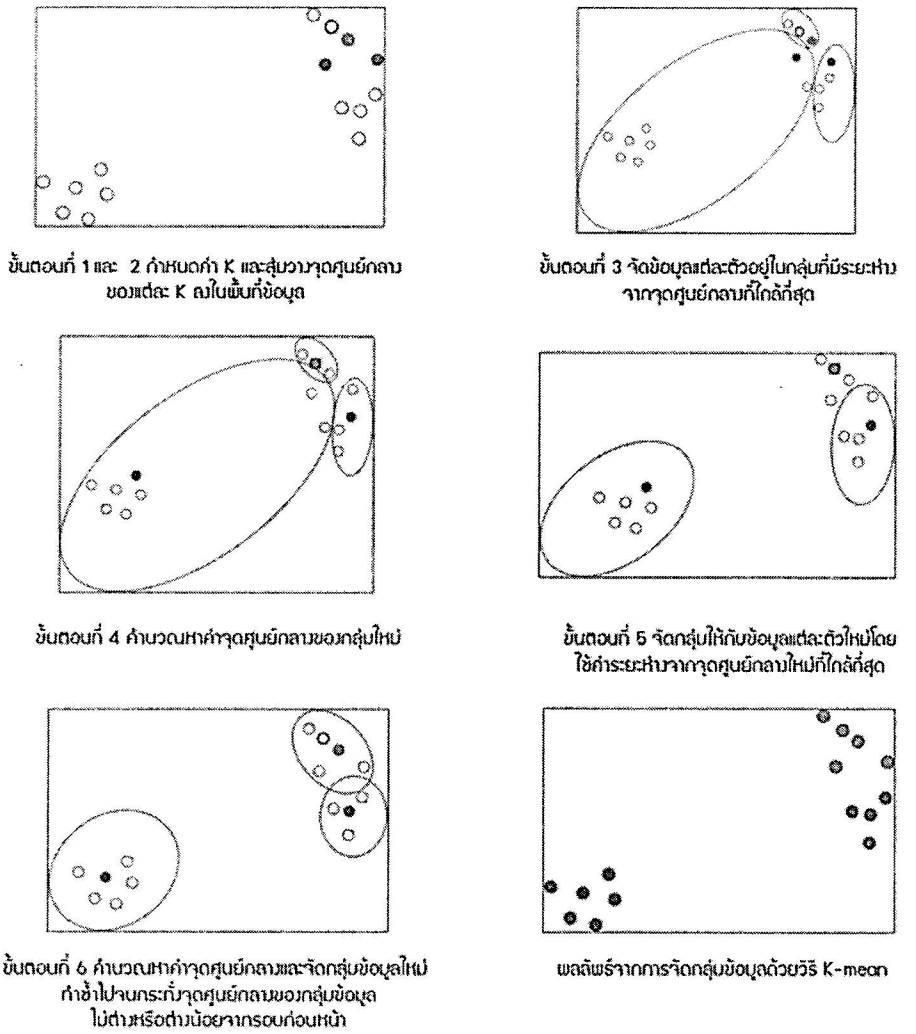
ขั้นตอนที่ 2 สุ่มวางตำแหน่งจุดศูนย์กลางของข้อมูลแต่ละกลุ่ม หรือเซ็นทรอยด์ (Centroid)

ขั้นตอนที่ 3 จากจุดศูนย์กลางแต่ละจุดจะมีการคำนวณระยะทางกับทุกข้อมูลในชุดข้อมูลซึ่งคำนวณโดยวิธียูคลิเดียน จากนั้นแต่ละข้อมูลจะถูกจัดอยู่ในกลุ่มของจุดศูนย์กลางที่มีระยะทางใกล้ที่สุดเท่านั้น

ขั้นตอนที่ 4 หลังจากทำการจัดกลุ่มข้อมูลใหม่แล้วทำการคำนวณค่าเฉลี่ยของสมาชิกในกลุ่มเพื่อกำหนดเป็นจุดศูนย์กลางของกลุ่มข้อมูลใหม่

ขั้นตอนที่ 5 ทำซ้ำ ข้อ 3 ถึง 4 จนกระทั่งค่าจุดศูนย์กลางของกลุ่มข้อมูลใหม่ได้ค่าไม่ต่างหรือต่างเพียงเล็กน้อยจากค่าจุดศูนย์กลางรอบก่อนหน้า

สมมติเราต้องการแบ่งกลุ่มข้อมูลที่กระจายกันอยู่ออกเป็น 3 โดยใช้ K-Means ขั้นตอนการทำงานแสดงได้ดังภาพที่ 2-5



ภาพที่ 2-5 การจัดกลุ่มข้อมูลด้วยวิธี K-Means

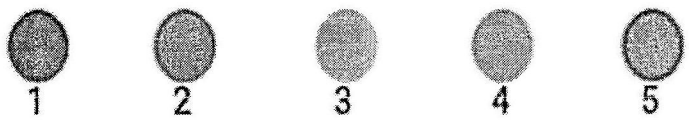
2.2.5.1.4 การจัดกลุ่มแบบโครงสร้างลำดับชั้น (Hierarchical Clustering) เป็นการจัดกลุ่มโดยไม่ต้องมีการกำหนดจำนวนกลุ่มที่ต้องการจัดกลุ่มข้อมูลก่อน เป็นการวิเคราะห์แบบเป็นขั้นตอน วิธีการที่นิยมคือวิธีการ “Agglomerative Hierarchical Cluster” การจัดกลุ่มแบบโครงสร้างลำดับชั้นนั้นจะประกอบด้วย 4 ขั้นตอนดังต่อไปนี้ (Mulika, 2019)

ตัวอย่าง การแบ่งกลุ่มนักเรียนในห้องโดยพิจารณาจากคะแนนของนักเรียนแต่ละคน โดยผลคะแนนแสดงดังตารางที่ 2-1

ตารางที่ 2-1 คะแนนของนักเรียนในชั้นเรียนเพื่อพิจารณาการแบ่งกลุ่ม

Student_id	Marks
1	10
2	7
3	28
4	20
5	35

ขั้นตอนที่ 1 กำหนดจำนวนกลุ่มที่ต้องการจัดกลุ่มของข้อมูลคือคลัสเตอร์ ดังนั้นมีนักเรียน 5 คนจึงทำให้มีคลัสเตอร์ทั้งหมด 5 คลัสเตอร์ ดังภาพที่ 2-6 แสดงจำนวนคลัสเตอร์เริ่มต้น



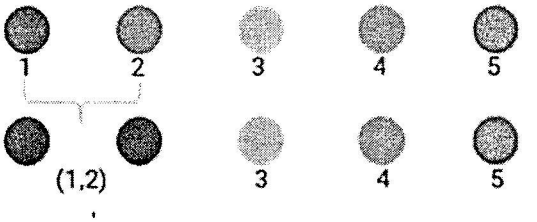
ภาพที่ 2-6 จำนวนคลัสเตอร์เริ่มต้น

ขั้นตอนที่ 2 ทำการสร้างเมตริกซ์ขนาด N*N จากชุดข้อมูล และคำนวณระยะห่างของแต่ละจุดข้อมูล โดยใช้สูตรเดียวกับที่ใช้คำนวณในสมการ K-Mean ดังสมการที่ (2-1) เช่น ระยะห่างระหว่างข้อมูลของคะแนนคนที่ 1 และ 2 = $\sqrt{(10-7)^2}$ = 3

ขั้นตอนที่ 3 มองหาค่าระยะห่างระหว่างข้อมูลที่น้อยที่สุดโดยทำการรวมข้อมูลเป็นคลัสเตอร์เดียวกัน และทำการปรับปรุงเมตริกซ์ จากตารางที่ 2-2 ค่าที่น้อยที่สุดคือ (1, 2) = 3 และเมื่อทำการรวมคลัสเตอร์ (1, 2) แสดงดังภาพที่ 2-7

ตารางที่ 2-2 ค่าที่น้อยที่สุดระหว่างของคลัสเตอร์ (1,2)

ID	1	2	3	4	5
1	0	3	18	10	25
2	3	0	21	13	28
3	18	21	0	8	7
4	10	13	8	0	15
5	25	28	7	15	0



ภาพที่ 2-7 ทำการรวมคลัสเตอร์ (1,2)

ขั้นตอนที่ 4 ทำการปรับปรุงตารางชุดข้อมูล เมื่อทำการรวมคลัสเตอร์แล้ว จะเหลือค่าที่มากที่สุดของทั้ง 2 คลัสเตอร์เอาไว้ จากตัวอย่าง คะแนนคนที่ 1 = 10 และคนที่ 2 = 7 เมื่อรวมคลัสเตอร์ (1,2) เข้าด้วยกัน คะแนนสอบที่จะเป็นตัวแทนของกลุ่มใหม่คือ 10 นั่นคือ (1,2) = 10 ดังตารางที่ 2-3 และปรับปรุงเมตริกซ์ดังตารางที่ 2-4

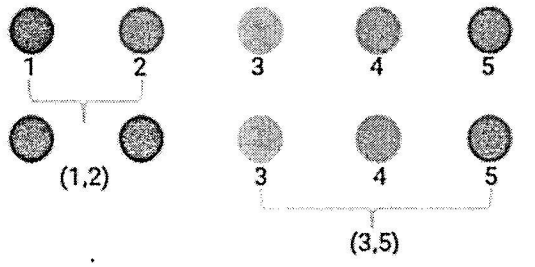
ตารางที่ 2-3 ผลลัพธ์การปรับปรุงตารางชุดข้อมูลเมื่อรวมคลัสเตอร์ (1, 2)

Student_id	Marks
(1,2)	10
3	28
4	20
5	35

ตารางที่ 2-4 ปรับปรุงค่าภายในเมตริกซ์เมื่อรวมคลัสเตอร์ (1, 2)

ID	(1,2)	3	4	5
(1,2)	0	18	10	25
3	18	0	8	7
4	10	8	0	15
5	25	7	15	0

ขั้นตอนที่ 5 ทำซ้ำขั้นตอนที่ 3-4 จนกว่าจะเหลือเพียงคลัสเตอร์เดียว
ขั้นตอนที่ 5.1 ทำการรวมคลัสเตอร์ 3, 5 และได้ผลลัพธ์ดังภาพที่ 2-8



ภาพที่ 2-8 ทำการรวมคลัสเตอร์ (3, 5)

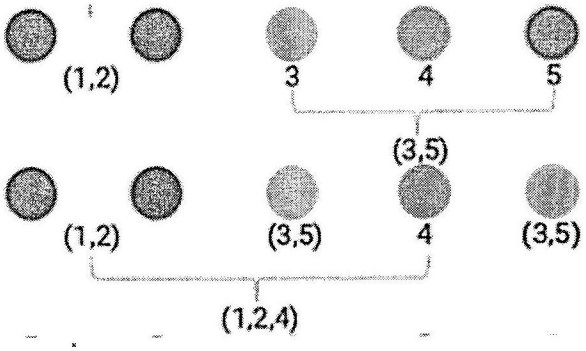
ตารางที่ 2-5 ผลลัพธ์การปรับปรุงตารางชุดข้อมูลเมื่อรวมคลัสเตอร์ (3,5)

Student_id	Marks
(1,2)	10
(3,5)	35
5	20

ตารางที่ 2-6 ปรับปรุงค่าภายในเมตริกซ์เมื่อรวมคลัสเตอร์ (3,5)

ID	(1,2)	(3,5)	4
(1,2)	0	25	10
(3,5)	25	0	15
4	10	15	0

ขั้นตอนที่ 5.2 ทำการรวมคลัสเตอร์ (1, 2, 4) และได้ผลลัพธ์ดังภาพที่ 2-9



ภาพที่ 2-9 การรวมคลัสเตอร์ (1, 2, 4)

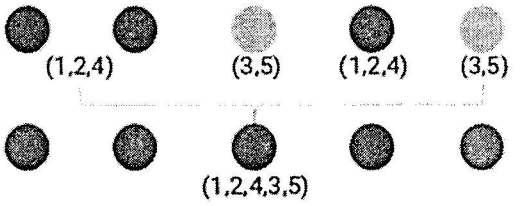
ตารางที่ 2-7 ผลลัพธ์การปรับปรุงตารางชุดข้อมูลเมื่อรวมคลัสเตอร์ (1, 2 ,4)

Student_id	Marks
(1,2,4)	20
(3,5)	35

ตารางที่ 2-8 ปรับปรุงค่าภายในเมตริกซ์เมื่อรวมคลัสเตอร์ (1,2,4)

ID	(1,2,4)	(3,5)
(1,2,4)	0	15
(3,5)	15	0

ขั้นตอนที่ 5.3 ทำการรวมคลัสเตอร์ (1, 2, 3, 4, 5) และได้ผลลัพธ์ดังภาพที่ 2-10

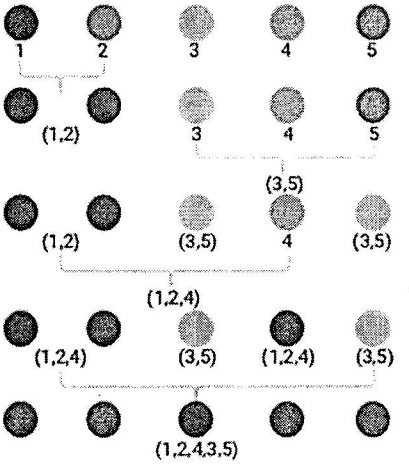


ภาพที่ 2-10 ทำการรวมคลัสเตอร์ (1,2,3,4,5)

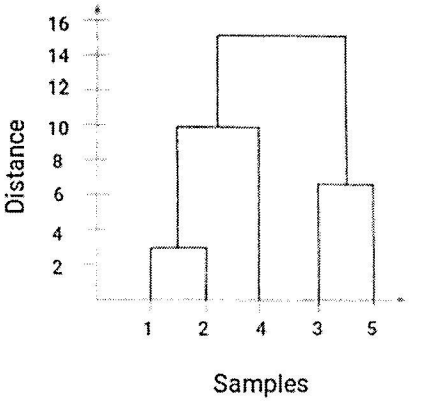
ตารางที่ 2-9 ผลลัพธ์การปรับปรุงตารางชุดข้อมูลเมื่อรวมคลัสเตอร์ (1,2,3,4,5)

Student_id	Marks
(1,2,3,4,5)	35

จากทุกขั้นตอนที่อธิบายมานี้สามารถสรุปได้ ดังภาพที่ 2-11



ภาพที่ 2-11 สรุปขั้นตอนการรวมคลัสเตอร์ในแต่ละขั้นตอน



ภาพที่ 2-12 กราฟเด็นโดรแกรมที่ได้จากวิธีการ Agglomerative Hierarchical Cluster

จากการแบ่งกลุ่มข้อมูลโดยใช้ Hierarchical Algorithm จะแสดงให้เห็นถึงความสัมพันธ์ของข้อมูลว่ามีข้อมูลใดมีความสัมพันธ์กันอย่างไร เหมาะกับการนำไปใช้วิเคราะห์ข้อมูลในกรณีที่ไม่ทราบจำนวนกลุ่มของข้อมูลที่ต้องการแบ่ง แต่เนื่องจากเป็นอัลกอริทึมที่ใช้เวลาในการทำคลัสเตอร์แต่ละกลุ่มนาน จึงไม่เหมาะกับข้อมูลที่มีปริมาณมาก สำหรับข้อมูลปริมาณมากจึงอาจต้องใช้ K-Means ในการจัดกลุ่มแทน

2.2.5.2 เทคนิคการหากฎความสัมพันธ์ (Association Rules)

กฎความสัมพันธ์เป็นวิธีการหนึ่งของการทำเหมืองข้อมูล โดยมีเป้าหมายเพื่อใช้วิเคราะห์หาความสัมพันธ์ที่ซ่อนอยู่ในกลุ่มข้อมูลขนาดใหญ่ ทั้งนี้เพื่อนำไปพยากรณ์เหตุการณ์ต่าง ๆ เช่น การซื้อสินค้าในซุปเปอร์มาเก็ต หรือ Market Basket Analysis ที่เป็นการค้นหาความสัมพันธ์ของรายการสินค้าที่ถูกซื้อ และมีประโยชน์ในการจัดเรียงสินค้าบนชั้นของร้านค้า เพื่ออำนวยความสะดวกให้กับลูกค้าได้ดียิ่งขึ้น รวมถึงสามารถนำไปใช้เป็นเครื่องมือในการส่งเสริมการขายสินค้า ในการหากฎความสัมพันธ์นี้มีอยู่หลายวิธี แต่สำหรับบทความนี้จะแสดงการหากฎความสัมพันธ์ด้วยวิธี Apriori ซึ่งจะมีอยู่ 2 ขั้นตอนใหญ่ ๆ คือ

5.2.5.2.1 การหา Frequent Itemset เป็นการหารูปแบบของข้อมูลที่เกิดขึ้นร่วมกันบ่อย ๆ ในฐานข้อมูลหรือ มากกว่าค่า Minimum Support ที่ผู้ใช้กำหนด ในขั้นตอนนี้จะแบ่งได้อีกเป็น 2 ขั้นตอนย่อย คือ

ก) การสร้างรูปแบบของ itemset (Join) จะใช้รูปแบบของ Itemset ที่มีค่ามากกว่า Minimum Support มาทำการสร้างรูปแบบของ Itemset ที่มีขนาดยาวมากขึ้นทีละหนึ่งขั้นไปเรื่อย ๆ

ข) การนับค่า Support (Count) หลังจากที่สร้างรูปแบบของ Itemset ได้แล้ว ขั้นถัดมาจะทำการคำนวณค่า Support ที่เกิดขึ้น โดยที่ Support คือ จำนวนเปอร์เซ็นต์ที่พบ Itemset ในฐานข้อมูล

5.2.5.2.2 การสร้าง Association Rule หลังจากที่หา Frequent Itemset ได้แล้วจะนำรูปแบบที่หาได้มาสร้างเป็นกฎความสัมพันธ์โดย เช่น Apple => Cereal หมายความว่าเมื่อลูกค้าซื้อ Apple แล้วลูกค่าจะซื้อ Cereal ร่วมไปด้วย

2.2.5.3 ตัวอย่างข้อมูลจากตารางที่ 2-10 เป็นข้อมูลที่เรียกว่า Transaction Database ในฐานข้อมูลนี้จะประกอบด้วย 4 Transaction และมีสินค้าที่ลูกค้าซื้อทั้งหมด 5 ประเภทคือ Apple, Beer, Cereal, Diapers และ Eggs การหากฎความสัมพันธ์ (Association Rules) สินค้าที่ลูกค้าซื้อ 2 ขั้นตอน ดังนี้

ตารางที่ 2-10 Transaction Database ที่จะใช้หากฎความสัมพันธ์

Transaction ID	Item
1	Apple, Cereal, Dispers
2	Beer, Cereal, Eggs
3	Apple, Beer, Cereal, Eggs
4	Beer, Eggs

2.2.5.3.1 การหา Frequent Itemset

ก) คำนวณหาค่า Support ของแต่ละ Item ดังแสดงในภาพถัดไป แต่ละแถวคือข้อมูลสินค้าแต่ละชนิดและแต่ละคอลัมน์มีความหมาย ดังนี้

(1) คอลัมน์ที่ 1 คือชื่อสินค้าหรือ Item หรือ Itemset

(2) คอลัมน์ที่ 2-5 คือหมายเลขของ Transaction ถ้ามีการซื้อสินค้าใน Transaction ใดที่คอลัมน์ของ Transaction นั้นจะมีหมายเลข 1 แต่ถ้าไม่มีจะเป็นเลข 0

(3) คอลัมน์ที่ 6 คือ ค่า Support

ตารางที่ 2-11 การคำนวณค่า Support ของสินค้าแต่ละชนิด

Items	Transaction ID				Support
	1	2	3	4	
Apple	1	0	1	0	2/4 = 50%
Beer	0	1	1	1	3/4 = 75%
Cereal	1	1	1	0	3/4 = 75%
Diapers	1	0	0	0	1/4 = 25%
Eggs	0	1	1	1	3/4 = 75%

จากค่า Support ที่คำนวณได้จะเห็นว่า Diapers มีค่า Support ต่ำกว่าค่า Minimum Support (25% < 50%) ดังนั้น Diapers จะถูกตัดออกและไม่นำไปพิจารณาสร้างเป็น Itemset ที่มีความยาว 2 ต่อไป ดังภาพถัดไปและเรียก Apple, Beer, Cereal และ Eggs ว่า Frequent Itemset

ตารางที่ 2-12 ทำการตัด Diapers เนื่องจากมีค่า Support ต่ำกว่าค่า Minimum Support

Items	Transaction ID				Support
	1	2	3	4	
Apple	1	0	1	0	2/4 = 50%
Beer	0	1	1	1	3/4 = 75%
Cereal	1	1	1	0	3/4 = 75%
~~Diapers~~	~~1~~	~~0~~	~~0~~	~~0~~	~~1/4 = 25%~~
Eggs	0	1	1	1	3/4 = 75%

ข) หา frequent Itemset ของ Itemset

(1) หา Frequent Itemset ของ Itemset ที่มีความยาว 2 โดยใช้การ Join ได้เป็น {Apple, Beer}, {Apple, Cereal}, {Apple, Eggs}, {Beer, Cereal}, {Cereal, Eggs} และเนื่องจากเป็นเซต ลำดับของข้อมูลไม่มีผล นั่นคือ {Apple, Beer} = {Beer, Apple} ส่วนการคำนวณค่า support ก็นำข้อมูล Transaction มาทำ Intersect กัน (หรือใช้ bit operation AND ก็ได้) เช่น การหา Support ของ {Apple, Beer} จะเกิดจาก

Transaction (Apple) = {1,0,1,0}

Transaction (Beer) = {0,1,1,1}

Transaction (Apple, Beer) = {0,0,1,0}

ค่า Support ของ Item Sets ที่มีความยาว 2 และ Itemset ที่มีค่า Support น้อยกว่า Minimum Support จะถูกตัดทิ้งไปทำให้เหลือ Frequent Itemset ที่มีความยาว 2 เพียงแค่ {Apple, Cereal}, {Beer, Cereal}, {Beer, Eggs}, {Cereal, Eggs}

ตารางที่ 2-13 ค่า Support ของ Itemset ที่มีความยาว 2

Items	Transaction ID				Support
	1	2	3	4	
{Apple, Beer}	0	0	1	0	1/4 = 25%
{Apple, Cereal}	1	0	1	0	2/4 = 50%
{Apple, Eggs}	0	0	1	0	1/4 = 25%
{Beer, Cereal}	0	1	1	0	2/4 = 50%
{Beer, Eggs}	0	1	1	1	3/4 = 75%
{Cereal, Eggs}	0	1	1	0	2/4 = 50%

ตารางที่ 2-14 ค่า Frequent Itemset ที่มีความยาว 2

Items	Transaction ID				Support
	1	2	3	4	
~~{Apple, Beer}~~	~~0~~	~~0~~	~~1~~	~~0~~	~~1/4 = 25%~~
{Apple, Cereal}	1	0	1	0	2/4 = 50%
~~{Apple, Eggs}~~	~~0~~	~~0~~	~~1~~	~~0~~	~~1/4 = 25%~~
{Beer, Cereal}	0	1	1	0	2/4 = 50%
{Beer, Eggs}	0	1	1	1	3/4 = 75%
{Cereal, Eggs}	0	1	1	0	2/4 = 50%

(2) หา Frequent Itemset ของ Itemset ที่มีความยาว 3 โดยมีเงื่อนไขว่าข้อมูล Item แรกจะต้องมีค่าเหมือนกันจึงสามารถทำการ Join กันได้ เช่น {Beer, Cereal, Eggs} เกิดจากการ Join {Beer, Cereal} และ {Beer, Eggs} แต่ไม่สามารถ Join {Apple, Cereal} และ {Beer, Cereal} ได้ ดังนั้นในขั้นตอนนี้จะมีเพียงแค่ Itemset เดียวคือ {Beer, Cereal, Eggs} และเป็น Frequent Itemset

ตารางที่ 2-15 ค่า Itemset ที่มีความยาวเท่ากับ 3

Items	Transaction ID				Support
	1	2	3	4	
{Beer, Cereaal, Eggs}	0	1	1	0	2/4=50%

จากตารางที่ 2-15 เราไม่สามารถสร้าง Itemset ที่มีความยาวมากขึ้นกว่านี้ได้อีก ขั้นตอนการหา Frequent Itemset จึงหยุด และได้ Frequent Itemset ทั้งหมดดังแสดงในตารางที่ 2-16

ตารางที่ 2-16 ค่า Frequent Itemset ทั้งหมด

Frequent itemset	Support	Size
{Apple}	2/4=50%	1
{Beer}	3/4=75%	1
{Cereal}	3/4=75%	1
{Eggs}	3/4=75%	1
{Apple, Cereal}	2/4=50%	2
{Beer, Cereal}	2/4=50%	2
{Beer, Eggs}	3/4=75%	2
{Cereal, Eggs}	2/4=50%	2
{Beer, Cereal, Eggs}	2/4=50%	3

2.2.5.3.2 การสร้างกฎความสัมพันธ์ (Association Rule)

การสร้างกฎความสัมพันธ์จาก Frequent Itemset ที่หาได้ โดยจะพิจารณา Frequent Itemset ที่มีความยาวมากกว่า 2 Item ขึ้นไปมาสร้างเป็นกฎความสัมพันธ์ เช่น {Apple, Cereal} จะสร้างกฎความสัมพันธ์ได้เป็น Apple => Cereal เป็นต้น โดยกฎความสัมพันธ์ทั่วไปจะแสดงเป็น

$$\text{LHS} \Rightarrow \text{RHS} \tag{2-2}$$

โดยที่ LHS (Left Hand Side) แสดงรูปแบบของ Itemset ด้านซ้ายของกฎความสัมพันธ์ และ RHS (Right Hand Side) แสดงรูปแบบของ Itemset ด้านขวาของกฎความสัมพันธ์ ซึ่งในการพิจารณาว่ากฎความสัมพันธ์ที่สร้างได้ดีหรือไม่จำเป็นจะต้องมีตัววัดประสิทธิภาพของกฎ ซึ่งได้แก่ Confidence และ Lift

ค่า Confidence แสดงความเชื่อมั่นของกฎความสัมพันธ์ที่เมื่อรูปแบบ LHS เกิดขึ้นแล้ว รูปแบบ RHS จะเกิดขึ้นด้วยเป็นจำนวนกี่เปอร์เซ็นต์ การคำนวณค่า Confidence หาได้จาก

$$\text{Confidence}(\text{LHS} \Rightarrow \text{RHS}) = \frac{support(LHS,RHS)}{support(LHS)} \tag{2-3}$$

โดยที่ Support (LHS, RHS) คือ ค่า Support ที่รูปแบบ LHS และ RHS เกิดขึ้นพร้อม ๆ กัน ตัวอย่างการหา Confidence ของกฎ Apple => Cereal แสดงได้ ดังนี้

$$\text{Confidence(Apple => Cereal)} = \frac{support(Apple,Cereal)}{support(Apple)}$$
$$= \frac{2/4}{2/4}$$
$$= 100\ \%$$

ค่า Lift คือค่าที่บ่งบอกว่าการเกิดรูปแบบ LHS และ RHS มีความสัมพันธ์กันแค่ไหน โดยถ้าค่า Lift เป็น 1 แสดงว่ารูปแบบ LHS และ RHS ไม่ขึ้นต่อกัน (Independent) ค่า Lift คำนวณได้จาก

$$\text{Lift(LHS => RHS)} = \frac{support(LHS,RHS)}{support(LHS)*support(RHS)} \quad (2\text{-}4)$$

ตัวอย่างของการหาค่า lift ของกฎ Apple => Cereal แสดงได้ ดังนี้

$$\text{Lift(LHS => RHS)} = \frac{support(Apple,Cereal)}{support(Apple)*support(Cereal)}$$
$$= \frac{2/4}{2/4*\ 3/4}$$
$$= 1.33$$

ค่า Confidence และ Lift ของกฎความสัมพันธ์ทั้งหมดที่สร้างได้แสดงดังตารางที่ 2-17 โดยทำการเรียงตามค่า Confidence และ Lift จากมากไปหาน้อย

ตารางที่ 2-17 กฎความสัมพันธ์ทั้งหมดที่สร้างได้พร้อมทั้งค่า Confidence และ Lift

NO	Frequent itemset	Confidence	Lift
1	Apple => Cereal	100%	1.33
2	Beer => Eggs	100%	1.33
3	Eggs => Beer	100%	1.33
4	Beer, Cereal => Eggs	100%	1.33
5	Cereal, Eggs => Beer	100%	1.33
6	Cereal => Apple	67%	1.33
7	Beer => Cereal, Eggs	67%	1.33
8	Eggs => Beer, Cereal	67%	1.33
9	Beer => Cereal	67%	1.33
10	Beer => Cereal	67%	1.33

ตารางที่ 2-17 (ต่อ)

NO	Frequent itemset	Confidence	Lift
11	Cereal => Eggs	67%	1.33
12	Eggs => Cereal	67%	1.33
13	Cereal => Beer, Eggs	67%	1.33
14	Beer,Eggs => Cereal	67%	1.33

2.2.5.4 การจำแนกประเภทข้อมูล (Classification)

การจำแนกข้อมูลเป็นเทคนิคหนึ่งของการทำเหมืองข้อมูลที่ต้องมีการเรียนรู้ข้อมูลในอดีต (Supervised Learning) เพื่อสร้างแบบจำลองสำหรับใช้ทดสอบกับข้อมูลใหม่ โดย (Rajashree., et al, 2011) กล่าวว่าการเรียนรู้ของเครื่อง (Machine Learning) โดยการทำเหมืองข้อมูลโดยส่วนใหญ่จะรองรับข้อมูลที่มีคุณลักษณะ (Feature) ที่เป็นลักษณะไม่ต่อเนื่อง (Discrete Data) เท่านั้น แต่ความเป็นจริงแล้วข้อมูลที่นำมาใช้จะมีคุณลักษณะเป็นข้อมูลแบบต่อเนื่อง (Continuous Data) ฉะนั้นจึงจำเป็นต้องมีการแบ่งข้อมูลแบบต่อเนื่องให้เป็นช่วง ๆ เพื่อทำให้ข้อมูลอยู่ในลักษณะไม่ต่อเนื่อง เพื่อจะได้นำไปใช้กับอัลกอริทึมที่รองรับข้อมูลเฉพาะชนิดไม่ต่อเนื่องได้ โดยเทคนิคการจำแนกประเภทข้อมูลมีหลายเทคนิค

2.2.5.4.1 เทคนิคต้นไม้ตัดสินใจ (Decision Tree)

ต้นไม้ตัดสินใจเป็นวิธีหนึ่งในการใช้เทคนิค Classification โดยต้นไม้ตัดสินใจเป็นแบบจำลองที่มีลักษณะเหมือนโครงสร้างของต้นไม้ที่มีลักษณะเป็นต้นไม้กลับหัว ซึ่งแต่ละโหนดจะแสดงคุณลักษณะ (Attribute) ที่ใช้ในการทดสอบ และกิ่งจะแสดงผลในการทดสอบและโหนดใบของต้นไม้จะแสดงคลาสหรือกฎในการแบ่งคลาสซึ่งเป็นผลลัพธ์ที่ได้จากการจำแนกข้อมูล (ชลนิศา, 2550) สูตรที่ใช้ในการคำนวณหาค่า Information Gain เริ่มจากการหาค่า Entropy ดังนี้

$$\text{Entropy(S)} = \sum_{i=0}^{c} -p_i log_2 P_i \qquad (2\text{-}5)$$

โดย

S หมายถึง attribute ที่นำมาวัดค่า Entropy

P_i หมายถึง สัดส่วนของจำนวนสมาชิกของกลุ่ม i กับจำนวนสมาชิกทั้งหมดของกลุ่มตัวอย่าง

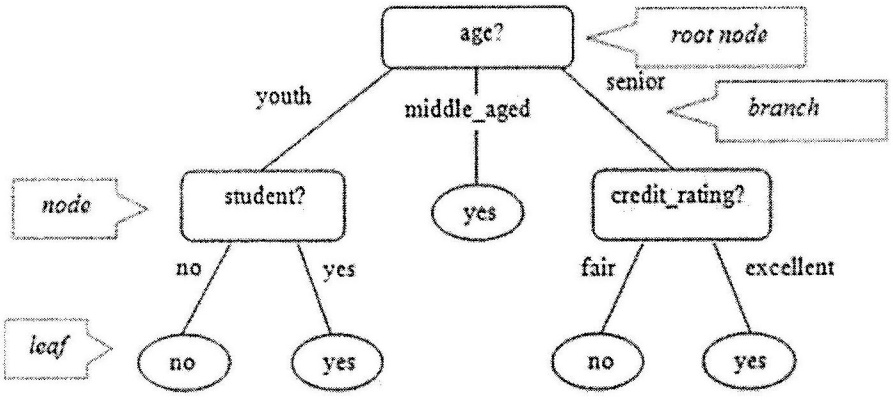
$$\text{Gain(S,A)} = \text{Entropy(S)} - \sum_{veValues(A)} \frac{Sv}{S} Entropy(Sv) \qquad (2\text{-}6)$$

โดย

A หมายถึง Attribute A

|Sv| หมายถึง สมาชิกของ Attribute A ที่มีค่า v

|S| หมายถึง จำนวนสมาชิกของกลุ่มตัวอย่าง



ภาพที่ 2-13 ต้นไม้ตัดสินใจ (Decision Tree)

ก) ส่วนประกอบต้นไม้ตัดสินใจ

(1) โหนดภายใน (Internal Node) หมายถึง คุณลักษณะต่าง ๆ ของข้อมูลใด ๆ ตกลงมาที่โหนด จะมีคุณลักษณะนี้เป็นตัวตัดสินใจว่าข้อมูลจะไปทิศทางใด โดยโหนดภายในที่มีจุดเริ่มต้นของต้นไม้ เรียกว่า โหนดราก

(2) กิ่ง (Branch) หมายถึง เป็นคุณลักษณะในโหนดภายในที่แตกกิ่งออกมา

(3) โหนดใบ (Leaf node) หมายถึง กลุ่มต่าง ๆ ซึ่งเป็นผลลัพธ์ในการจำแนกประเภทข้อมูล

ข) ปัญหา Over-Fitting เป็นปัญหาที่พบในการสร้างแบบจำลองต้นไม้ตัดสินใจ ลักษณะปัญหานี้คือแบบจำลองที่ได้มี Node หรือ Branch ที่มีความลึกทำให้แบบจำลองซับซ้อนเกินความจำเป็นหรือการที่จำนวนข้อมูลฝึกสอนมีน้อยเกินกว่าที่จะจัดกลุ่มได้ตามต้องการซึ่งอัลกอริทึม C4.5 ได้ใช้เทคนิค Pruning เพื่อลดขนาดและความซับซ้อนของแบบจำลองต้นไม้ตัดสินใจ

2.2.5.4.2 เทคนิคเรียนรู้แบบนาอีฟเบย์ (Naïve Bayes)

การเรียนรู้แบบเบย์เป็นเทคนิคที่ใช้ทฤษฎีความน่าจะเป็น (Probability) ตามทฤษฎีกฎของเบย์ (Bayes' Theorem) เพื่อหาว่าสมมติฐานใดน่าจะถูกต้องที่สุดโดยใช้ความรู้ก่อนหน้า (Prior Knowledge) เข้ามาช่วยในการเรียนรู้ได้ ได้แก่ ความน่าจะเป็นก่อนหน้าสำหรับสมมติฐานหนึ่ง ๆ ร่วมกับข้อมูลเพื่อหาสมมติฐานที่ดีที่สุด การเรียนรู้แบบเบย์อาศัยหลักการของการคำนวณความน่าจะเป็นของแต่ละสมมติฐาน โดยการเรียนรู้แบบเบย์เป็นการเรียนรู้เพิ่มได้ เนื่องจากตัวอย่างใหม่ที่ได้มาถูกนำมาปรับเปลี่ยนการแจกแจงซึ่งมีผลต่อการเพิ่ม หรือลดความน่าจะเป็นทำให้มีการเรียนรู้ที่เปลี่ยนไป วิธีการนี้ตัวแบบจะถูกปรับเปลี่ยนไปตามตัวอย่างใหม่ที่ได้โดยผนวกกับความรู้เดิมที่มี ซึ่งการทำนายค่าคลาสเป้าหมายของตัวอย่างใช้ความน่าจะเป็นมากที่สุดของทุกสมมติฐาน ซึ่งพบว่าวิธีนี้ให้ประสิทธิภาพในการเรียนรู้ได้ดีไม่ด้วยกว่าวิธีการเรียนรู้ประเภทอื่น ๆ

การจำแนกประเภทข้อมูลด้วยวิธี Naïve Bayes เป็นวิธีหนึ่งที่นิยมเนื่องจากการสร้างโมเดลไม่ซับซ้อน เหมาะกับกรณีของเซตตัวอย่างที่มีจำนวนมากและสมบัติ (Attribute) ของตัวอย่างไม่ขึ้นต่อกัน โดยเป็นการใช้ทฤษฎีความน่าจะเป็น (Probability) ดังสมการดังนี้

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2\text{-}7)$$

กำหนดให้

$P(A|B)$ หมายถึง ความน่าจะเป็นที่จะเกิดเหตุการณ์ B ขึ้นก่อนแล้วจะมีเหตุการณ์ A ตามมา

$P(A \cap B)$ หมายถึง ความน่าจะเป็นที่เกิดเหตุการณ์ A และ B พร้อมกัน

$P(B)$ หมายถึง ความน่าจะเป็นที่เหตุการณ์ B เกิดขึ้น

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (2\text{-}8)$$

กำหนดให้

$P(B|A)$ หมายถึง ความน่าจะเป็นที่จะเกิดเหตุการณ์ A ขึ้นก่อนแล้วจะมีเหตุการณ์ B ตามมา

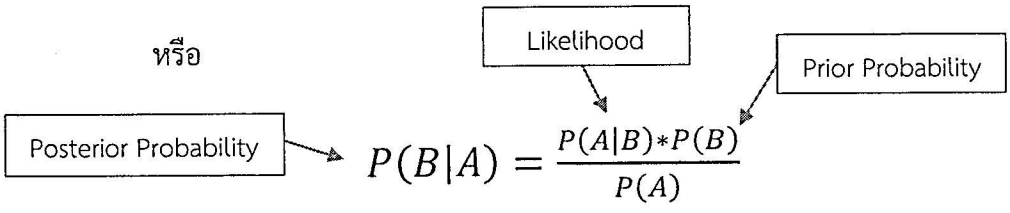
$P(A \cap B)$ หมายถึง ความน่าจะเป็นที่เกิดเหตุการณ์ A และ B พร้อมกัน

$P(A)$ หมายถึง ความน่าจะเป็นที่เหตุการณ์ A เกิดขึ้น

จากสมการที่ (2-7) และสมการที่ (2-8) ทั้งสองสมการมีค่าของ $P(A \cap B)$ เหมือนกัน ดังนั้นสามารถเขียนสมการของ $P(A \cap B)$ ได้ ดังนี้

$$P(A \cap B) = P(A|B) * P(B) = P(B|A) * P(A) \quad (2\text{-}9)$$

หรือ



(2-10)

รูปแบบของสมการ Bayes theorem แบ่งเป็น 3 กลุ่ม ดังนี้

Posterior Probability : ความน่าจะเป็นของข้อมูลที่มีแอททริบิวต์ A จะอยู่ในคลาส B

Likelihood : ความน่าจะเป็นของเทรนนิ่งดาต้าที่อยู่ในคลาส B และมีแอททริบิวต์ A

Prior Probability : P(A) ความน่าจะเป็นของคลาส A และ P(B) ความน่าจะเป็นของคลาส B

จากสมการที่ (2-10) จะพบว่านี่คือสมการของ Bayes theorem หรือทฤษฎีของเบย์ ซึ่งการหาความน่าจะเป็นของข้อมูลสามารถเปลี่ยน A และ B ให้เป็น A และ C โดย A หมายถึง แอททริบิวต์ และ C หมายถึง คลาส ดังสมการต่อไปนี้

$$P(C|A) = \frac{P(A|C)*\mathrm{P(C)}}{P(A)} \tag{2-11}$$

กำหนดให้

$P(C|A)$ หมายถึง ความน่าจะเป็นที่จะมีแอททริบิวต์เป็น A จะมีคลาส C

$P(A|C)$ หมายถึง ความน่าจะเป็นของเทรนนิ่งดาต้าที่มีคลาส C และมีแอททริบิวต์ A โดย $A = a_1 \cap a_2 \cap \ldots \cap a_M$ และ M คือ จำนวนแอททริบิวต์ในเทรนนิ่งดาต้า

$P(C)$ หมายถึง ความน่าจะเป็นของคลาส C

$P(A)$ หมายถึง ค่าความน่าจะเป็นของ $A = a_1 \cap a_2 \cap \ldots \cap a_M$

จากสมการที่ (2-5) สามารถเขียนในอีกรูปแบบหนึ่งโดยแทนค่าของแอททริบิวต์ A ตามจำนวนของแอททริบิวต์ ซึ่งจะได้สมการดังนี้

$$P(C|A) = \frac{P(a_1|C)*\,P(a_2|C)*\ldots..*P(a_M|C)\,*\mathrm{P(C)}}{P(A)} \tag{2-12}$$

ปัญหาที่มักพบโมเดล Naïve Bayes กรณีถ้าไม่มีรูปแบบของแอททริบิวต์เกิดขึ้นในข้อมูลของเทรนนิ่งดาต้า (Training Data) ค่าความน่าจะเป็นบางค่าเป็น 0 ดังนั้นการใช้งานโมเดล Naïve Bayes ที่มีความน่าจะเป็นเท่ากับ 0 จะทำให้ค่าที่พยากรณ์เกิดปัญหาเนื่องจากผลลัพธ์ที่ได้จะเท่ากับ 0 ตัวอย่างเช่น P(101501=2.5|GPAX=High) = 0/38 ผลลัพธ์ที่ได้จะเท่ากับ 0 ซึ่งปัญหาดังกล่าวสามารถแก้ไขได้โดยใช้เทคนิควิธีที่เรียกว่า Laplace Smoothing ซึ่งเป็นเทคนิคการปรับให้เรียบใช้กับปัญหาความน่าจะเป็นที่เป็นศูนย์ในเทคนิค Naïve Bayes ด้วยการเพิ่มความถี่ของข้อมูลเข้าไปอีก 1 ค่า ดังสมการที่ (2-13)

$$P(C|A) = \frac{number\ of\ A + \propto}{N + \propto * K} \tag{2-13}$$

เมื่อ

$\propto$ หมายถึง ค่าที่เพิ่มขึ้น ($\propto = 1$)

A หมายถึง จำนวนแอททริบิวต์ของ a_x ในคลาส C

K หมายถึง ค่าความเป็นไปได้ทั้งหมดของ A

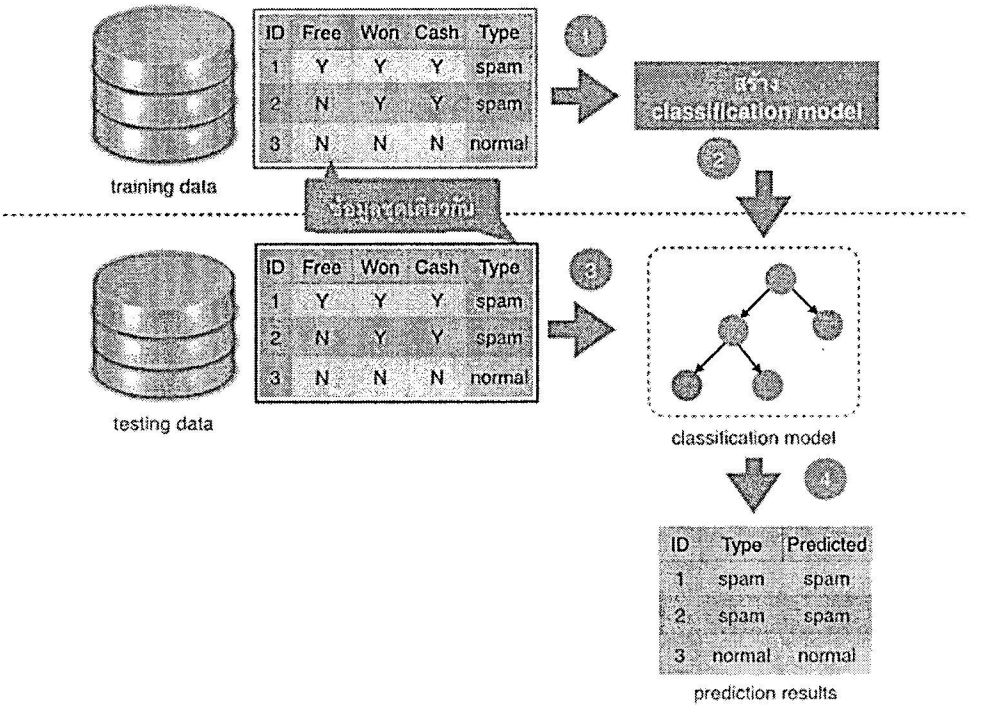
N หมายถึง จำนวนแอทริบิวต์ทั้งหมดของ A ในคลาส C

2.2.6 การทดสอบประสิทธิภาพโมเดลจำแนกประเภทข้อมูล

ขั้นตอนการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้าไมน์นิงก่อนที่จะนำผลลัพธ์ที่ได้ไปใช้งานต่อไป ก็จะต้องมีการวัดประสิทธิภาพของผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ที่ได้ตั้งไว้ในขั้นตอนแรกหรือมีความน่าเชื่อถือมากน้อยเพียงใด ซึ่งอาจจะย้อนกลับไปยังขั้นตอนก่อนหน้าเพื่อเปลี่ยนแปลงแก้ไขเพื่อให้ได้ผลลัพธ์ตามที่ต้องการได้ สำหรับการสร้างโมเดลด้วยเทคนิคจำแนกประเภทข้อมูล (Classification) มีการทดสอบประสิทธิภาพของโมเดลอยู่ 3 แบบ ได้แก่ 1) Self-Consistency Test 2) Split Test และ 3) Cross-Validation Test เป็นต้น

2.2.6.1 การวัดประสิทธิภาพโมเดลด้วยเทคนิค Self-Consistency Test

วิธี Self Consistency Test หรือบางครั้งเรียกว่า Use Training Set นี้เป็นวิธีการที่ง่ายที่สุด โดยข้อมูลที่ใช้ในการสร้างโมเดลและข้อมูลที่ใช้ในการทดสอบโมเดลเป็นข้อมูลชุดเดียวกัน กระบวนการนี้เริ่มจากสร้างโมเดลด้วยข้อมูลเทรนนิงดาต้า (Training Data) หลังจากนั้นนำโมเดลที่สร้างได้มาทำนายข้อมูลเทรนนิงดาต้าชุดเดิม ซึ่งวิธีการนี้เหมาะสำหรับใช้ในการทดสอบประสิทธิภาพเพื่อดูแนวโน้มของโมเดลที่สร้างขึ้น ถ้าได้ผลการวัดที่น้อยแสดงว่าโมเดลไม่เหมาะสมกับข้อมูลจึงไม่ควรจะนำไปทดสอบด้วยวิธีการแบ่งข้อมูลแบบต่าง ๆ

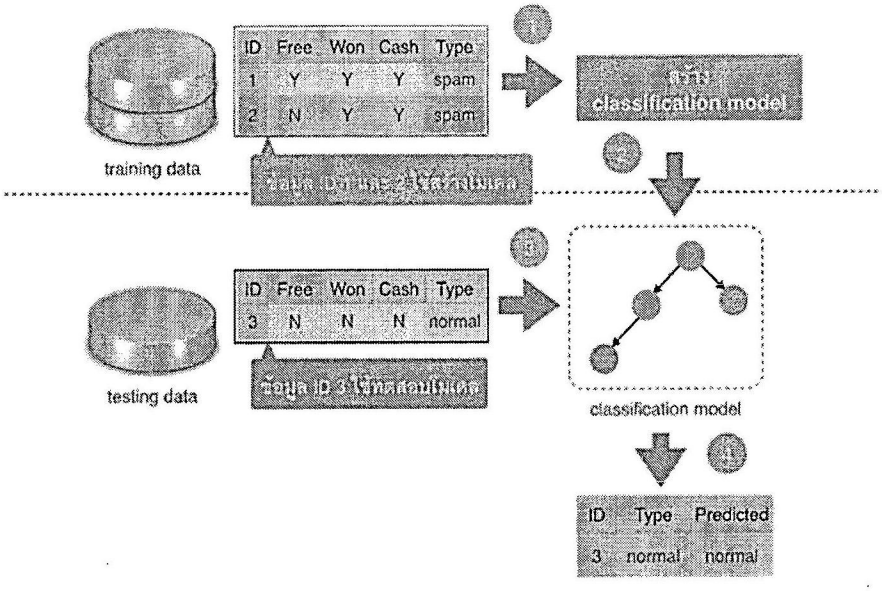


ภาพที่ 2-14 การวัดประสิทธิภาพด้วยเทคนิค Self-Consistency Test

2.2.6.2 การวัดประสิทธิภาพโมเดลด้วยเทคนิค Split Test

การวัดประสิทธิภาพแบบ Split Test เป็นการแบ่งข้อมูลด้วยการสุ่มออกเป็น 2 ชุด โดยข้อมูลชุดที่ 1 เป็น Training Data สำหรับสร้างโมเดล และข้อมูลชุดที่ 2 เป็น Testing Data สำหรับการ

ทดสอบประสิทธิภาพของโมเดล ตัวอย่างเช่น 70% ต่อ 30% หรือ 80% ต่อ 20% แต่การทดสอบแบบ Split Test นี้ทำการสุ่มข้อมูลเพียงครั้งเดียวซึ่งในบางครั้งถ้าการสุ่มข้อมูลที่ใช้ในการทดสอบที่มีลักษณะคล้ายกับข้อมูลที่ใช้สร้างโมเดลทำให้ผลการวัดประสิทธิภาพได้ออกมาดี ในทางตรงข้ามถ้าการสุ่มข้อมูลที่ใช้ในการทดสอบที่มีลักษณะแตกต่างกับข้อมูลที่ใช้สร้างโมเดลมากทำให้ผลการวัดประสิทธิภาพได้ออกมาแย่ ดังนั้นจึงควรใช้วิธี Split Test หรือทำการสุ่มหลาย ๆ ครั้ง แต่ข้อดีของวิธีการนี้คือใช้เวลาในการสร้างโมเดลน้อยซึ่งเหมาะกับชุดข้อมูลที่มีขนาดใหญ่มาก



ภาพที่ 2-15 การวัดประสิทธิภาพด้วยเทคนิค Split Test

2.2.6.3 การวัดประสิทธิภาพโมเดลด้วยเทคนิค Cross-validation Test

การวัดประสิทธิภาพแบบ Cross-validation Test เป็นวิธีที่นิยมในการทดสอบประสิทธิภาพของโมเดลเนื่องจากผลที่ได้มีความน่าเชื่อถือ การวัดประสิทธิภาพด้วยวิธี Cross-validation นี้จะทำการแบ่งข้อมูลออกเป็นหลายส่วนเท่า ๆ กัน (k-fold Cross-validation) เพื่อทำการทดสอบประกอบด้วยข้อมูลส่วน Training Set และ Testing Set ตัวอย่างเช่น วิธี 10-fold Cross-validation ซึ่งแบ่งข้อมูลออกเป็น 10 ส่วนเท่า ๆ กัน แล้วนำข้อมูลที่แบ่งไว้ใช้ในการสร้างตัวแบบพยากรณ์และการทดสอบตัวแบบพยากรณ์ เวียนกันไปจนครบทุกส่วนของข้อมูล ซึ่งกระบวนการทดสอบจะทำการให้ข้อมูลชุดที่ 1 เป็นข้อมูลทดสอบ (Testing Set) และข้อมูลชุดที่ 2-10 ชุดเทรนนิ่ง (Training Set) วนไปเรื่อย ๆ จนกระทั่งชุดที่ 10 เป็นชุดทดสอบ (Testing Set) และชุดที่ 1-9 เป็นชุดเทรนนิ่ง (Training Set) ดังภาพที่ 2-16

	Training Set									Testing Set
รอบที่1	2	3	4	5	6	7	8	9	10	1
รอบที่2	1	3	4	5	6	7	8	9	10	2
รอบที่3	1	2	4	5	6	7	8	9	10	3
รอบที่4	1	2	3	5	6	7	8	9	10	4
รอบที่5	1	2	3	4	6	7	8	9	10	5
รอบที่6	1	2	3	4	5	7	8	9	10	6
รอบที่7	1	2	3	4	5	6	8	9	10	7
รอบที่8	1	2	3	4	5	6	7	9	10	8
รอบที่9	1	2	3	4	5	6	7	8	10	9
รอบที่10	1	2	3	4	5	6	7	8	9	10

ภาพที่ 2-16 การแบ่งกลุ่มข้อมูลทดสอบ (10-fold Cross-validation)

2.3 ระบบสนับสนุนการตัดสินใจ

เทอร์บันและอรอนสัน (Turban and Aronson, 2001) ได้สรุปความหมายของระบบสนับสนุนการตัดสินใจ โดยยังคงรักษาจุดมุ่งหมายสำคัญ คือ เพื่อช่วยและปรับปรุงการตัดสินใจ ระบบสนับสนุนการตัดสินใจ คือ ระบบสารสนเทศที่ใช้คอมพิวเตอร์ (Computer-Based Information System : CBIS) ในการช่วยตัดสินใจ โดยระบบมีความสามารถเชิงโต้ตอบ ยืดหยุ่น ปรับใช้ โดยเฉพาะอย่างยิ่งพัฒนาเพื่อช่วยแก้ปัญหาที่มีลักษณะเฉพาะที่ไม่มีโครงสร้างของผู้บริหาร ระบบมีการใช้ข้อมูลทำให้ใช้งานง่ายและรวมเข้ากับลักษณะการทำงานของผู้ใช้ นอกจากนี้ระบบอาจใช้ตัวแบบ ซึ่งสร้างขึ้นมาเพื่อสามารถทางานในเชิงโต้ตอบ ระบบช่วยในทุกขั้นของการตัดสินใจและอาจรวมทั้งส่วนขององค์ความรู้

เทอร์บัน (2001) กล่าวว่า ระบบสนับสนุนการตัดสินใจ เป็นระบบสารสนเทศคอมพิวเตอร์เชิงโต้ตอบที่ยืดหยุ่นและปรับเปลี่ยนได้ ทำงานโดยการใช้กฎเกณฑ์ของการตัดสินใจที่อยู่ในลักษณะของแบบจำลองร่วมกับข้อมูลที่มีการจัดเก็บอยู่ในฐานข้อมูล พร้อมกับทัศนะของผู้ตัดสินใจในการที่จะหาแนวทางการตัดสินใจที่สามารถนำไปปฏิบัติเมื่อต้องการแก้ปัญหาที่ไม่อาจใช้วิธีการต่าง ๆ ที่กำหนดไว้ล่วงหน้าช่วยได้โดยตรงดังนั้นระบบสนับสนุนการตัดสินใจที่มีแบบจำลองซับซ้อนมากกว่าจะช่วยเพิ่มประสิทธิผลให้การตัดสินใจได้ดียิ่งขึ้น

2.3.1 ประเภทของสภาวการณ์ที่เกี่ยวข้องกับกับตัดสินใจ

ณัฏฐพร (2543) กล่าวว่า ประเภทของสภาวการณ์ที่เกี่ยวข้องกับกับตัดสินใจ มี 3 ประเภท คือ 1.การตัดสินใจภายใต้สภาวการณ์ที่แน่นอน (Certainty Condition) 2.การตัดสินใจภายใต้สภาวการณ์ที่มีความเสี่ยง (Risk Condition) 3.การตัดสินใจใต้สภาวการณ์ที่ไม่แน่นอน (Uncertainty Condition)

2.3.1.1 การตัดสินใจภายใต้สภาวการณ์ที่แน่นอน (Certainty Condition) เป็นการตัดสินใจที่ผู้ตัดสินใจมีความรู้ มีข้อมูลและสารสนเทศประกอบการตัดสินใจอย่างครบถ้วนสมบูรณ์และสามารถคาดการณ์ได้ว่าหากตัดสินใจดำเนินการในการแก้ไขปัญหาในแบบใดแบบหนึ่งไปแล้วจะได้ผลลัพธ์ออกมาอย่างไร สภาวการณ์เหล่านี้มักจะเกิดกับการตัดสินใจแก้ปัญหาแบบมีโครงสร้าง

2.3.1.2 การตัดสินใจภายใต้สภาวการณ์ที่มีความเสี่ยง (Risk Condition) เป็นการตัดสินใจที่ผู้ตัดสินใจมีความรู้มีข้อมูลหรือสารสนเทศประกอบการตัดสินใจเพียงบางส่วนไม่สามารถคาดการณ์ถึงผลของการตัดสินใจได้อย่างชัดเจนเพราะจะต้องพิจารณาถึงความเสี่ยงของทางเลือกในการตัดสินใจแต่ละทางเลือก ดังนั้นผู้ตัดสินใจจะต้องคาดการณ์และหาค่าความน่าจะเป็นของผลลัพธ์จากการตัดสินใจในแต่ละทางเลือกให้ครบถ้วนสมบูรณ์จึงจะตัดสินใจได้อย่างถูกต้องแม่นยำ

2.3.1.3 การตัดสินใจใต้สภาวการณ์ที่ไม่แน่นอน (Uncertainty Condition) เป็นการตัดสินใจที่ผู้ตัดสินใจไม่มีความรู้ไม่มีข้อมูลหรือสารสนเทศใด ๆ ประกอบการตัดสินใจ ทำให้ไม่ทราบเลยว่าโอกาสและความน่าจะเป็นที่จะเกิดผลลัพธ์ของการตัดสินใจในแต่ละทางเลือกเป็นอย่างไร โดยในสภาวการณ์แบบนี้มักจะเกิดกับการตัดสินใจในการแก้ไขปัญหาแบบไม่มีโครงสร้าง

2.3.2 กระบวนการการตัดสินใจ

ณัฏฐพร (2543) กล่าวว่า กระบวนการการตัดสินใจคือ การรับรู้ปัญหา การวิเคราะห์ปัญหา การกำหนดทางเลือก การประเมินและเปรียบเทียบทางเลือกและการเลือกทางเลือกไปปฏิบัติ รายละเอียดดังนี้

2.3.2.1 การรับรู้ปัญหา เป็นขั้นตอนในการเก็บรวบรวมข้อเท็จจริง ข้อมูลและสารสนเทศต่าง ๆ ที่เกี่ยวข้องกับปัญหาที่ต้องทำการตัดสินใจ

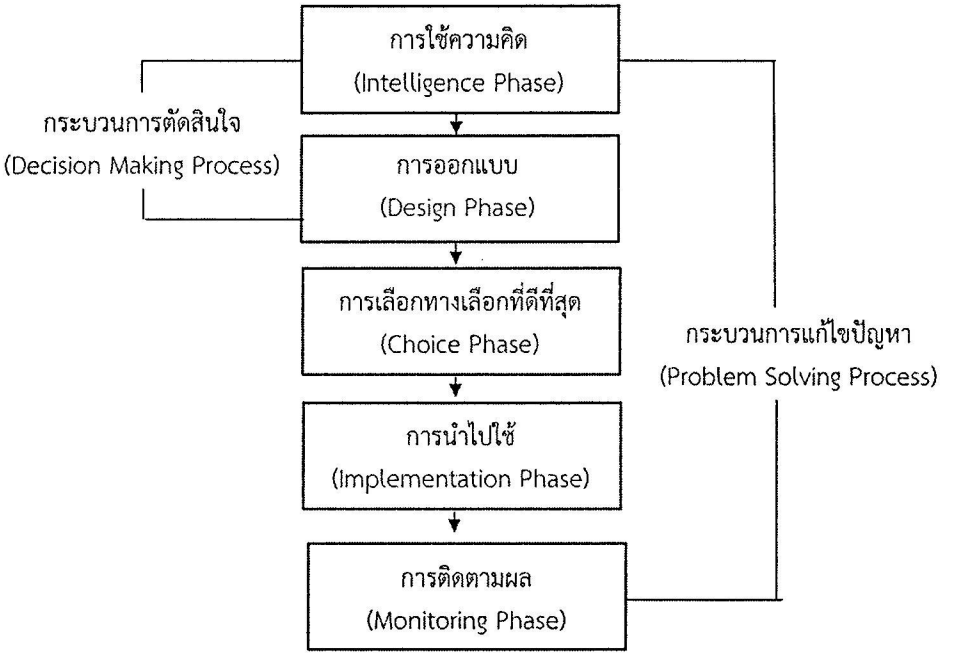
2.3.2.2 การวิเคราะห์ปัญหา เป็นขั้นตอนที่ผู้อำนาจในการตัดสินใจทำการวิเคราะห์หาสาเหตุของปัญหา และปัจจัยที่เกี่ยวข้องกับปัญหา จากข้อเท็จจริง ข้อมูลหรือสารสนเทศที่ได้มาจากขั้นตอนแรก

2.3.2.3 การกำหนดทางเลือก เป็นขั้นตอนที่ผู้มีอำนาจในการตัดสินใจพยายามค้นหาวิธีการหรือทางเลือกที่จะใช้ในการแก้ไขปัญหาตามสาเหตุของปัญหาและข้อเท็จจริง หรือข้อมูลที่ได้จากการวิเคราะห์มาจากขั้นตอนที่สอง

2.3.2.4 การประเมินและเปรียบเทียบทางเลือก เป็นขั้นตอนที่ผู้มีอำนาจในการตัดสินใจทำการพิจารณาหาวิธีการหรือทางเลือกในการแก้ปัญหาที่เหมาะสมที่สุดด้วยวิธีการประเมินและเปรียบเทียบถึงผลลัพธ์ที่ได้และข้อดีข้อเสียของแต่ละวิธีการหรือแต่ละทางเลือกในการแก้ไขปัญหาที่มีการกำหนดไว้

2.3.2.5 การเลือกทางเลือกไปปฏิบัติ เป็นขั้นตอนที่ผู้มีอำนาจในการตัดสินใจเลือกที่จะนำวิธีการหรือทางเลือกในการแก้ไขปัญหาที่คิดว่าเหมาะสมที่สุด ไปดำเนินการเพื่อแก้ไขปัญหาที่มีอยู่ให้ประสบความสำเร็จ

กิตติ (2550) กระบวนการตัดสินใจ (Decision Making Process) คือ การกำหนดขั้นตอนในการตัดสินใจแก้ปัญหาที่เกิดขึ้นภายในองค์กรอย่างมีหลักเกณฑ์ ด้วยการกำหนดขั้นตอนต่าง ๆ โดยผู้ที่กำหนดขั้นตอนคือ George Huber ได้นำมารวมเข้ากับกระบวนการแก้ปัญหา จึงทำให้กระบวนการตัดสินใจมีทั้งหมด 5 ขั้นตอน ดังภาพที่ 2-17

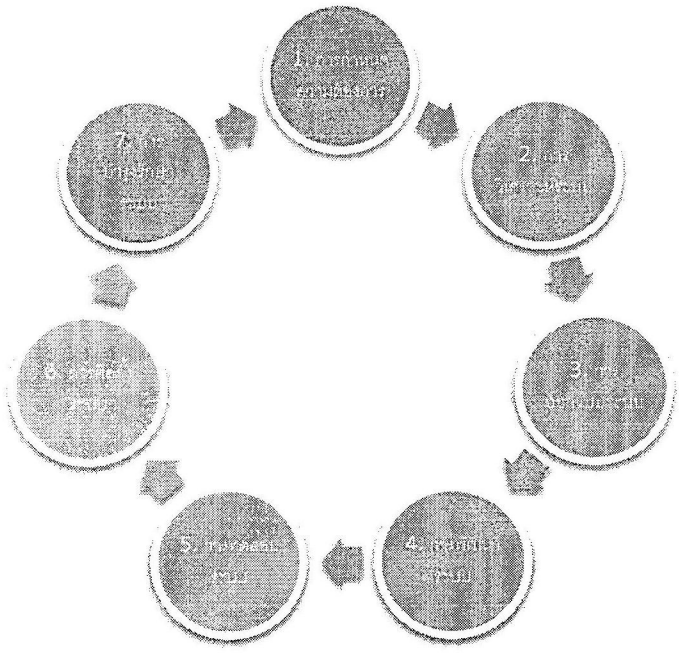


ภาพที่ 2-17 กระบวนการตัดสินใจและแก้ปัญหา (กิตติ, 2550)

ดังนั้นสรุปความหมายของระบบสนับสนุนการตัดสินใจ (Decision Support System : DSS) ได้ว่า คือ ระบบสารสนเทศที่สามารถโต้ตอบกับผู้ใช้โดยที่ระบบนี้จะรวบรวมข้อมูลที่มีการจัดเก็บอยู่ในฐานข้อมูล และแบบจำลองในการตัดสินใจที่สำคัญ เพื่อช่วยผู้บริหารในการตัดสินปัญหาแบบกึ่งโครงสร้างและไม่มีโครงสร้างเพื่อตัดสินใจแก้ปัญหาหรือวางแผนอย่างมีระบบ

2.4 ขั้นตอนการพัฒนาระบบ

การพัฒนาระบบมีขั้นตอนหลากหลายขั้นตอน เพื่อให้การพัฒนามีประสิทธิภาพและสำเร็จลุล่วงตามจุดประสงค์และระยะเวลาที่กำหนด จะมีการกำหนดขั้นตอนที่เรียกว่าวงจรการพัฒนาระบบ (System Development Life Cycle : SDLC) ซึ่งเป็นวิธีการพัฒนาระบบแบบดั้งเดิมและมีการกำหนดกรอบการทำงานที่ชัดเจน โดยมีลำดับขั้นตอนทั้งหมด 7 ขั้นตอนดังต่อไปนี้



ภาพที่ 2-18 วงจรพัฒนาระบบ (System Development life Cycle : SDLC)

2.4.1 การกำหนดความต้องการ (Requirement Definition)

เป็นขั้นตอนการค้นหาปัญหาและศึกษาทำความเข้าใจปัญหา ข้อเท็จจริงที่เกิดขึ้นจากการทำงานในระบบเดิม โดยนักวิเคราะห์ระบบจะต้องทำความเข้าใจปัญหาอย่างถ่องแท้ คิดหาแนวทางและวัตถุประสงค์ในการแก้ปัญหา ศึกษาความเป็นไปได้ในการแก้ปัญหา รวบรวมความต้องการและสรุปข้อกำหนดต่าง ๆ ให้ชัดเจน ถูกต้องและเป็นที่ยอมรับทั้งสองฝ่าย

2.4.2 การวิเคราะห์ระบบ (System Analysis)

เป็นขั้นตอนหลังจากผู้บริหารได้ตัดสินใจที่จะพัฒนาระบบหรือปรับปรุงระบบเดิม นักวิเคราะห์ระบบจะต้องดำเนินการวิเคราะห์ระบบงานปัจจุบัน (Current System) เพื่อนำมาพัฒนาแนวคิดสำหรับระบบใหม่ (New System) วัตถุประสงค์หลักในการวิเคราะห์ระบบคือจะต้องศึกษาและทำความเข้าใจในความต้องการต่าง ๆ ที่ได้รวบรวมมาจากขั้นตอนการกำหนดความต้องการโดยนักวิเคราะห์ระบบจะต้องนำข้อมูลความต้องการมาวิเคราะห์ เพื่อประเมินว่าควรมีอะไรบ้างที่ระบบใหม่ต้องดำเนินการ ด้วยการพัฒนาแบบจำลองลอจิคัล (Logical Model) ขึ้นมา ซึ่งได้แก่แบบจำลองแผนภาพกระแสข้อมูล (Data Flow Diagram) แบบจำลองกระบวนการ (Process Model) และแบบจำลองข้อมูล (Data Model) เป็นต้น

2.4.3 การออกแบบระบบ (System Design)

นักวิเคราะห์ระบบจำต้องออกแบบระบบที่จะพัฒนาให้สอดคล้องกับความต้องการที่ได้ระบุไว้ในเอกสารขั้นตอนการวิเคราะห์ ที่เป็นแบบจำลองเชิงตรรกะมาพัฒนาเป็นแบบจำลองเชิงกายภาพ โดยแบบจำลองเชิงตรรกะที่ได้จากขั้นตอนการวิเคราะห์ มุ่งเน้นว่ามีอะไรที่ต้องทำในระบบ ในขณะที่แบบจำลองเชิงกายภาพจะนำแบบจำลองเชิงตรรกะมาพัฒนาต่อด้วยการมุ่งเน้นว่าระบบจำดำเนินงาน

อย่างไร เพื่อให้เกิดผลตามความต้องการ โดยการออกแบบระบบจะประกอบด้วยการออกแบบที่เกี่ยวข้องกับฮาร์ดแวร์ ซอฟต์แวร์และระบบเครือข่าย การออกแบบรายงาน การออกแบบหน้าจอนำเข้าข้อมูล รออกแบบรูปแบบข้อมูลที่นำเข้าและรูปแบบการรับข้อมูล การออกแบบผังระบบงาน การออกแบบฐานข้อมูล การสร้างต้นแบบและการออกแบบโปรแกรม

2.4.4 การพัฒนาระบบ (System Development)

เป็นขั้นตอนที่มีการนำเอาระบบที่ได้ออกแบบไว้จากขั้นตอนออกแบบมาทบทวนเพื่อกำหนดการจัดทำซอฟต์แวร์ การออกแบบซอฟต์แวร์ การเขียนโปรแกรมและการทดสอบโปรแกรม ในกระบวนการนี้ทีมงานโปรแกรมเมอร์จะต้องพัฒนาโปรแกรมตามที่นักวิเคราะห์ระบบได้ออกแบบไว้ การเขียนชุดคำสั่งเพื่อสร้างระบบงานทางคอมพิวเตอร์ โดยโปรแกรมเมอร์สามารถนำเครื่องมือเข้ามาช่วยในการพัฒนาโปรแกรม เพื่อช่วยให้ระบบงานสามารถพัฒนาได้เร็วขึ้นและมีคุณภาพ และในกระบวนการนี้จะต้องทำเอกสารโปรแกรมควบคู่ไปกับการพัฒนาโปรแกรมเพื่อทำให้ง่ายต่อการตรวจสอบและแก้ไข ข้อกำหนดเกี่ยวกับการพัฒนาโปรแกรม ผู้บริหารขององค์กรจะเป็นผู้ตัดสินใจเลือกวิธีการพัฒนาโปรแกรมบางองค์กรอาจมีทีมงานพัฒนาโปรแกรมในองค์กร หรือซื้อซอฟต์แวร์สำเร็จรูปมาใช้หรือจ้างบริษัทรับพัฒนาระบบโดยเฉพาะ

2.4.5 การทดสอบระบบ (System Testing)

เมื่อโปรแกรมได้พัฒนาขึ้นมาแล้ว ยังไม่สามารถนำระบบไปใช้ไปใช้งานได้ทันที จำเป็นต้องดำเนินการทดสอบระบบก่อนที่จะนำระบบไปใช้งานจริง การทดสอบเบื้องต้นด้วยการสร้างข้อมูลจำลองขึ้นมาเพื่อตรวจสอบการทำงานของระบบงาน หากพบข้อผิดพรากก็ทำการปรับแก้ไขให้ถูกต้อง การทดสอบระบบจะมีการตรวจสอบไวยากรณ์ของภาษาที่ใช้และตรวจสอบว่าระบบทำงานตรงกับความต้องการของผู้ใช้หรือไม่

2.4.6 การติดตั้งระบบ (System Implement)

เมื่อทำการทดสอบระบบจนมั่นใจว่าระบบที่ได้รับการดสอบนั้นพร้อมที่จะนำไปติดตั้งเพื่อใช้งานบนสถานการณ์จริง จึงนำระบบไปติดตั้ง การติดตั้งระบบคือการเปลี่ยนการทำงานจากระบบเดิมไปเป็นระบบงานใหม่ แต่การเปลี่ยนแปลงไปสู่สิ่งใหม่ย่อมมีผลกระทบต่อผู้ใช้งานบางกลุ่ม ที่ยังคงความคุ้นเคยกับวิธีการดำเนินงานแบบเก่า รวมทั้งข้อจำกัดในเรื่องของความพร้อมในการเปลี่ยนแปลง ดังนั้นจึงควรเลือกแนวทางที่เหมาะสมในการติดตั้ง ซึ่งแบ่งออกเป็น 4 แนวทาง ดังนี้ 1) การติดตั้งแบบทันทีทันใด (Direct Installation) 2) การติดตั้งแบบขนาน (Parallel Installation) 3) การติดตั้งแบบนำร่อง (Single Location Installation/Pilot Installation) และ 4) การติดตั้งแบบทยอยติดตั้งเป็นระยะ (Phased Installation)

2.4.7 การบำรุงรักษาระบบ (System Maintenance)

หลังจากที่ระบบงานที่พัฒนาขึ้นมาได้ถูกนำไปใช้งาน หากพบข้อผิดพลาดหรือข้อบกพร่องจากการทำงานของระบบงาน นักวิเคราะห์ระบบจึงจำเป็นต้องดำเนินการติดตามและแก้ไขให้ถูกต้อง รวมถึงกรณีที่ข้อมูลจัดเก็บมีปริมาณมากขึ้น การขยายระบบเครือข่ายเพื่อรองรับเครื่องลูกข่ายที่มีจำนวนมากขึ้น บางกรณีอาจจำเป็นต้องเขียนโปรแกรมเพิ่มเติม หากผู้ใช้มีความต้องการเพิ่มขึ้น ดังนั้นในขั้นตอนของการกำหนดความต้องการนักวิเคราะห์ระบบจึงจำเป็นต้องมีการจัดทำเอกสารข้อตกลงร่วมกันทั้งสองฝ่ายถึงขอบเขตในการพัฒนาระบบงานและกรณีที่มีการแก้ไขหรือพัฒนาระบบงานเพิ่มเติม

2.5 การวิเคราะห์ข้อมูลด้วย RapidMiner Studio

โปรแกรม RapidMiner เป็นซอฟต์แวร์ Data Science ที่พัฒนาขึ้นโดยบริษัทที่ชื่อว่า Rapid-I ในประเทศเยอรมนีและเมื่อปลายปี 2013 ได้เปลี่ยนชื่อบริษัทจาก Rapid-I เป็น RapidMiner แทน โดยทำการย้ายสำนักงานใหญ่มาอยู่ประเทศสหรัฐอเมริกา โดย RapidMiner ใช้สำหรับการเตรียมข้อมูล การเรียนรู้เครื่อง การเรียนรู้ลึก การทำเหมืองข้อความ และการวิเคราะห์การทำนาย (Predictive Analysis) เป็นซอฟต์แวร์ที่ช่วยในการจัดส่งข้อมูลและลดข้อผิดพลาดจนแทบจะไม่จำเป็นต้องเขียนโค้ดเพิ่ม แต่ที่ทำให้เป็นเครื่องมือนัก Data Science นิยมเลือกใช้เป็นเพราะว่าตัว RapidMiner มีขั้นตอนพร้อมสำหรับการทำ Data mining (ขุดข้อมูล) และ Machine learning ซึ่งรวมไปถึงการโหลดและการแปลงข้อมูล (ETL) การประมวลผลล่วงหน้าและการวาดภาพจากข้อมูล การวิเคราะห์เชิงพยากรณ์และการสร้างแบบจำลองทางสถิติ การประเมินผลและการปรับใช้ต่าง ๆ ล้วนเป็นสิ่งที่เหล่า Data Scientist จำเป็นต้องทำในการเข้าใจข้อมูลมากขึ้น

โปรแกรม RapidMiner ใช้งานง่ายสำหรับผู้ใช้งานที่ไม่มีพื้นฐานเป็นเพราะว่า หน้าตาของซอฟต์แวร์เข้าใจง่าย และการทำงานนั้นเป็นแบบ Drag and Drop สำหรับการวิเคราะห์ในขั้นเบสิค RapidMiner ให้ยูสเซอร์ได้ออกแบบ Workflow ในการวิเคราะห์ข้อมูลในรูปแบบของภาพ ซึ่งทุกอย่างจะเกิดขึ้นในหน้า Design View

2.5.1 ข้อดีของโปรแกรม RapidMiner Studio 9

2.5.1.1 รองรับการใช้งานไฟล์ได้หลายประเภท เช่น ไฟล์ Excel 2007

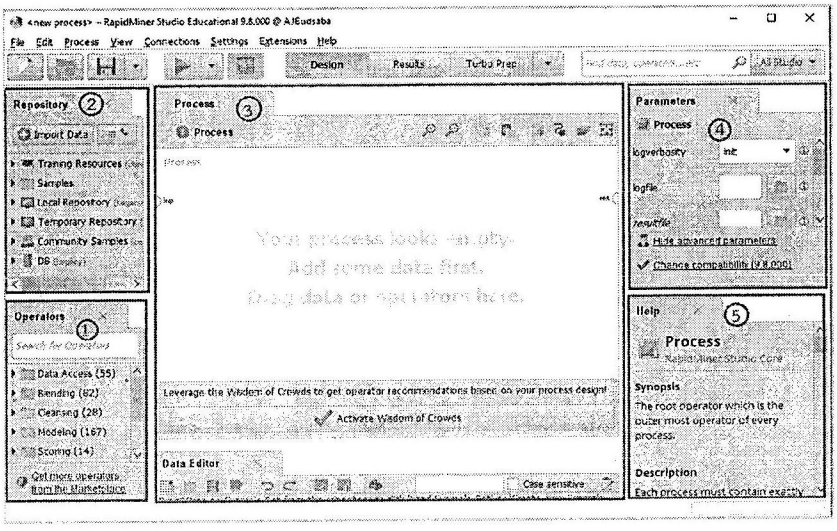
2.5.1.2 สามารถแสดงข้อมูลได้หลายรูปแบบ เช่น scatter plot 3D

2.5.1.3 สามารถแสดงผลโมเดลที่สวยงามและแก้ไขการแสดงผลให้สามารถอ่านได้ง่ายขึ้น

2.5.1.4 สามารถบันทึกไฟล์โมเดลออกเป็นไฟล์ภาพประเภทต่าง ๆ เช่น PNG, JPG หรือ PDF

2.5.1.5 มีวิธีการเตรียมข้อมูล (Preprocess) และการวิเคราะห์ได้หลากหลายรูปแบบ

2.5.2 ส่วนประกอบของหน้าต่าง Design ใน RapidMiner หน้า Design RapidMiner Studio 7 ประกอบด้วย 5 ส่วนหลัก ๆ ได้แก่ Operator, Repository, Process, Parameter และ Help



ภาพที่ 2-19 ส่วนประกอบหน้าต่าง Design ใน RapidMiner

2.5.2.1 Operators เป็นส่วนที่ใช้เก็บตัวโอเปอร์เรเตอร์ ที่ใช้ในการทำงานทั้งหมด ซึ่งจัดเป็นกลุ่ม ๆ โดยกลุ่มที่ใช้งานคล้ายคลึงกันจะจะอยู่ในกลุ่มเดียวกัน มี 8 กลุ่ม โอเปอร์เรเตอร์ ดังนี้

2.5.2.1.1 Data Access

2.5.2.1.2 Blending

2.5.2.1.3 Cleansing

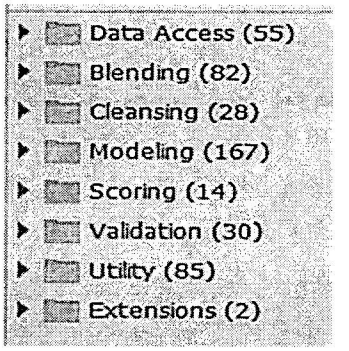
2.5.2.1.4 Modeling

2.5.2.1.5 Scoring

2.5.2.1.6 Validation

2.5.2.1.7 Utility

2.5.2.1.8 Extensions



ภาพที่ 2-20 Operators

2.5.2.2 Repository ส่วนนี้เป็นส่วนจัดการไฟล์ RapidMiner จะจัดการข้อมูลจาก 3 แหล่ง คือ DB (ดาต้าเบส) Local (ในเครื่องคอมพิวเตอร์ที่ใช้อยู่) และ Cloud Repository (ในคลาวด์) โดยเก็บไฟล์ Data Set และ Process ต่าง ๆ แยกเก็บคนละโฟลเดอร์กัน

2.5.2.3 Process เป็นหน้าหลักในการทำงานในการสร้างโปรเซสสำหรับทำ Machine Learning ของซอฟต์แวร์นี้ โดยจะนำโอเปรอเรเตอร์มาประกอบเพื่อสร้างโปรเซสขึ้นมาตามวัตถุประสงค์ของโจทย์ที่ตั้งไว้

2.5.2.4 Parameters เป็นส่วนที่กำหนดพารามิเตอร์ที่เป็นรายละเอียดของโอเปอร์เรเตอร์ที่เลือกใช้งาน เช่น โอเปอเรเตอร์ Set Role เป็น Operator ที่มีพารามิเตอร์ที่เกี่ยวข้องสองพารามิเตอร์ คือ แอทติบิวเนม ซึ่งเป็นพารามิเตอร์ที่ใช้กำหนดเอาท์พุทของโปรเซสว่าจะเลือกใหน ซึ่งจากตัวอย่างเลือก time_btw_stop เป็นเอาท์พุทของโปรเซสและพารามิเตอร์ Target Role เพื่อระบุว่าพารามิเตอร์เอาท์พุทเป็น label เท่านั้น

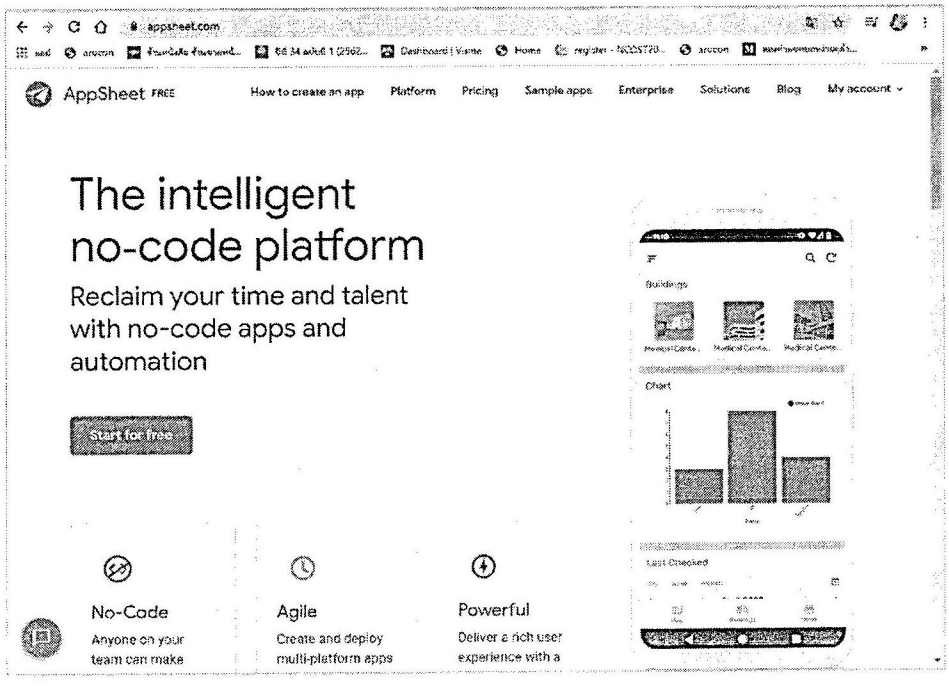
2.5.2.5 Help เป็นส่วนช่วยเหลือซึ่งจะแสดงรายละเอียดของตัวโอเปอเรเตอร์ที่เลือกใช้งานอยู่ ส่วนช่วยเหลือของ RapidMiner Studio จะบอกเพียงหน้าที่และรายละเอียดคร่าว ๆ ของโอเปอเรเตอร์หากต้องการดูรายละเอียดมากกว่านี้ต้องไปที่ Jum to Tutorai Process ซึ่งจะ Link ไปยังเวปไซต์ที่มีรายละเอียดของเกี่ยวกับ Operator ที่ใช้อยู่ เช่น โอเปอเรเตอร์ชื่อ Linear Regression ใน

หน้า Help ก็จะบอกว่าเป็น โอเปอเรเตอร์ ที่ใช้คำนวณข้อมูลจากจาก Data Set โดยใช้วิธี Linear Regression

2.6 การพัฒนาแอพพลิเคชั่นด้วย Appsheet

Appsheet เป็นเครื่องมือที่ช่วยในการสร้าง Mobile Application สำหรับใช้ในองค์กรโดยแทบไม่ต้องเขียนโค้ด ซึ่งมีฟังก์ชันการทำงานได้เทียบเท่าซอฟแวร์สำนักงานพวก Enterprise Resource Planning (ERP) ตั้งแต่การทำเอกสาร การทำบัญชี การทำระบบจองห้องประชุม การเก็บข้อมูลความพึงพอใจพนักงาน การตรวจวัดสต็อกสินค้า การติดตามยอดขายของฝ่ายแผนกการตลาด โดยสามารถเปิดใช้งานผ่าน Web Browser ได้ทั้งในสมาร์ทโฟนและคอมพิวเตอร์ตั้งโต๊ะ โดย Appsheet สามารถเชื่อมต่อฐานข้อมูลโดยใช้ Spread Sheet ได้ ซึ่งแตกต่างจากการพัฒนาแอปพลิเคชั่นแบบเดิมที่ต้องสร้างฐานข้อมูลเฉพาะ

โดยทั่ว ๆ ไป Appsheet สามารถสร้างแอพพลิเคชันสำหรับบันทึกข้อมูลทั่ว ๆ ไป โดยรองรับข้อมูลได้หลายแบบ เช่น ข้อมูลพิกัด Location ภาพถ่าย ลายเซ็น เวลาปัจจุบัน บาร์โค้ด QRcode, Barcode RFID และอื่น ๆ รองรับการทำ Automation เบื้องต้น เช่น การส่งอีเมล สร้างไฟล์เอกสารอัตโนมัติ หรือแม้แต่การแจ้งเตือนผ่าน Line และถ้าติดตั้งผ่าน Appsheet Engine จะทำให้สามารถรองรับการใช้งานแบบออฟไลน์และส่ง Push Notification ผ่านสมาร์ทโฟนละ Appsheet สามารถเชื่อมต่อ Good Sheet เข้ากับ Google Data Studio หรือโปรแกรมอื่น ๆ เพื่อทำรายงานหรือ Dashboard ได้



ภาพที่ 2-21 เครื่องมือพัฒนาแอพพลิเคชันด้วย Appsheet

2.6.1 ประโยชน์ของ Appsheet

2.6.1.1 เพิ่มประสิทธิภาพในการสื่อสารด้วย (Real-Time Collaboration)

Appsheet สามารถติดตามการทำงานแบบ Real-time โดยที่ทุกคนเห็นข้อมูลตรงกันได้ผ่านสมาร์ทโฟน อาทิเช่นมีบริษัทขนส่งที่นึงใช้ Appsheet เพื่อสร้างแอพพลิเคชันติดตามการทำงานของพนักงานขนส่งในระบบ (ในบริษัทนี้มีพนักงานขนส่งเกือบ 100 คน) โดยพนักงานขนส่งสามารถรายงานสถานะการส่งต่าง ๆ ผ่านสมาร์ทโฟน และหัวหน้างานก็สามารถดูรายงานผ่าน Dashboard ที่ห้องประชุมได้เลย

2.6.1.2 ลดข้อผิดพลาด และการทุจริตที่เกิดจากพนักงาน (Human Error and Fraud Prevention) Appsheet ไม่ต่างอะไรกับการคีย์ข้อมูลผ่าน Spread Sheet แต่ Appsheet สามารถคีย์ข้อมูลออนไลน์ได้ และมีช่องให้กรอกแบบแอพพลิเคชันแทนที่จะเป็นตาราง Spread Sheet เฉย ๆ

2.6.1.3 ลดงานซ้ำซ้อน (Duplicate Work) และการใช้กระดาษ (Paper Process) โดยเมื่อบันทึกข้อมูลใน Appsheet แล้วข้อมูลที่บันทึกสามารถปริ้นออกมาเป็นรายงานทำออกมาเป็นกราฟหรือแม้แต่ส่งเป็นอีเมลอัตโนมัติทุก ๆ ได้ ซึ่งเป็นฟีเจอร์ที่ช่วยลดงานซ้ำซ้อน

2.7 การสร้างแดชบอร์ดด้วย Google Data Studio

ข้อมูลเป็นสิ่งสำคัญและมีประโยชน์กับการแข่งขันทางธุรกิจ โดยทั่วไปจะมีการเก็บข้อมูลให้ได้มากที่สุด เพราะยิ่งข้อมูลมากก็จะทำให้เราตัดสินใจเรื่องต่าง ๆ ได้อย่างถูกต้องแม่นยำขึ้น แต่ในอดีตปัญหาอย่างหนึ่งของการนำเอาข้อมูลจำนวนมากวิเคราะห์นั้นไม่ใช่เรื่องง่ายหรือจะเอามาทำเป็น Dashboard ก็ไม่ใช่เรื่องง่ายอีกเช่นกัน ปัจจุบันมีเครื่องมือมากมายในการแปลงข้อมูลออกมาเป็นกราฟรูปแบบต่าง ๆ ที่ทำให้การอ่านทำความเข้าใจข้อมูลง่ายขึ้น Dashboard คือหน้ากระดานที่ใช้ในการสรุปข้อมูลแบบ Executive ในมุมมองต่าง ๆ เพื่อให้สามารถดูได้ง่าย ใช้เวลาในการตีความสั้นและสามารถตอบโจทย์ได้ ใช้ในการติดตามเรื่องที่สนใจ เพื่อเห็นการเปลี่ยนแปลงของข้อมูลตลอดเวลา ซึ่ง Google Data Studio เป็นเครื่องมือหนึ่งที่ฟรีใช้งานง่าย และสามารถต่อเข้ากับข้อมูลได้หลากหลายรูปแบบด้วยขั้นตอนที่ง่ายโดยเฉพาะคนที่ใช้เครื่องมือของ Google อยู่แล้ว ไม่ว่าจะเป็น Analytics, Google Ads, หรือ Google Search Console ยิ่งง่ายและสะดวกมาก ซึ่งการใช้ Data Studio นั้นช่วยลดเวลาการเตรียมข้อมูล และการทำพรีเซนต์ไปได้มาก

2.7.1 การใช้งาน Google Data Studio

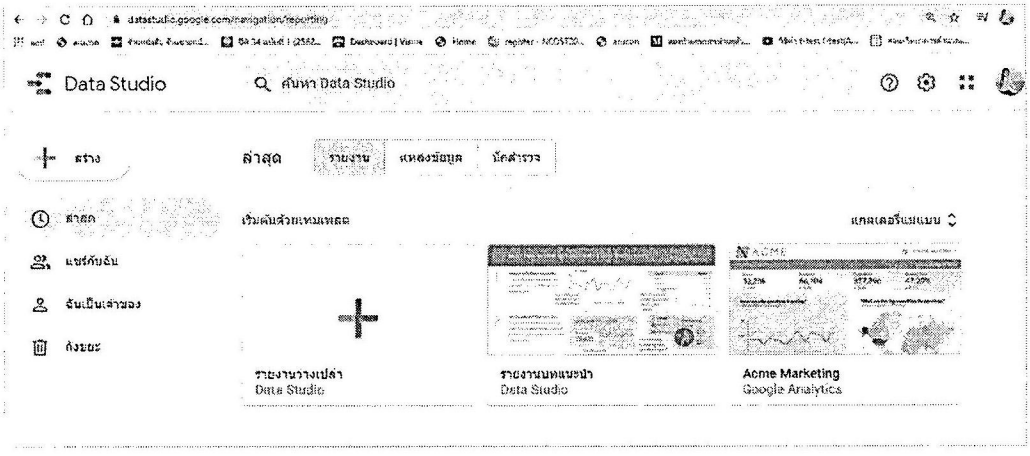
Data Studio เป็นเครื่องมีที่ให้ใช้งานฟรี และการเริ่มต้นใช้งานก็ง่ายมาก เพียงแต่เราจะต้องมี Google Account และ Data Source รวมกับความเข้าใจความหมายของข้อมูลที่เรามีได้แก่เรื่อง Dimensions และ Metrics ต่าง ๆ เพราะถ้าไม่เข้าใจ เราก็จะไม่รู้ว่าจะเอาข้อมูลชุดไหนอะไรออกมาเพื่อใช้งานตามที่เราต้องการ

Google Account เพื่อทำการ Log in เข้าไป datastudio.google.com

Data Source คือแหล่งข้อมูลที่จะให้ Data Studio ไปเชื่อมต่อและดึงข้อมูลออกมาเพื่อสร้าง Dashboard สำหรับ Data Source จะเป็น Google Analytics, Google Ads หรือ Google Search Console ก็ได้ ซึ่งถ้ายังไม่มี Data Source สามารถลง Google Analytics Demo Account ได้

2.7.2 เครื่องมือเบื้องต้นต่าง ๆ ใน Data Studio

เมื่อทำการ Log in เข้าที่ Data Studio จะเห็นหน้าตาของเครื่องมือตามภาพที่ 2-39 ซึ่งไม่ได้มีเมนูอะไรที่ซับซ้อนวุ่นวายชวนเวียนหัวเหมือนกับเครื่องมืออื่น ๆ



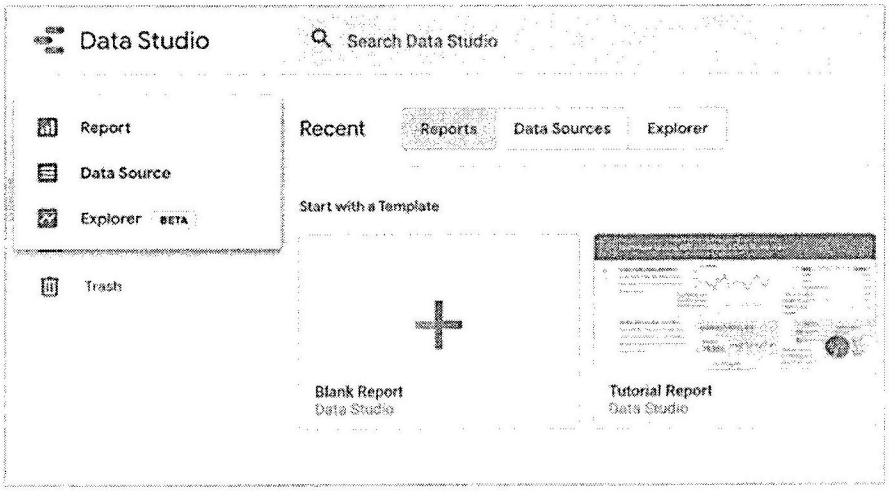
ภาพที่ 2-22 โปรแกรม Google Data Studio

เมนู Report โดยปกติค่าตั้งต้นการใช้งาน Data Studio จะตั้งต้นที่เมนูนี้ก่อนอยู่แล้ว ซึ่งจะแสดงลิสต์รีพอร์ททั้งหมดที่เราได้สร้างไว้ ตัวอย่างนี้เรายังไม่ได้สร้างรีพอร์ทอะไร เราก็จะเห็นเฉพาะ Untitled Report ซึ่งระบบสร้างรอเอาไว้ให้เรา ส่วนที่สองคือ Data Source ส่วนนี้จะเป็นส่วนสำหรับจัดการ Data Source เช่นการ Connect หรือการ Remove

2.7.3 การเชื่อมต่อกับ Data Source

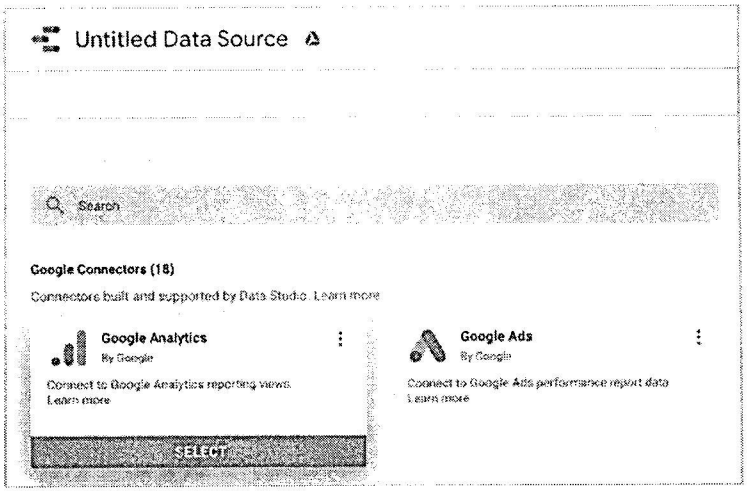
ผู้ใช้สามารถเชื่อต่อData Source กับ Google Analytics Demo Account หรือเชื่อมต่อกับ Google Analytics โดยมีขั้นตอนการเชื่อมต่อมีดังนี้

2.7.3.1 กด Create ที่ปุ่มด้านซ้ายบนของน้าจอ



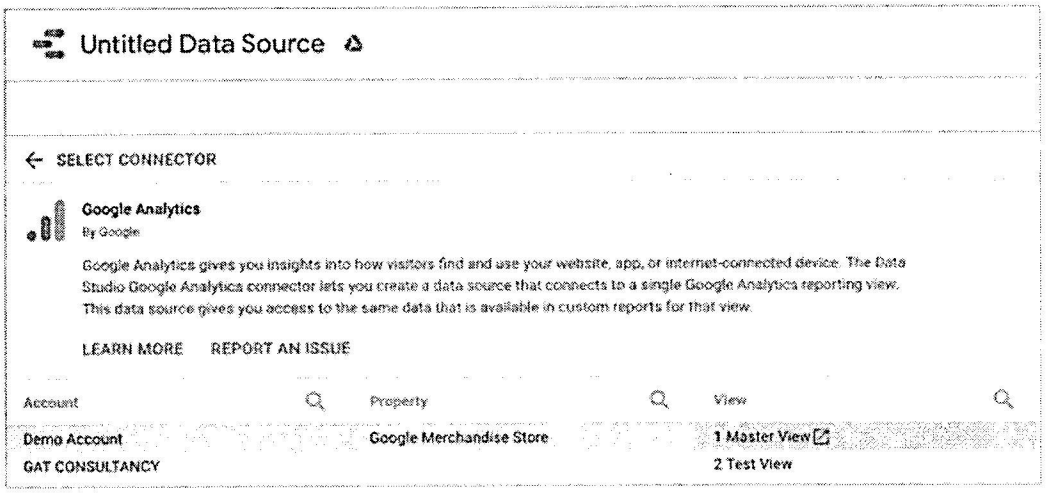
ภาพที่ 2-23 การเชื่อมต่อ Data Source

2.7.3.2 เลือก Google Analytics ซึ่งเป็น Data Source แรกที่แสดงอยู่ในลิสต์ Data Source ทั้งหมด



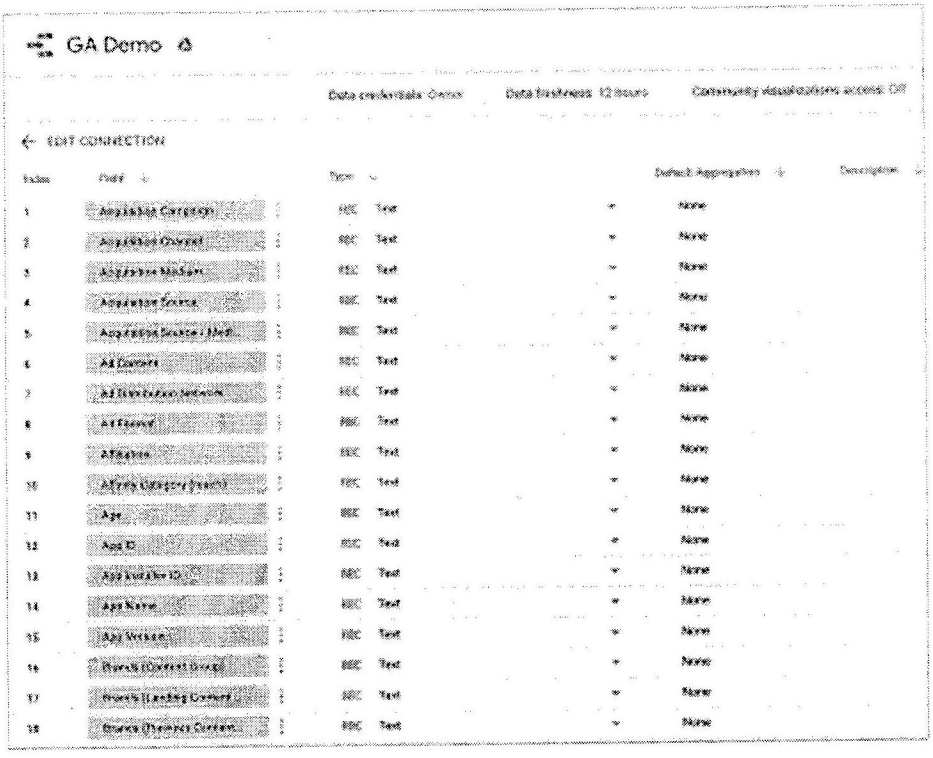
ภาพที่ 2-24 เลือก Google Analytic

2.7.3.3 เลือกที่ Demo Account>Google Merchandise Store>Master View แล้วกดปุ่ม Connect สีฟ้าที่ด้านขวาบน



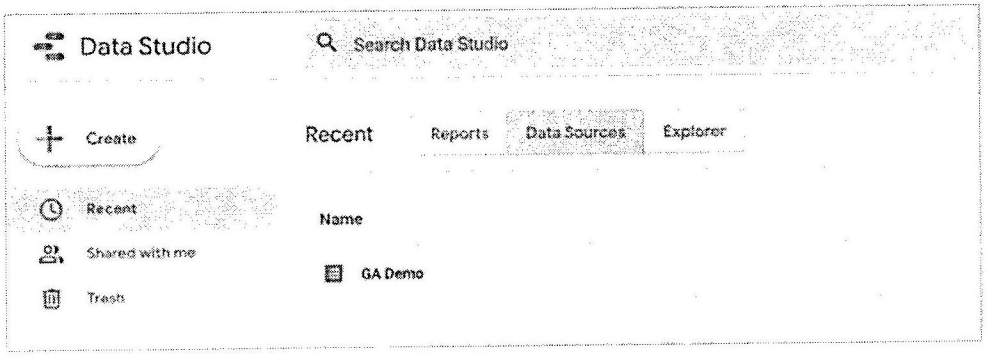
ภาพที่ 2-25 เชื่อมต่อ Google Analytics

2.7.3.4 หลังจาก Connect แล้ว จะเห็นลิสต์ของข้อมูล Dimension และ Metrics ทั้งหมดของ Google Analytics ซึ่งสามารถแก้ไข Data Source ด้านบนซ้าย



ภาพที่ 2-26 ผลการเชื่อมต่อ Google Analytics

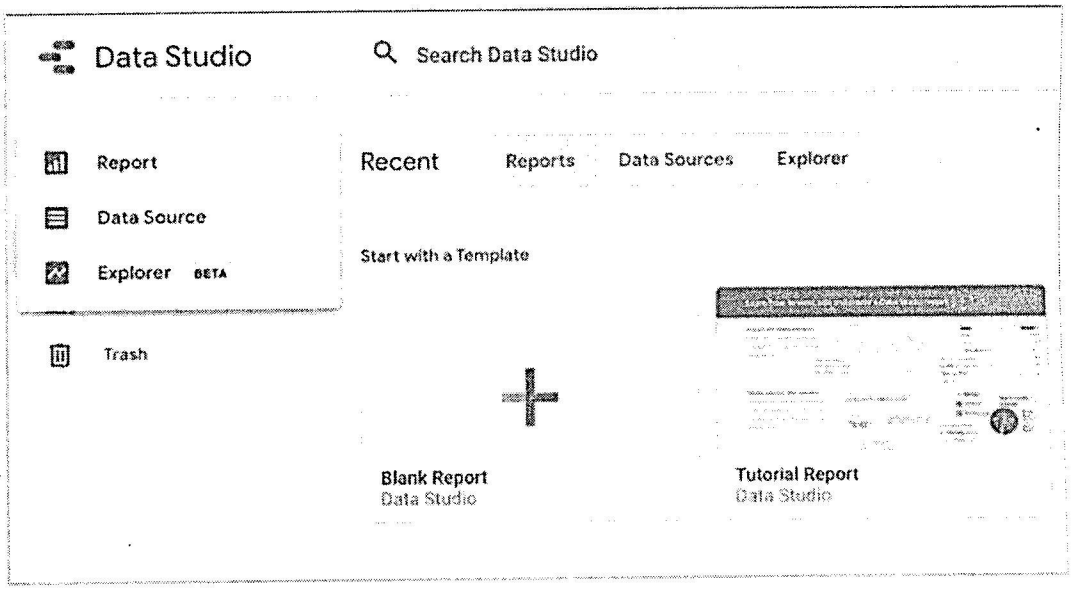
2.7.3.5 กด Create Report



ภาพที่ 2-27 Data Studio

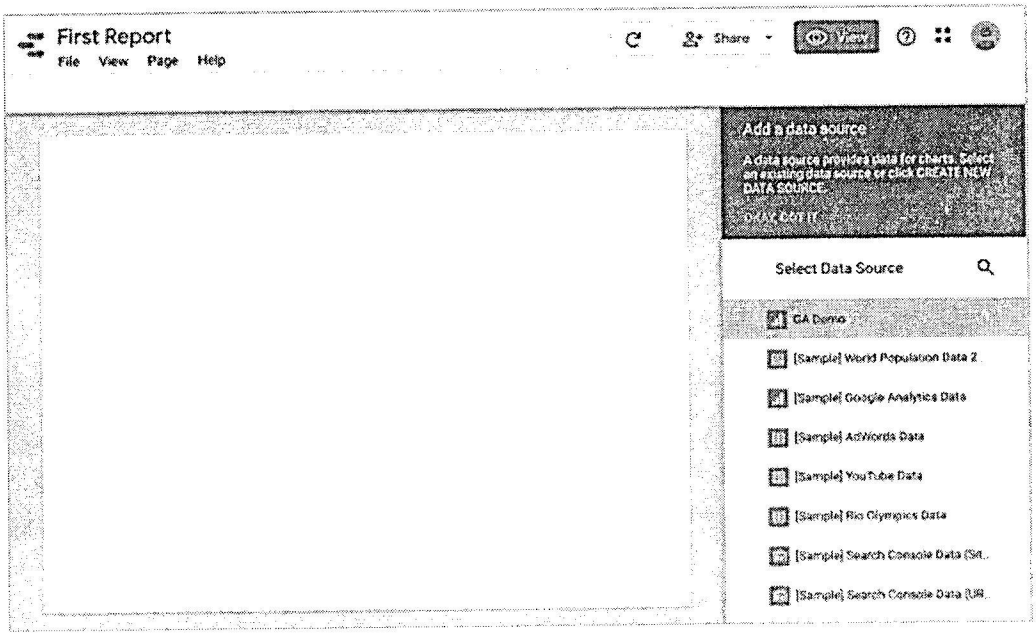
2.7.4 การสร้างรีพอร์ท

2.7.4.1 กด Create แล้วเลือกที่ Report



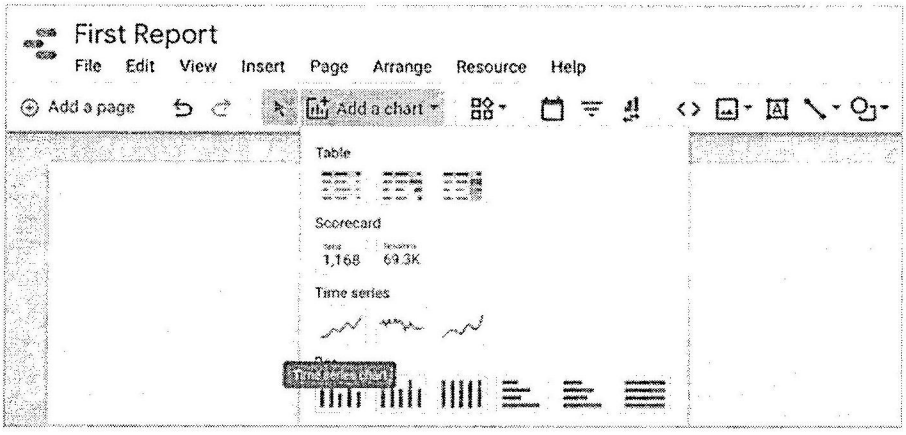
ภาพที่ 2-28 การสร้างรายงาน

2.7.4.2 เมื่อเข้ามาหน้าแรกของ Report ก็จะเห็นเป็นหน้าว่าง ๆ แต่จะเห็นลิสต์ของ Data Source อยู่ด้านขวามือ ตรงนี้ Data Source ที่เราเพิ่มมาก่อนหน้าจะแสดงอยู่ในลิสต์ด้วย ให้เลือก Data Source ที่เราสร้างไว้



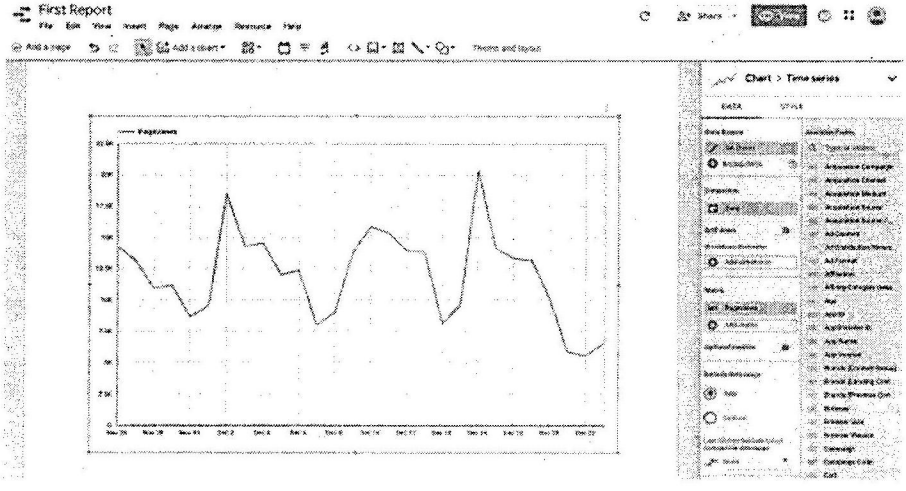
ภาพที่ 2-29 การเพิ่ม Data Source

2.7.4.3 กด Add a chart ที่เมนูด้านบน แล้วเลือก Time Series Chart อันแรก ทำการทดลองสร้าง Chart แรกด้วยการลากเพื่อกำหนดขนาดของ Chart



ภาพที่ 2-30 การเพิ่มรูปแบบของกราฟในรูปแบบต่าง ๆ

2.7.4.4 Data Studio จะสร้าง Chart พร้อมข้อมูลในแบบอัตโนมัติโดยเลือกชุดข้อมูลที่คิดว่าเหมาะสมที่สุดออกมาเป็นตัวอย่างให้ดูก่อน



ภาพที่ 2-31 รูปแบบของกราฟของ Data Studio

2.8 การจัดการฐานข้อมูลด้วย Google Sheet

Google Sheets ก็เป็น Apps ในกลุ่มของ Google Drive ซึ่งเป็นนวัตกรรมใหม่ของ Google มีลักษณะการทำงานคล้าย ๆ กับ Excel มีการสร้าง Column Row สามารถใส่ข้อมูลต่าง ๆ ลงไปใน Cell ได้ คำนวณสูตรต่าง ๆ ได้ แต่วิธีการใช้สูตรคำนวณจะแตกต่างจาก Excel ไม่ต้องติดตั้งที่เครื่อง สามารถใช้งานบน Web ได้ โดยไฟล์จะถูกบันทึกไว้ที่ Server ของ Google ทำให้สามารถเปิดใช้งาน

ได้ไม่ว่าจะอยู่ที่ใดเพียงมี Web browser และอินเตอร์เน็ต สามารถแชร์ไฟล์ให้ผู้อื่นร่วมใช้งานได้ และมีระบบ Real Time Save อัตโนมัติ นอกจากนี้ยังสามารถ Save หรือ Export ออกมาใช้งานกับ Excel ที่เครื่องของเราได้อีกด้วย ทำให้การทำงานสะดวกสบายมากยิ่งขึ้น โดยการล็อกอินเข้าใช้งานในเว็บไซต์ Google ด้วย Google Account หรือ gmail ก็สามารถเข้าไปทำงาน

2.8.1 คุณสมบัติของ Google Sheet

2.8.1.1 สร้างสเปรดชีตใหม่ เปิด และแก้ไขสเปรดชีตจากเครื่องใดก็ได้จากทางหน้าเว็บหรืออุปกรณ์อื่น

2.8.1.2 แชร์สเปรดชีต และทำงานร่วมกันกับคนอื่น ๆ บนสเปรดชีตเดียวกันและในเวลาเดียวกัน

2.8.1.3 ทำงานได้ทุกเวลา แม้ขณะที่ไม่มีอินเทอร์เน็ต

2.8.1.4 จัดรูปแบบเซลล์ ป้อน/จัดเรียงข้อมูล และการทำงานต่าง ๆ บนสเปรดชีต

2.8.1.5 ระบบคลาวด์ (Cloud) จะบันทึกการแก้ไขข้อมูลโดยอัตโนมัติ

2.9 งานวิจัยที่เกี่ยวข้อง

2.9.1 งานวิจัยในประเทศ

พงษ์เทพ และคณะ (2558) วิจัยเรื่องการใช้กฎความสัมพันธ์ร่วมกับฟัซซี่กฎความสัมพันธ์เพื่อคาดการณ์ผลการเรียนของนักศึกษา ในงานวิจัยนี้ใช้เทคนิคกฎความสัมพันธ์ (Association Rule Mining) เพื่อหาความสัมพันธ์ของผลการเรียนของแต่ละรายวิชา และใช้เทคนิคฟัซซี่กฎความสัมพันธ์ (Fuzzy Association Rule Mining) เพื่อคาดการณ์ผลการเรียนรวมโดยเฉลี่ยว่าอยู่ในระดับสูง ปานกลาง หรือต่ำ เพื่อนำผลไปคัดกรองนักศึกษาที่เรียนวิชาและได้เกรดตามกฎที่ค้นพบ ซึ่งคาดการณ์ว่าจะได้ผลการเรียนรวมโดยเฉลี่ยต่ำ ให้ปรับวิธีการเรียนตั้งแต่ต้น เพื่อที่จะมีผลการเรียนผ่านเกณฑ์เมื่อเรียนจนครบหลักสูตร

ณัฐธิดา และคณะ (2554) ได้ศึกษาวิจัยเรื่องการหาปัจจัยที่ส่งผลต่อความเสี่ยงของนักศึกษาเรียนอ่อนด้วยเทคนิคกฎความสัมพันธ์ กรณีศึกษา : มหาวิทยาลัยสงขลานครินทร์ โดยใช้เทคนิคกฎความสัมพันธ์ทำนายความเสี่ยงของการเรียนอ่อนของนักศึกษาชั้นปีที่ 1 ในคณะต่าง ๆ ของมหาวิทยาลัยสงขลา ซึ่งได้ใช้ฐานข้อมูลมหาวิทยาลัย ตั้งแต่ปี พ.ศ. 2550-2552 มาศึกษา ผลการวิจัยพบว่าปัจจัยที่ส่งผลต่อผลการศึกษาของนักศึกษาและคณะแตกต่างกันและความถูกต้องของผลลัพธ์ก็แตกต่างกัน เนื่องมาจากข้อมูลนำเข้าของนักศึกษาในแต่ละคณะต่างกัน ดังนั้นผลลัพธ์และถูกต้องของกฎความสัมพันธ์จึงขึ้นอยู่กับข้อมูลที่นำมาใช้ในการเรียนรู้ เมื่อมองในภาพรวมแล้วพบว่า ปัจจัยที่ส่งผลต่อผลการศึกษาของนักศึกษาเรียนอ่อน คือ วิธีการเข้าศึกษา คะแนนสอบเข้าศึกษาของวิชาต่าง ๆ และเพศของนักศึกษา จากผลการวิจัยนี้ได้พัฒนาโปรแกรมต้นแบบเพื่อทำนายความเสี่ยงต่อการเรียนอ่อนของนักศึกษาในแต่ละคณะ โดยใช้ปัจจัยของนักศึกษาแต่ละคณะมาทำการพัฒนา ผลของการวิจัยฉบับนี้สามารถนำไปใช้ในการทำนายความเสี่ยงที่จะเรียนอ่อนของนักศึกษาเพื่อช่วยป้องกันปัญหาทางผลการเรียนได้

ณรงค์ศักดิ์ และคณะ (2554) ได้ศึกษาวิจัยเรื่องการประยุกต์ใช้เอฟพี-กโรธกับงานแนะแนวการศึกษาต่อในระดับอุดมศึกษา โดยงานวิจัยได้ประยุกต์กฎความสัมพันธ์ซึ่งเป็นหนึ่งในเทคนิคของ

เหมืองข้อมูลเพื่อช่วยแนะแนวการศึกษาให้กับนักเรียนที่จะศึกษาต่อในระดับอุดมศึกษา โดยนำผลการเรียนเฉลี่ยของเจ็ดรายวิชาหลัก ได้แก่ วิชาคณิตศาสตร์ เคมี ชีววิทยา ฟิสิกส์ ภาษาไทย สังคมศึกษา และวิชาภาษาอังกฤษ ในระดับชั้นมัธยมศึกษาตอนปลายสายวิทยาศาสตร์-คณิตศาสตร์มาค้นหาความสัมพันธ์กับการศึกษาต่อของนักเรียน โดยอัลกอริทึมเอพี-กโรธ (FP-Growth Algorithm) เพื่อนำรูปแบบความสัมพันธ์เหล่านั้นมาแนะแนวการศึกษาให้กับนักเรียนที่จะศึกษาต่อในระดับอุดมศึกษา ซึ่งมีค่าเฉลี่ยความถูกต้องรวมเป็น 89.87%

ธนพัฒน์ (2563) ได้ศึกษาวิจัยเรื่องการพัฒนาโมเดลทำนายแผนการเรียนในการศึกษาต่อระดับมัธยมศึกษาตอนปลายของนักศึกษาโรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม : การประยุกต์ใช้เทคนิคเอ็นเซ็มเบิลโหวตร่วมระหว่างเครือข่ายในประสาท ซัพพอร์ตเวกเตอร์แมชชีน และต้นไม้ตัดสินใจ กลุ่มตัวอย่าง ได้แก่ นักเรียนรุ่นที่ 50 ถึงรุ่นที่ 53 จำนวน 787 คน ผลการวิจัยพบว่าโมเดล S-E ซึ่งเป็นโมเดลที่สามารถใช้ทำนายกลุ่มการเรียนที่เหมาะสมกับนักเรียน 2 กลุ่มการเรียนหลัก ได้แก่ กลุ่มการเรียนวิทย์ และกลุ่มการเรียนศิลป์ ที่ประสิทธิภาพการทำนายด้วยเทคนิคเอ็นเซ็มเบิลโหลตถูกต้องสูงที่สุดคิดเป็นร้อยละ 80.05 และผลการเรียนเฉลี่ยกลุ่มสาระการเรียนรู้คณิตศาสตร์ระดับมัธยมศึกษาตอนต้นมีความสำคัญและมีอิทธิพลต่อการเลือกแผนการเรียนเป็นอย่างมาก

นนทวัฒน์ และคณะ (2562) ได้ศึกษาวิจัยเรื่องการวิเคราะห์ปัจจัยที่มีผลต่อการพ้นสภาพของนักศึกษาระดับปริญญาตรีด้วยเทคนิคการทำเหมืองข้อมูล โดยทำการวิเคราะห์ข้อมูลเพื่อสร้างแบบจำลองโดยใช้เทคนิคจำแนกข้อมูลด้วยผังต้นไม้ตัดสินใจและใช้อัลกอริทึม J48 ในการทดสอบแบบจำลองด้วยวิธีการตรวจสอบไขว้ และวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งเป็นร้อยละโดยวิเคราะห์ข้อมูลจากข้อมูลนักศึกษาคณะวิทยาศาสตร์มหาวิทยาลัยราชภัฏบุรีรัมย์ที่เข้าศึกษาระหว่างปีการศึกษา พ.ศ. 2556-2559 จำนวน 3,604 ชุดข้อมูล ผลการวิเคราะห์ปัจจัยที่มีผลต่ออัตราการพ้นสภาพจำนวน 11 ปัจจัย คือ ปัจจัยที่เกี่ยวกับการเรียนในระดับมัธยมศึกษาและปัจจัยระหว่างเรียนในมหาวิทยาลัย โดยวัดค่าความถูกต้องได้ค่าร้อยละ 96.73 ความแม่นยำค่าร้อยละ 96.6 ค่าความระลึกได้ร้อยละ 96.7 และค่าความถ่วงดุลได้ร้อยละ 96.5

ซอและ และคณะ (2561) ได้ศึกษาวิจัยเรื่องการวิเคราะห์ปัจจัยที่มีผลต่อการพ้นสภาพของนักศึกษาโดยใช้เทคนิคเหมืองข้อมูล จากข้อมูลของนักศึกษาที่กำลังศึกษาและสำเร็จการศึกษาแล้วระหว่างปีการศึกษา 2555-2558 หลักสูตรวิทยาการคอมพิวเตอร์ จำนวน 97 ระเบียน และหลักสูตรเทคโนโลยีสารสนเทศ จำนวน 202 ระเบียน ประกอบด้วย 26 คุณลักษณะโดยใช้เทคนิคต้นไม้ตัดสินใจ เทคนิคเครือข่ายประสาทเทียบแบบย้อยกลับ และเทคนิคซัพพอร์ตเวกเตอร์แมชชีนสำหรับสร้างตัวแบบพยากรณ์และเปรียบเทียบตัวแบบด้วยการทดสอบประสิทธิภาพ 10-fold Cross Validation ผลการวิจัยพบว่า ข้อมูลภูมิหลังไม่ใช่ปัจจัยในการทำนายการพ้นสภาพของทั้ง 2 หลักสูตร โดยปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาเทคโนโลยีสารสนเทสมี 3 ปัจจัย ได้แก่ ผลการเรียนรายวิชาพื้นฐานทางฟิสิกส์ ผลการรเยนรายวิชาแพลตฟอร์มเทคโนโลยี และผลการเรียนเฉลี่ยปีที่ 2 ภาคเรียนที่ 1 และหลักสูตรวิทยาการคอมพิวเตอร์มีเพียง 1 ปัจจัย ได้แก่ ผลการเรียนรายวิชาโครงสร้างข้อมูล

ภัทรพงศ์ (2553) ได้ศึกษาวิจัยเรื่องการวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาระดับปริญญาตรีโดยใช้คอมมิตตีแมชชีน ซึ่งเป็นการนำเสนอการวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาปริญญาตรี โดยใช้คอมมิตตีแมชชีน ภายใต้หลักการทำงานของเหมืองข้อมูลโดยใช้

ชุดข้อมูลนักศึกษาที่เข้าศึกษาระหว่างปี พ.ศ. 2546-2549 ของมหาวิทยาลัยราชภัฏเลย จำนวน 11 แอทริบิวต์และ 12,865 ชุดข้อมูล ซึ่งคอมมิตตีแมชชีนเป็นกรทำงานระหว่าง SVM ร่วมกับ C5.0 โดยได้ทำการทดลองวัดประสิทธิภาพความถูกต้องเปรียบเทียบกับนิวรอลเน็ตเวิร์กและ C5:0 ซึ่งแบบจำลองแบบคอมมิตตีแมชชีน มีประสิทธิภาพในการจำแนกข้อมูลสูงที่สุด มีค่าเท่ากับ 75.32 และได้ทำการวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาระดับปริญญาตรี ผลการทดลองพบว่ามีปัจจัยที่มีความสำคัญ จำนวน 3 ปัจจัย คือ ขนาดโรงเรียนเดิม อาชีพของบิดาและอาชีพของมารดา และสามารถนำมาเขียนเป็นกฎการจำแนกสร้างความสัมพันธ์ของแอททริบิวต์ ซึ่งเทคนิคนี้สามารถประยุกต์ใช้กับการวิเคราะห์จำแนกข้อมูลที่มีความสัมพันธ์กับข้อมูลอื่น ๆ ได้เป็นอย่างดี

ชณิดาภา และคณะ (2561) ได้ศึกษาวิจัยเรื่องการวิเคราะห์การทำนายการลาออกกลางคันของนักศึกษาระดับปริญญาตรี โดยใช้เทคนิควิธีการทำเหมืองข้อมูล ซึ่งใช้ข้อมูลนักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏอุบลราชธานี ระหว่างปีการศึกษา 2558-2560 จำนวน 11 แอททริบิวต์ จำนวน 13,729 ชุดข้อมูล โดยผลการวิเคราะห์ค่าน้ำหนักด้วยวิธี Information Theory พบว่า 1) ปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา จำนวน 8 ปัจจัย และ 2) ผลการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลโมเดลที่สร้างด้วยเทคนิค Naïve Bayes มีประสิทธิภาพสูงสุดที่ความถูกต้องร้อยละ 93.58 เทคนิค Decision Tree ความถูกต้องร้อยละ 93.52 และเทคนิค K-Nearest Neighbors ความพถูกต้องร้อยละ 87.95 ตามลำดับ และมีปัจจัยที่เกี่ยวข้องสูงสุด 5 อันดับ ได้แก่ การกู้ยืมกองทุนเพื่อการศึกษา สาขาวิชา เกรดเฉลี่ย อาชีพของมารดา และอาชีพของบิดา

ชุติมา และคณะ (2553) ได้ศึกษาวิจัยเรื่องการพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา ซึ่งเป็นการวิจัยที่มีความมุ่งหมายเพื่อสร้างตัวแบบสำหรับหาปัจจัยที่ส่งผลต่อการตัดสินใจเลือกสาขาวิชาเรียนของนักศึกษา และพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจอย่างอัตโนมัติแบบออนไลน์สำหรับแนะแนวทางการเลือกสาขาวิชาเรียนของนักศึกษาในระดับอุดมศึกษา ตัวแบบเชิงความน่าจะเป็นได้สร้างขึ้นภายใจเทคนิคการจัดทำเหมืองข้อมูล ซึ่งในการทดสอบแบบจำลองที่ได้จะทำการทดสอบผลบนพื้นฐานของวิธี K-ford Cross Validation โดยใช้โปรแกรม Weka ในการสร้างแบบจำลอง หลังจากนั้นได้นำผลลัพธ์ที่ดีที่สุดจากเทคนิคข่ายงานเบย์ มาสร้างตัวแบบในขั้นสุดท้ายและได้เรียบเทียบความถูกต้องของตัวแบบกับผลการวิเคราะห์ความถดถอยเชิงพหุคูณ ในการวิจัยครั้งนี้ใช้ข้อมูลด้านการเรียนของกลุ่มตัวอย่างจากนักศึกษาในระดับอุดมศึกษาจากมหาวิทยาลัยทั้งภาครัฐบาลและเอกชนจำนวน 9 มหาวิทยาลัย ผลการวิจัยที่ได้แสดงให้เห็นว่า ตัวแบบความน่าจะเป็นโดยใช้เทคนิคการจัดทำเหมืองข้อมูลตามวิธีของข่ายงานเบย์ สามารถบ่งบอกตัวแปรสำคัญที่มีผลต่อการตัดสินใจในการเลือกสาขาวิชาเรียนของนักศึกษาในระดับอุดมศึกษาได้ และให้ค่าความแม่นยำในการทำนายสูงถึงร้อยละ 91.35 จากตัวแบบดังกล่าวทำให้ทราบว่าตัวแปรสำคัญที่มีต่อการตัดสินใจในการเลือกสาขาวิชาเรียนของนักศึกษาในระดับอุดมศึกษา ได้แก่ เกรียดเฉลี่ยวิชาคณิตศาสตร์ก่อนศึกษา เกรดเฉลี่ยวิชาคณิตศาสตร์ขณะกำลังศึกษา เกรดเฉลี่ยวิชาเขียนโปรแกรม ความรู้ในการพัฒนาระดับคอมพิวเตอร์และซอฟต์แวร์ ความรู้ด้านวิชาเรียน 1 ความรู้ด้านวิชากรเรียน 2 และความถนัด นอกจากนั้นตัวแปรรายวิชาที่ค้นพบที่มีผลต่อการเลือกสาขาวิชานั้น มีความสอดคล้องกับผลที่ได้จากการวิเคราะห์ความ

ถดถอยเชิงพหุคูณ จึงน่าเชื่อว่าตัวแปรที่ได้จากการพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจด้วยข่ายงานเบย์ดังกล่าวมีความน่าเชื่อถืออยู่ในระดับที่ยอมรับได้

ภรัณยา และคณะ (2563) ได้พัฒนาระบบวิเคราะห์ข้อมูลเศรษฐกิจชุมชนเพื่อสนับสนุนการตัดสินใจด้วยเทคนิคเหมืองข้อมูล วัตถุประสงค์เพื่อ 1) ออกแบบและพัฒนาระบบวิเคราะห์ข้อมูลเศรษฐกิจชุมชนด้วยเทคนิคเหมืองข้อมูล และ 2) ประเมินประสิทธิภาพการใช้งานและการยอมระบบระบบกลุ่มเป้าหมายของการวิจัย ได้แก่ ผู้เชี่ยวชาญด้านการพัฒนาระบบ จำนวน 3 คน และกลุ่มเจ้าหน้าที่องค์การบริหารส่วนตำบลจำนวน 15 คน เครื่องมือในการวิจัย ได้แก่ ระบบวิเคราะห์ข้อมูลเศรษฐกิจชุมชนด้วยเทคนิคเหมืองข้อมูล และแบบประเมินประสิทธิภาพของระบบ สถิติที่ใช้ ได้แก่ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน ผลการวิจัยพบว่า 1) ระบบวิเคราะห์ข้อมูลเศรษฐกิจชุมชนด้วยเทคนิคเหมืองข้อมูลแบ่งออกเป็น 4 ด้าน และด้านการวิเคราะห์ข้อมูลที่นำตัวแบบที่ได้จากอัคลอริทึม Random Tree ซึ่งประสิทธิภาพสูงสุดให้ค่าความถูกต้องร้อยละ 98.13 ค่าความแม่นยำเท่ากับร้อยละ 98.10 และ 2) ผลการประเมินประสิทธิภาพการใช้งานและการยอมรับระบบอยู่ในระดับมาก ($\bar{X}$=4.38, S.D=0.69)

2.9.2 งานวิจัยต่างประเทศ

Tekin (2014) ได้ศึกษาเรื่องการทำนายเกรดเฉลี่ยที่ของนักเรียนที่จบการศึกษา: ด้วยวิธีการทำเหมืองข้อมูล การศึกษานี้มีวัตถุประสงค์เพื่อใช้เทคนิคการทำนายด้วยการทำเหมืองข้อมูลเพื่อช่วยสถาบันการศึกษากับการคาดการณ์ผลการเรียน (GPAs) ของนักเรียนที่จบการศึกษา ถ้านักศึกษาคาดว่า จะมีผลการเรียน GPAs ต่ำ สามารถทำการปรับปรุงประสิทธิภาพการศึกษาได้ ผลการวิจัยวิเคราะห์เปรียบเทียบผลลัพธ์บ่งชี้ว่า เทคนิค SVM ผลการคาดการณ์ให้ผลความแม่นยำ 97.98% เครื่องจักรเรียนรู้เอ็กซ์ให้ผลแม่นยำ 94.92 ประเมินโดยใช้เกณฑ์ของสัมประสิทธิ์สหสัมพันธ์ที่โครงข่ายประสาทเทียมอัตราคาดเดาถูกต้อง 93.76%

Sen (2012) ได้ศึกษาเรื่องการคาดการณ์และวิเคราะห์การทดสอบความรู้ในระดับมัธยมศึกษาด้วยวิธีการทำเหมืองข้อมูล โดยทำการศึกษาปัจจัยที่นำไปสู่ความสำเร็จ (หรือความล้มเหลว) ของนักเรียนในการทดสอบเป็นปัญหาที่น่าสนใจและท้าทาย ตั้งแต่การทดสอบส่วนกลางและความก้าวหน้าทางวิชาการจะถือว่าเป็นแนวความคิดที่เกี่ยวข้องกับการวิเคราะห์ปัจจัยพื้นฐานของความสำเร็จ การทดสอบอาจช่วยให้เข้าใจและอาจปรับปรุงผลสัมฤทธิ์ทางการเรียน ในการศึกษานี้ใช้ข้อมูลจากการศึกษาในระดับมัธยม ในตุรกีพัฒนาแบบจำลองในการทำนายผลการทดสอบในระดับมัธยมศึกษาและการใช้การวิเคราะห์การคาดการณ์ที่สำคัญที่สุด ผลการศึกษาพบว่าวิธีการตัดสินใจ C5 ต้นไม้การตัดสินใจมีความแม่นยำ 95% และเครือข่ายประสาทเทียม (กับความถูกต้องของ 89%) ใน รูปแบบโลจิสติกแบบถดถอยความถูกต้องโดยรวมของ 82% การวิเคราะห์ก่อนหน้านี้ไม่ว่าจะเป็นนักเรียนที่มีทุนการศึกษา จำนวนนักเรียนปีก่อนหน้า เกรดเฉลี่ยอยู่ในกลุ่มพยากรณ์ที่สำคัญที่สุดคือคะแนนการทดสอบ

Oancea (2013) ได้ศึกษาเรื่องการคาดการณ์ผลการเรียนของนักเรียนในระดับอุดมศึกษาโดยใช้เครือข่ายประสาท ปัญหาสำคัญในการศึกษาจากปัญหาของนักเรียนหลังจากเข้าศึกษา นักเรียนหลายคนออกจากมหาวิทยาลัยจากหลากหลายเหตุผล : ความรู้เบื้องหลังที่ไม่ดีในด้านการศึกษา เกรดต่ำมาก และไม่สามารถผ่านการตรวจสอบ การขาดทรัพยากรทางการเงิน ผลคาดการณ์ของนักเรียน

เป็นเรื่องสำคัญสำหรับการบริหารของมหาวิทยาลัยที่ต้องการหลีกเลี่ยงปรากฏการณ์ออกกลางคัน โดยใช้เครือข่ายประสาทเทียมเพื่อทำนายผลการเรียนของนักเรียนปีที่ 1 เป้าหมายกลุ่มตัวอย่างคือ นักเรียน 1000 คน จาก "Nicolae Titulescu" มหาวิทยาลัยของบูคาเรสต์ จากข้อมูลบัณฑิตย้อนหลัง 3 รุ่น จำนวน 800 คน ใช้ 200 สำหรับกลุ่มทดลองเครือข่าย เครือข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น (MLP) ใช้สำหรับงานที่มีความซับซ้อน ข้อมูลป้อนเข้ามีประวัตินักเรียนเมื่อลงทะเบียนที่มหาวิทยาลัยรวมถึงข้อมูลเกี่ยวกับอายุนักเรียน GPA ที่โรงเรียนมัธยมศึกษา ช่องว่างระหว่างโรงเรียนมัธยมศึกษาและอุดมศึกษาที่ลงทะเบียน หลังจากฝึกอบรมเครือข่ายเรารับ Mean square error : MSE ประมาณ 1.7% ความสามารถในการทำนายผลของนักเรียนเป็นของช่วยเหลือที่ดีในการบริหารจัดการมหาวิทยาลัยเพื่อดำเนินการก่อนเพื่อหลีกเลี่ยงปรากฏการณ์การออกจากการศึกษา

บทที่ 3
วิธีการดำเนินการวิจัย

การวิจัยเพื่อพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี เป็นการวิจัยและพัฒนา (Research and Development) โดยดำเนินการวิจัยตามจุดประสงค์ของงานวิจัย ได้แก่ 1) เพื่อวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียน 2) เพื่อพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะ 3) เพื่อออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี 4) เพื่อพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี และ 5) เพื่อศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ซึ่งได้กำหนดแผนการวิจัยและขั้นตอนการดำเนินงาน ตลอดจนประชากรและกลุ่มตัวอย่างสำหรับการวิจัย ซึ่งแบ่งออกเป็น 5 ระยะ ดังนี้

3.1 ระยะที่ 1 การวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.2 ระยะที่ 2 พัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.3 ระยะที่ 3 ออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.4 ระยะที่ 4 พัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.5 ระยะที่ 5 ศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.1 ระยะที่ 1 การวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.1.1 วัตถุประสงค์การวิจัยระยะที่ 1

3.1.1.1 ศึกษาแนวคิด หลักการวิเคราะห์และออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา

3.1.1.2 เพื่อวิเคราะห์ปัจจัยการออกลางคันของนักศึกษา

3.1.1.3 เพื่อวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.1.2 ขอบเขตการวิจัยระยะที่ 1

3.1.2.1 การวิเคราะห์ปัจจัยการออกลางคันของนักศึกษา

3.1.2.2.1 ประชากร ได้แก่ นักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่เข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 3,913 คน

3.1.2.2.2 กลุ่มเป้าหมาย ได้แก่ นักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่เข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 3,913 คน

3.1.2.2 การวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.1.2.2.1 ประชากร ได้แก่ นักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี เข้าศึกษาในปีการศึกษา 2555-2557 ที่สำเร็จการศึกษา จำนวน 2,379 คน และนักศึกษาที่ลงทะเบียนเรียนในภาคการศึกษา 2/2563 จำนวน 3,579 คน

3.1.2.2.2 กลุ่มตัวอย่าง ได้แก่ นักศึกษาสาขาภาษาอังกฤษ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่เข้าศึกษาในปีการศึกษา 2555-2557 ซึ่งจบการศึกษาแล้ว จำนวน 141 คน ได้จากวิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling)

3.1.3 วิธีการดำเนินการวิจัยระยะที่ 1

3.1.3.1 ศึกษาแนวคิด หลักการวิเคราะห์และออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา

การศึกษา วิเคราะห์ และสังเคราะห์ปัจจัยการเรียนรู้ ผู้วิจัยได้ดำเนินการศึกษาวิเคราะห์สังเคราะห์จากเอกสาร บทความวิชากรและบทความวิจัยทั้งในประเทศและต่างประเทศที่เกี่ยวข้องโดยแสดงรายละเอียดดังตารางที่ 3-1

ตารางที่ 3-1 การสังเคราะห์การคัดเลือกปัจจัยการเรียนรู้

ปัจจัยการเรียนรู้	การเปรียบเทียบประสิทธิภาพการคัดเลือกคุณลักษณะสำหรับการจำแนกประเภทข้อมูล	การคัดเลือกคุณลักษณะเพื่อสร้างโมเดลสำหรับการพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วยเทคนิคเหมืองข้อมูล	การประยุกต์ใช้กฎความสัมพันธ์เพื่อวิเคราะห์ความเสี่ยงการออกกลางคันของนักศึกษาสาขาเทคโนโลยีสารสนเทศ	การพัฒนาขั้นตอนวิธีแบบผสมผสานสำหรับพยากรณ์ผลสัมฤทธิ์ทางการเรียน	การวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาที่มีต่อผลการเรียนปกติโดยใช้ต้นไม้ตัดสินใจ	การใช้กฎความสัมพันธ์ ร่วมกับทัชซึ่กฎความสัมพันธ์เพื่อคาดการณ์ผลการเรียนของนักศึกษา
เพศ	✓				✓	
อายุ	✓					
Curriculum type ประเภทหลักสูตร					✓	
Program สาขาวิชา	✓				✓	
จำนวนพี่น้อง	✓					
รูปแบบการเข้า	✓					
กิจกรรมเข้าแถว	✓					
กิจกรรมHome Room	✓					
กิจกรรมชมรม	✓					
ระบบการเรียน	✓					
สถานภาพบิดา	✓					
รายได้บิดา	✓				✓	
อาชีพบิดา	✓				✓	
สถานภาพมารดา	✓					

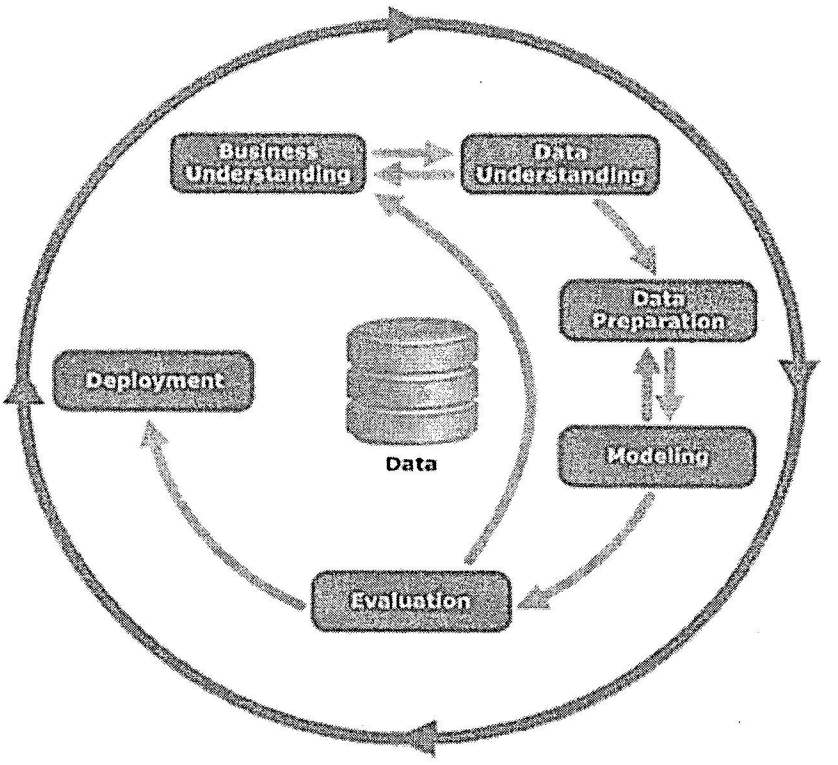
ตารางที่ 3-1 (ต่อ)

ปัจจัยการเรียนรู้	การเปรียบเทียบประสิทธิภาพการคัดเลือกคุณลักษณะสำหรับการจำแนกประเภทข้อมูล	การคัดเลือกคุณลักษณะเพื่อสร้างโมเดลสำหรับการพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วยเทคนิคเหมืองข้อมูล	การประยุกต์ใช้กฎความสัมพันธ์เพื่อวิเคราะห์ความเสี่ยงการออกกลางคันของนักศึกษาสาขาเทคโนโลยีสารสนเทศ	การพัฒนาขั้นตอนวิธีแบบผสมผสานสำหรับพยากรณ์ผลสัมฤทธิ์ทางการเรียน	การวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาที่มีต่อผลการเรียนปกติโดยใช้ต้นไม้ตัดสินใจ	การใช้กฎความสัมพันธ์ ร่วมกับฟัซซีกฎความสัมพันธ์เพื่อคาดการณ์ผลการเรียนของนักศึกษา
รายได้มารดา	✓				✓	
อาชีพมารดา	✓				✓	
สถานภาพผู้ปกครอง					✓	
ความพิการ	✓					
ความสามารถพิเศษ	✓					
วุฒิการศึกษาเดิม	✓				✓	
จังหวัด	✓		✓		✓	
ผลการเรียนแต่ละวิชา		✓ 16 วิชา	✓ 16 วิชา	✓		✓
เกรดเฉลี่ยสะสม	✓	✓			✓	✓
สถานะการออกกลางคัน					✓	

จากตารางที่ 3-1 การสังเคราะห์การคัดเลือกปัจจัยการเรียนรู้ พบว่า การคัดเลือกปัจจัยการเรียนรู้เพื่อนำปัจจัยไปทำการวิเคราะห์ข้อมูลใดข้อมูลหนึ่งจะขึ้นอยู่กับจุดประสงค์ของการวิเคราะห์ข้อมูลนั้น ๆ และจำนวนฐานข้อมูลที่ผู้วิจัยสามารถเข้าถึงได้

3.1.3.2 การวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษา

การดำเนินการวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษา ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลตามกระบวนการของ CRISP-DM ซึ่งมีกระบวนการดำเนินการดังนี้



ภาพที่ 3-1 การวิเคราะห์ข้อมูลตามกระบวนการของ CRISP-DM

3.1.3.2.1 ทำความเข้าใจปัญหา (Business Understanding) เป็นขั้นตอนที่เน้นไปที่การเข้าใจปัญหาและแปลงปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูลทางดาต้าไมน์นิง พร้อมทั้งวางแผนในการดำเนินการคร่าว ๆ ซึ่งผู้วิจัยทำการศึกษาฐานข้อมูลนักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่เข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 5 คณะ ได้แก่ คณะวิทยาศาสตร์และเทคโนโลยี คณะครุศาสตร์ คณะมนุษยศาสตร์และสังคมศาสตร์ คณะเทคโนโลยีอุตสาหกรรม และคณะวิทยาการจัดการ นักศึกษาทั้งหมด จำนวน 3,913 คน ประกอบด้วยข้อมูลนักศึกษาที่ออกกลางคัน จำนวน 1,291 คน และไม่ออกกลางคัน จำนวน 2,622 คน

3.1.3.2.2 การทำความเข้าใจข้อมูล (Data Understanding) ซึ่งผู้วิจัยได้ทำการศึกษาข้อมูลภูมิหลังของนักศึกษาข้อมูลผู้ปกครองข้อมูลเกรดเฉลี่ยและสถานะของนักศึกษาที่เข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 3,913 คน โดยแยกตัวแปรทั้งหมด 12 ตัวแปร ได้แก่ ตัวแปรต้น 11 ตัวแปร ตัวแปรตาม 1 ตัวแปร

3.1.3.2.3 การเตรียมข้อมูล (Data Preparation) ขั้นตอนนี้เป็นขั้นตอนที่ทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมา (Raw Data) เพื่อปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่เหมาะสมกลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการแปลงข้อมูลนี้อาจจะต้องมีการทำข้อมูลให้ถูกต้อง (Data Cleaning) เช่น การแปลงข้อมูลให้อยู่ในช่วง (Scale) เดียวกัน หรือการเติมข้อมูลที่ขาดหายไป เป็นต้น

ตารางที่ 3-2 แอททริบิวต์สำหรับการวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษาระดับปริญญาตรี

แอททริบิวต์	คำอธิบาย	รายละเอียด
Sex	เพศ	0=ชาย 1=หญิง
Old_ed	ข้อมูลการจบการศึกษาเดิม	11=ม.6 1=ปวช
Fat_status	สถานะบิดา	0=ถึงแก่กรรม 1=มีชีวิตอยู่ 9=ไม่ระบุ
Mor_status	สถานะมารดา	0=ถึงแก่กรรม 1=มีชีวิตอยู่ 9=ไม่ระบุ
Fat_reveue	รายได้ของบิดา	1 = น้อยกว่า 150,000 บาทต่อปี(น้อยกว่า 12,500 บาทต่อเดือน) 2 = 150,000 - 300,000 บาทต่อปี (12,500-25,000 บาทต่อเดือน) 3 = มากกว่า 300,000 บาทต่อปี (มากกว่า 25,000 บาทต่อเดือน) 0 = ไม่มีรายได้ 9 = ไม่ระบุ
Mot_reveue	รายได้ของมารดา	
Fat_occup	อาชีพของบิดา	00=ไม่ระบุ 01=รับราชการ 02=รัฐวิสาหกิจ 03=พนักงานหน่วยงานเอกชน 04=ค้าขาย,ธุรกิจส่วนตัว 05=เกษตร,ประมง 06=ไม่มีเงินได้ 07=อื่น ๆ 08=พนักงานราชการ
Mot_occup	อาชีพของมารดา	
Id_program	สาขาที่เรียน	101=การศึกษาปฐมวัย 137=ภาษาอังกฤษ 141=เทคโนโลยีการศึกษาฯ..... 547=การจัดการโลจิสติกส์

ตารางที่ 3-2 (ต่อ)

แอททริบิวต์	คำอธิบาย	รายละเอียด
Id_faculty	คณะที่เรียน	1= คณะวิทยาศาสตร์และเทคโนโลยี 2= คณะครุศาสตร์ 3= คณะมนุษยศาสตร์และสังคมศาสตร์ 4= คณะวิทยาการจัดการ 5= คณะเทคโนโลยีอุตสาหกรรม
GPA	เกรดเฉลี่ย	0.00-4.00
Status	สถานะการออกกลางคัน	Yes=ออก No=ไม่ออก

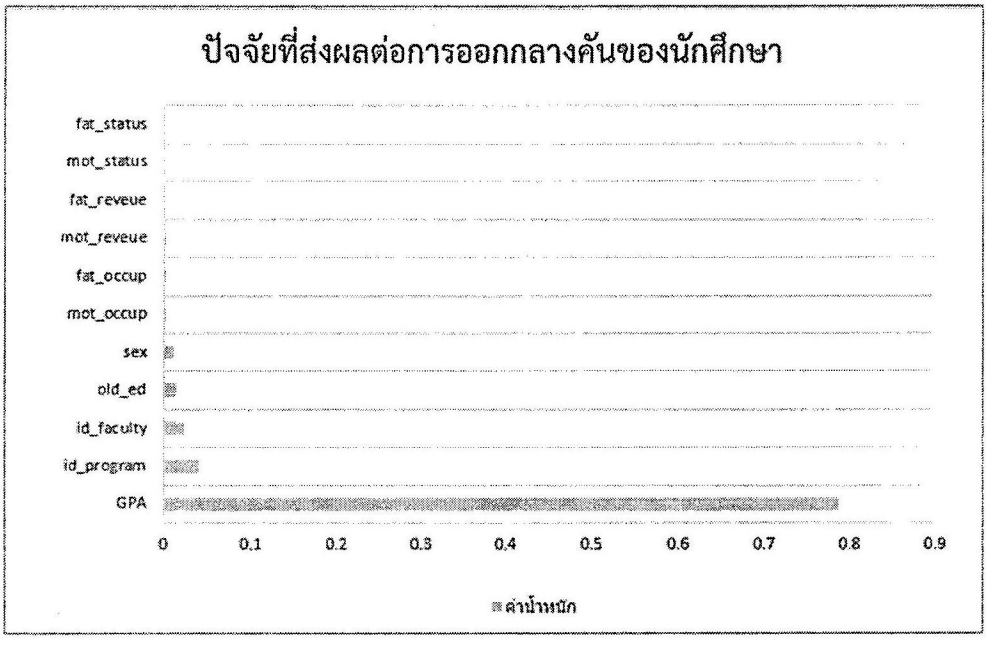
3.1.3.2.4 ดำเนินการวิเคราะห์ปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษาโดยใช้เทคนิคการคัดเลือกคุณลักษณะ (Feature Selection) ด้วยเทคนิค Information Gain

ก) การวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษาด้วยเทคนิค Information Gain เป็นการคัดเลือกคุณลักษณะของปัจจัยสำคัญที่ส่งผลต่อการออกกลางคันของนักศึกษา จากการวิเคราะห์ปัจจัยทั้งหมด 11 ปัจจัย ได้แก่ ข้อมูลภูมิหลังของนักศึกษาข้อมูลผู้ปกครอง ข้อมูลเกรดเฉลี่ย และสถานะของนักศึกษาที่เข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 3,913 คน สามารถแสดงค่าน้ำหนัก (Weight) ของแต่ละปัจจัยได้ดังตารางที่ 3-3 โดยปัจจัยที่สำคัญมากต่อการออกกลางคันของนักศึกษาจะมีค่าน้ำหนักสูงและปัจจัยที่มีความสำคัญต่อการออกกลางคันน้อยจะมีน้ำหนักน้อยตามลำดับ

ตารางที่ 3-3 อัตราค่าน้ำหนักของปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา

attribute	weight
GPA	0.788
Id_program	0.041
Id_faculty	0.024
Old_ed	0.016
Sex	0.012
Mot_occup	0.004
Fat_occup	0.004
Mot_reveue	0.003
Fat_reveue	0.001
Mot_status	0.000
Fat_status	0.000

จากตารางที่ 3-3 อัตราค่าน้ำหนักการวิเคราะปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษาด้วยเทคนิค Information Gain พบว่า ปัจจัยที่ส่งผลต่อการออกลางคัน 5 ปัจจัยแรก ได้แก่ เกรดเฉลี่ย (GPA) สาขาวิชา (Id_program) คณะที่สังกัด (Id_faculty) และเพศ (Sex) โดยปัจจัยอื่น ๆ ส่งผลต่อการออกกลางคันของนักศึกษาน้อยมากรายละเอียดแสดงดังภาพที่ 3-2



ภาพที่ 3-2 ปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา

ผลการวิเคราะห์ปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา พบว่า ปัจจัยที่สำคัญที่สุด ได้แก่ เกรดเฉลี่ย (GPAX) ซึ่งส่งผลต่อการออกลางคันถึงร้อยละ 78.8 ดังนั้นปัจจัยที่ต้องทำการแก้ไขอย่างเร่งด่วนจึงได้แก่ เกรดเฉลี่ย (GPAX) ดังนั้นในขั้นตอนถัดไปผู้วิจัยจึงทำการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษา เพื่อเป็นการหาแนวทางการป้องกัน แก้ปัญหาและส่งเสริมผลการเรียนของนักศึกษาให้นักศึกษามีผลการเรียนสูงขึ้นและจะสามารถลดอัตราการออกกลางคันของนักศึกษาต่อไปได้

3.1.3.3 การวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.1.3.3.1 ทำความเข้าใจปัญหา (Business Understanding) เป็นขั้นตอนแรกในกระบวนการ CRISP-DM ที่เน้นไปที่การเข้าใจปัญหาและแปลงปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูลทางดาต้าไมน์นิงพร้อมทั้งวางแผนในการดำเนินการคร่าว ๆ ซึ่งการพยากรณ์การเรียนของนักศึกษาเพื่อลดการออกกลางคันของนักศึกษาจึงทำการศึกษาข้อมูลนักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี เข้าศึกษาในปีการศึกษา 2555-2557 และจบการศึกษาแล้วและพร้อมทั้งศึกษาข้อมูลนักศึกษาที่กำลังศึกษาในปัจจุบันด้วย

3.1.3.3.2 การทำความเข้าใจข้อมูล (Data Understanding) ขั้นตอนนี้เริ่มจากการเก็บรวบรวมข้อมูล หลังจากนั้นจะเป็นการทำความเข้าใจกับข้อมูลที่จะนำมาใช้ความมีลักษณะอย่างไร ตรวจสอบข้อมูลที่ได้ทำการรวบรวมมาได้เพื่อดูความถูกต้องของข้อมูล และพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์ ซึ่งผู้วิจัยได้ทำการศึกษาข้อมูลภูมิหลัง ข้อมูลเกรดเฉลี่ยแต่ละเทอมและข้อมูลผลการเรียนแต่ละรายวิชาในแผนการเรียนของนักศึกษาที่เข้าศึกษาในปีการศึกษา 2555-2557 และจบการศึกษาแล้วจากประชากรทั้งหมด 2,379 คน เนื่องจากโครงสร้างหลักสูตรของแต่ละสาขาวิชามีความแตกต่างกัน ซึ่งการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาจึงต้องวิเคราะห์ข้อมูลแยกแต่ละสาขา ดังนั้นผู้วิจัยจึงทำการเลือกกลุ่มตัวอย่างเพื่อวิเคราะห์ปัจจัยการผลกรณ์ผลการเรียน ได้แก่ นักศึกษาสาขาวิชาภาษาอังกฤษที่เข้าศึกษาในปีการศึกษา 2555-2557 ซึ่งสำเร็จการศึกษาแล้ว จำนวน 141 คน โดยแยกตัวแปรทั้งหมด 34 ตัวแปร ได้แก่ ตัวแปรต้น 33 ตัวแปร ตัวแปรตาม 1 ตัวแปร

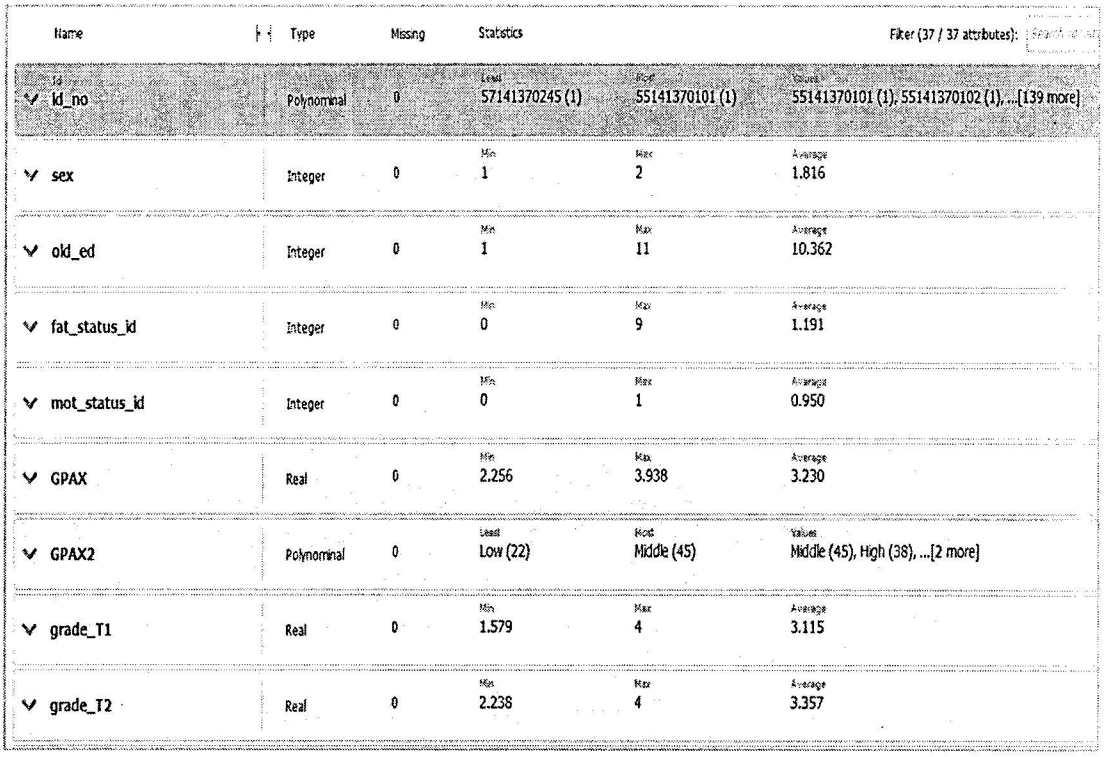
3.1.3.3.3 การเตรียมข้อมูล (Data Preparation) ขั้นตอนนี้เป็นขั้นตอนที่ทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมา (Raw Data) เพื่อปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่เหมาะสมกลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการแปลงข้อมูลนี้อาจจะต้องมีการทำข้อมูลให้ถูกต้อง (Data Cleaning) เช่น การแปลงข้อมูลให้อยู่ในช่วง (Scale) เดียวกัน หรือการเติมข้อมูลที่ขาดหายไป เป็นต้น

ตารางที่ 3-4 รายละเอียดข้อมูลแอททริบิวต์สำหรับการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียน

แอททริบิวต์	คำอธิบาย	รายละเอียด
Sex	เพศ	0=ชาย 1=หญิง
Old_ed	ข้อมูลการจบการศึกษาเดิม	11=ม.6 1=ปวช
Fat_status	สถานะบิดา	0=ถึงแก่กรรม 1=มีชีวิตอยู่ 9=ไม่ระบุ
Mor_status	สถานะมารดา	0=ถึงแก่กรรม 1=มีชีวิตอยู่ 9=ไม่ระบุ
Grade_T1	ผลการเรียนเฉลี่ย ปี 1 เทอม 1	VH = 3.50-4.00 H = 3.00-3.49 M = 2.50-2.99 L = 2.00-2.49 VL = 0.00-1.99
Grade_T2	ผลการเรียนเฉลี่ย ปี 1 เทอม 2	
Grade_T3	ผลการเรียนเฉลี่ย ปี 2 เทอม 1	
Grade_T4	ผลการเรียนเฉลี่ย ปี 2 เทอม 2	
Grade_T5	ผลการเรียนเฉลี่ย ปี 3 เทอม 1	

ตารางที่ 3-4 (ต่อ)

แอททริบิวต์	คำอธิบาย	รายละเอียด
ผลการเรียน 24 วิชา (5 เทอม)	24 วิชา (5 เทอม) ได้แก่ 101101 101201 101301 101401 101501 105131 105142 202101 202102 202207 202211 202307 202505 202506 202507 202702 901101 901201 902101 902301 903101 903301 904201 904301	4.0 = A 3.5 = B+ 3.0 = B 2.5 = C+ 2.0 = C 1.5 = D+ 1.0 = D 0.0 = F
GPAX	เกรดเฉลี่ยจบ	Under = 0.00-1.99 vLow = 2.00-2.24 Low = 2.25-2.80 Middle = 2.81-3.19 Good = 3.20-3.53 High = 3.54-4.00

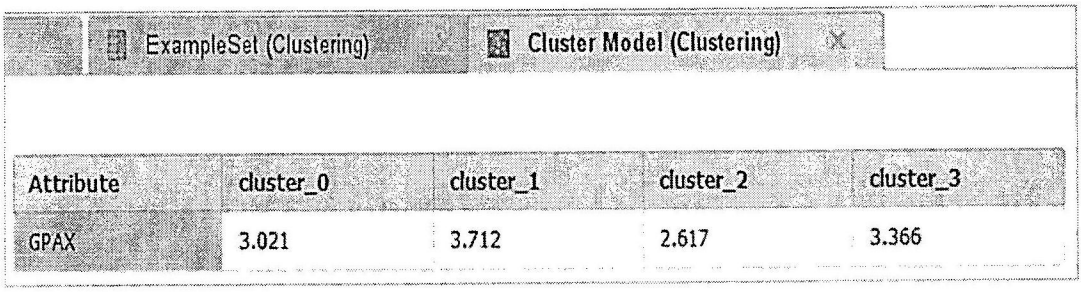


Name	Type	Missing	Statistics		Filter (37 / 37 attributes):
id_no	Polynominal	0	Least: 57141370245 (1)	Most: 55141370101 (1)	Values: 55141370101 (1), 55141370102 (1), ...[139 more]
sex	Integer	0	Min: 1	Max: 2	Average: 1.816
old_ed	Integer	0	Min: 1	Max: 11	Average: 10.362
fat_status_id	Integer	0	Min: 0	Max: 9	Average: 1.191
mot_status_id	Integer	0	Min: 0	Max: 1	Average: 0.950
GPAX	Real	0	Min: 2.256	Max: 3.938	Average: 3.230
GPAX2	Polynominal	0	Least: Low (22)	Most: Middle (45)	Values: Middle (45), High (38), ...[2 more]
grade_T1	Real	0	Min: 1.579	Max: 4	Average: 3.115
grade_T2	Real	0	Min: 2.238	Max: 4	Average: 3.357

ภาพที่ 3-3 แสดงรายละเอียดและลักษณะของข้อมูลแต่ละคุณลักษณะ

ก) การแปลงข้อมูล (Data Transformation) เป็นการเตรียมข้อมูลให้อยู่ในรูปแบบที่สามารถประมวลผลได้ เนื่องจากข้อมูลมีทั้งแบบตัวเลขและแบบตัวอักษรซึ่งไม่อยู่ในรูปแบบที่สามารถประมวลผลได้ จึงต้องทำการแทนค่าข้อมูลให้อยู่ในรูปแบบที่สามารถประมวลผลได้ เช่น เกรดเฉลี่ยแต่ละเทอมเป็นข้อมูลเชิงปริมาณที่สามารถแปลงข้อมูลตัวเลข (Numeric) ให้เป็นกลุ่มข้อมูล (Nnominal) โดยแบ่งข้อมูลเป็นช่วง ๆ และอยู่ในรูปแบบของไอเท็มและแอททริบิวต์

ข) จากการศึกษาข้อมูลผลการเรียน (GPAX) ของนักศึกษาภาษาอังกฤษที่เข้าศึกษาในปีการศึกษา 2555-2557 และสำเร็จการศึกษาแล้ว จำนวน 141 คน พบว่าผลการเรียนของนักศึกษาขณะจบการศึกษานักศึกษามีผลการเรียนระหว่าง 2.25-3.94 ซึ่งผู้วิจัยได้ทำการแบ่งกลุ่มผลการเรียนด้วยเทคนิค k-mean (k = 4) เพื่อแบ่งกลุ่มผลการเรียน 4 กลุ่ม ดังนี้ กลุ่มสูง (High) กลุ่มดี (Good) กลุ่มปานกลาง (Medium) กลุ่มอ่อน (Low) ซึ่งผลการจัดกลุ่มของผลการเรียนแสดงดังภาพที่ 3-4 และดังตารางที่ 3-5


ExampleSet (Clustering) | Cluster Model (Clustering)

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
GPAX	3.021	3.712	2.617	3.366

ภาพที่ 3-4 จุดศูนย์กลางของผลการเรียนแต่ละกลุ่ม

ตารางที่ 3-5 ข้อมูลผลการเรียนที่จัดกลุ่มด้วยเทคนิค K-mean

Cluster	จุดศูนย์กลางแต่ละคลัสเตอร์	ผลการเรียน (GPAX)	จำนวน(คน)
Cluster 0	3.021	2.81-3.16	45
Cluster 1	3.712	3.54-3.94	38
Cluster 2	2.617	2.25-2.78	22
Cluster 3	3.366	3.20-3.52	36
รวมทั้งหมด			**141**

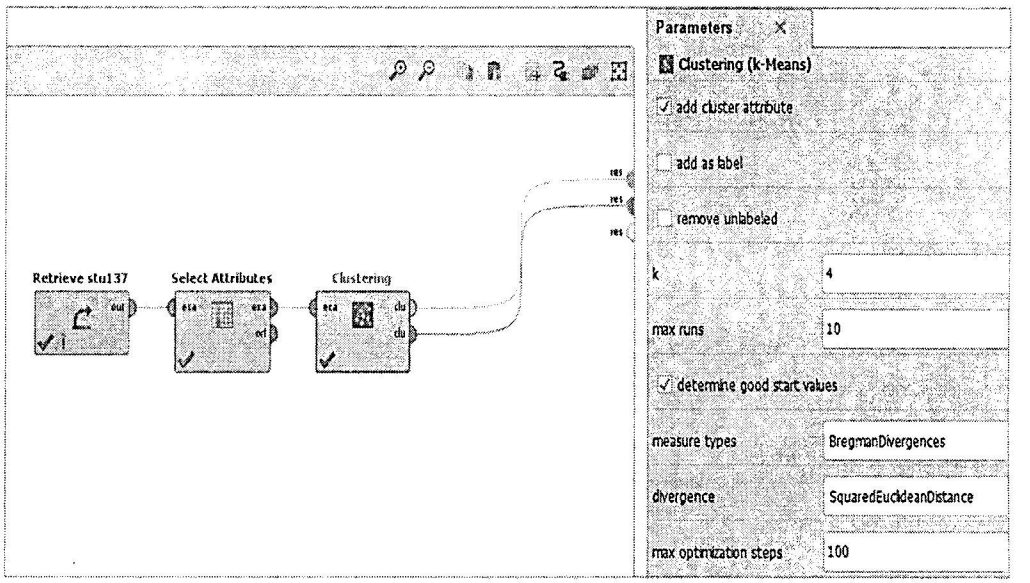
จากตารางที่ 3-5 การจัดกลุ่มผลการเรียนด้วยเทคนิค k-mean ซึ่งเป็นการหาความสัมพันธ์ของข้อมูลผลการเรียน (GPAX) ให้ข้อมูลที่มีลักษณะเดียวกันให้อยู่กลุ่มเดียวกันซึ่งหมายถึงข้อมูลดังกล่าวมีความสัมพันธ์กันสูง และกลุ่มที่มีความสัมพันธ์กันน้อยแสดงว่าข้อมูลดังกล่าวจะอยู่คนละกลุ่ม โดยผู้วิจัยทำการศึกษาการจัดกลุ่มผลการเรียน (GPAX) พบว่า Cluster 0 มีจุดศูนย์กลางของกลุ่มเท่ากับ 3.021 และมีจำนวนสมาชิก 45 คน Cluster 1 มีจุดศูนย์กลางของกลุ่มเท่ากับ 3.712 และมีจำนวนสมาชิก 38 คน Cluster 2 มีจุดศูนย์กลางของกลุ่มเท่ากับ 2.617 และมีจำนวนสมาชิก 22 คน และ Cluster 3 มีจุดศูนย์กลางของกลุ่มเท่ากับ 3.366 และมีจำนวนสมาชิก 36 คน

ผลของจากการจัดกลุ่มผลการเรียนนักศึกษาสาขาวิชาภาษาอังกฤษที่สำเร็จการศึกษาจำนวน 141 คน สามารถจัดกลุ่มผลการเรียนได้ จำนวน 4 กลุ่ม ได้แก่ กลุ่มสูง (High) กลุ่มดี (Good) กลุ่มปานกลาง (Medium) กลุ่มอ่อน (Low) แสดงรายละเอียดดังตารางที่ 3-6

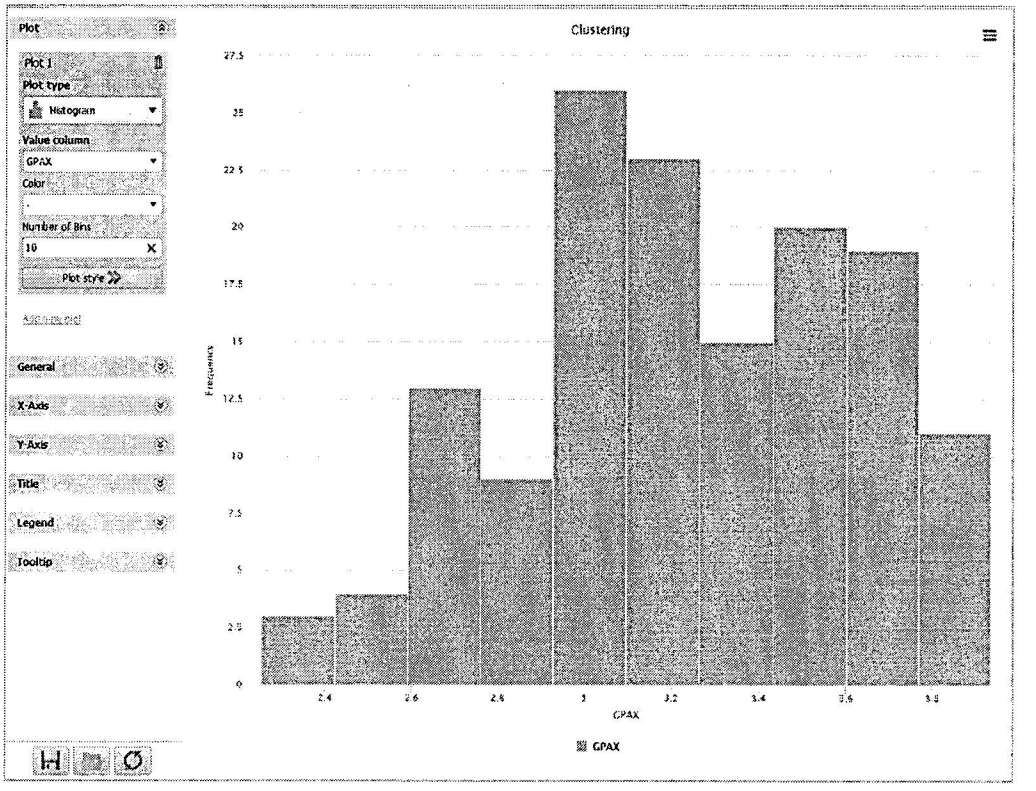
ตารางที่ 3-6 การจัดกลุ่มผลการเรียน (GPAX)

รายการ	กลุ่มผลการเรียน	ผลการเรียนแต่ละกลุ่ม	จำนวน(คน)
เกรดเฉลี่ยจบ(GPAX)	Low	2.25-2.80	22
	Middle	2.81-3.19	45
	Good	3.20-3.53	36
	High	3.54-4.00	38

วิธีการแบ่งกลุ่มผลการเรียน (GPAX : 2.25-3.93) ด้วยเทคนิค k-mean (k = 4) ด้วยโปรแกรม RapidMiner แสดงดังภาพที่ 3-6



ภาพที่ 3-5 Clustering (k-Mean)



ภาพที่ 3-6 กราฟผลการเรียน

ตารางที่ 3-7 ค่าเฉลี่ยของผลการเรียน (GPAX)

GPAX	ผลลัพธ์
min	2.25
Max	3.93
Average	3.23
deviation	0.386

จากการศึกษาข้อมูลผลการเรียน (GPAX) ของนักศึกษาภาษาอังกฤษที่เข้าศึกษาในปีการศึกษา 2555-2557 และสำเร็จการศึกษาแล้วจำนวน 141 คน สามารถจัดกลุ่มผลการเรียนด้วยเทคนิค k-mean ได้จำนวน 4 กลุ่ม ได้แก่ กลุ่มสูง (High) กลุ่มดี (Good) กลุ่มปานกลาง (Medium) กลุ่มอ่อน (Low) สามารถแสดงรายละเอียดของผลการเรียนแต่ละกลุ่มดังตารางที่ 3-8

ตารางที่ 3-8 การจัดกลุ่มข้อมูลผลการเรียนนักศึกษาที่สำเร็จการศึกษา

No	ID_no	Cluster	GPAX
1	S55xxx101	cluster_0	2.887573
2	S55xxx102	cluster_0	2.97093
3	S55xxx103	cluster_0	3.16568
4	S55xxx104	cluster_0	3.106508
5	S55xxx105	cluster_0	3.068047
6	S55xxx106	cluster_0	3.150887
7	S55xxx107	cluster_0	2.943786
8	S55xxx108	cluster_3	3.399408
9	S55xxx109	cluster_1	3.606508
10	S55xxx110	cluster_3	3.381656
11	S55xxx111	cluster_0	3.04142
12	S55xxx113	cluster_3	3.209302
13	S55xxx114	cluster_2	2.69774
14	S55xxx115	cluster_2	2.550279
15	S55xxx116	cluster_0	3.08284
16	S55xxx119	cluster_1	3.60355
17	S55xxx122	cluster_2	2.613372
18	S55xxx123	cluster_0	2.914201
19	S55xxx124	cluster_2	2.789772
20	S55xxx127	cluster_2	2.287709
21	S55xxx128	cluster_2	2.719298

ตารางที่ 3-8 (ต่อ)

No	ID_no	Cluster	GPAX
22	S55xxx129	cluster_2	2.477401
23	S55xxx130	cluster_1	3.568047
24	S55xxx131	cluster_3	3.443786
25	S55xxx132	cluster_0	2.956395
26	S55xxx133	cluster_1	3.822485
27	S55xxx134	cluster_0	3.00862
28	S55xxx135	cluster_0	3.029069
29	S55xxx136	cluster_2	2.255747
30	S55xxx137	cluster_0	3.061046
31	S55xxx138	cluster_1	3.653846
32	S55xxx139	cluster_0	3.12426
33	S55xxx140	cluster_0	2.822674
34	S55xxx142	cluster_0	3.061046
35	S55xxx143	cluster_2	2.419075
36	S55xxx144	cluster_2	2.611731
37	S55xxx146	cluster_3	3.375739
38	S55xxx147	cluster_3	3.213017
39	S55xxx149	cluster_3	3.218023
40	S55xxx150	cluster_3	3.52071
41	S55xxx151	cluster_3	3.363905
42	S55xxx152	cluster_1	3.78994
43	S55xxx153	cluster_1	3.639053
44	S55xxx155	cluster_1	3.639053
45	S55xxx156	cluster_3	3.24852
46	S55xxx157	cluster_3	3.505917
47	S55xxx158	cluster_3	3.437869
48	S55xxx159	cluster_2	2.714285
49	S55xxx160	cluster_0	3.147928
50	S56xxx101	cluster_1	3.872781
51	S56xxx102	cluster_1	3.83136
52	S56xxx103	cluster_1	3.816568
53	S56xxx104	cluster_0	2.843195
54	S56xxx106	cluster_1	3.56213

ตารางที่ 3-8 (ต่อ)

No	ID_no	Cluster	GPAX
55	S56xxx107	cluster_0	3.156804
56	S56xxx109	cluster_2	2.671597
57	S56xxx110	cluster_3	3.239644
58	S56xxx111	cluster_0	3.111111
59	S56xxx112	cluster_2	2.676829
60	S56xxx113	cluster_3	3.275147
61	S56xxx114	cluster_2	2.571005
62	S56xxx115	cluster_1	3.710059
63	S56xxx116	cluster_0	2.931952
64	S56xxx117	cluster_1	3.74852
65	S56xxx118	cluster_1	3.872781
66	S56xxx119	cluster_1	3.659763
67	S56xxx120	cluster_2	2.621301
68	S56xxx121	cluster_2	2.701183
69	S56xxx123	cluster_1	3.837278
70	S56xxx124	cluster_1	3.875739
71	S56xxx126	cluster_0	3.032544
72	S56xxx129	cluster_0	2.819526
73	S56xxx130	cluster_0	3.053254
74	S56xxx131	cluster_1	3.763313
75	S56xxx132	cluster_3	3.52071
76	S56xxx133	cluster_3	3.334319
77	S56xxx134	cluster_2	2.639053
78	S56xxx135	cluster_0	3.049707
79	S56xxx136	cluster_0	2.932748
80	S56xxx137	cluster_2	2.7071
81	S56xxx141	cluster_1	3.718934
82	S56xxx142	cluster_0	3.130177
83	S56xxx143	cluster_3	3.244186
84	S56xxx144	cluster_0	3.046783
85	S56xxx145	cluster_3	3.502923
86	S56xxx146	cluster_0	3.070175
87	S56xxx147	cluster_3	3.494082

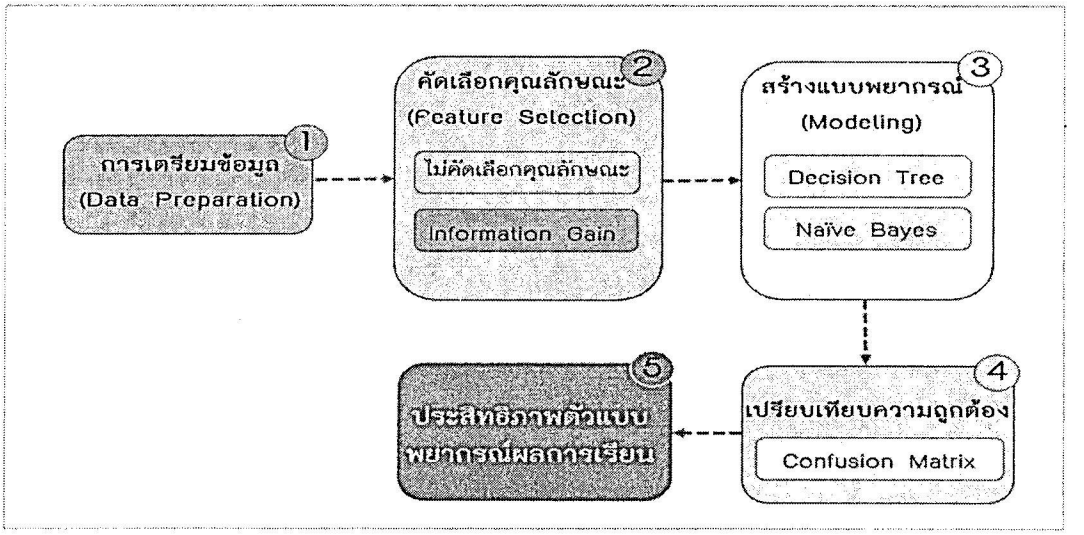
ตารางที่ 3-8 (ต่อ)

No	ID_no	Cluster	GPAX
88	S56xxx148	cluster_1	3.656804
89	S57xxx101	cluster_1	3.766272
90	S57xxx103	cluster_0	2.964497
91	S57xxx105	cluster_3	3.517751
92	S57xxx106	cluster_0	3.091715
93	S57xxx107	cluster_1	3.56213
94	S57xxx109	cluster_3	3.473372
95	S57xxx110	cluster_0	2.955621
96	S57xxx111	cluster_3	3.434911
97	S57xxx112	cluster_0	3.10355
98	S57xxx113	cluster_3	3.497041
99	S57xxx114	cluster_0	2.961538
100	S57xxx115	cluster_1	3.757396
101	S57xxx116	cluster_2	2.74852
102	S57xxx120	cluster_1	3.662721
103	S57xxx121	cluster_3	3.236686
104	S57xxx122	cluster_1	3.573964
105	S57xxx123	cluster_1	3.609467
106	S57xxx124	cluster_1	3.56213
107	S57xxx127	cluster_0	2.869822
108	S57xxx128	cluster_0	2.923076
109	S57xxx129	cluster_3	3.449704
110	S57xxx132	cluster_0	2.973372
111	S57xxx135	cluster_1	3.547337
112	S57xxx138	cluster_1	3.718934
113	S57xxx139	cluster_1	3.937869
114	S57xxx140	cluster_1	3.881656
115	S57xxx141	cluster_3	3.369822
116	S57xxx142	cluster_3	3.414201
117	S57xxx145	cluster_3	3.245562
118	S57xxx203	cluster_1	3.698224
119	S57xxx207	cluster_3	3.538461
120	S57xxx209	cluster_3	3.239644

ตารางที่ 3-8 (ต่อ)

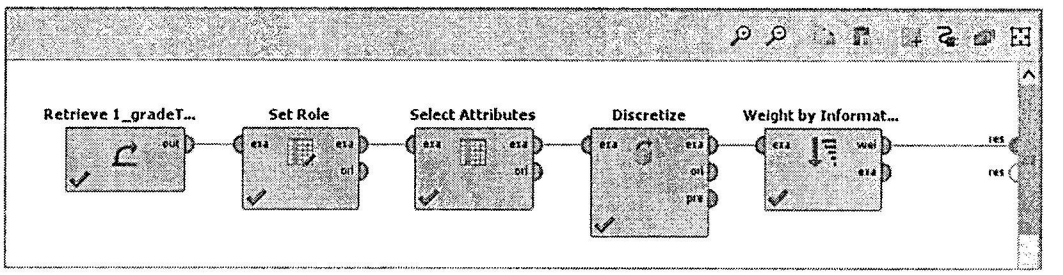
No	ID_no	Cluster	GPAX
121	S57xxx211	cluster_2	2.778106
122	S57xxx213	cluster_0	3.056213
123	S57xxx215	cluster_0	3.112426
124	S57xxx217	cluster_0	3.0
125	S57xxx221	cluster_3	3.366863
126	S57xxx223	cluster_0	2.985207
127	S57xxx224	cluster_0	3.139053
128	S57xxx227	cluster_3	3.375739
129	S57xxx228	cluster_3	3.281065
130	S57xxx230	cluster_0	3.100591
131	S57xxx233	cluster_2	2.587719
132	S57xxx234	cluster_3	3.26923
133	S57xxx235	cluster_0	3.002958
134	S57xxx236	cluster_1	3.559171
135	S57xxx239	cluster_2	2.739644
136	S57xxx240	cluster_3	3.33136
137	S57xxx241	cluster_1	3.760355
138	S57xxx242	cluster_1	3.872781
139	S57xxx243	cluster_3	3.28994
140	S57xxx244	cluster_3	3.266272
141	S57xxx245	cluster_1	3.609467

3.1.3.3.4 ดำเนินการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียน โดยใช้เทคนิคการคัดเลือกคุณลักษณะ (Feature Selection) ที่สำคัญกับการพยากรณ์เหมืองข้อมูลด้วยเทคนิค Information Gain กับเทคนิค Decision Tree และเทคนิค Naïve Bayes แสดงขั้นตอนการคัดเลือกคุณลักษณะแสดงดังภาพที่ 3-7



ภาพที่ 3-7 ขั้นตอนการคัดเลือกคุณลักษณะ

ก) การวิเคราะห์ปัจจัยพยากรณ์ผลการเรียนด้วยเทคนิค Information Gain เทคนิคการคัดเลือกคุณลักษณะที่สำคัญสำหรับการพยากรณ์ผลการเรียนด้วยเทคนิค Information Gain (IG) จากคุณลักษณะทั้งหมด 33 คุณลักษณะ ทำให้ได้ค่าน้ำหนัก (Weight) ในแต่ละปัจจัยที่สำคัญออกมา โดยผลการวิจัยสามารถแสดงขั้นตอนการทำงานด้วย Rapid Miner แสดงดังภาพที่ 3-8 และแสดงผลการวิเคราะห์ค่าน้ำหนักของปัจจัยการพยากรณ์ผลการเรียนแต่ละปัจจัยดังตารางที่ 3-9



ภาพที่ 3-8 การคัดเลือกคุณลักษณะที่สำคัญด้วยเทคนิค Information Gain

ตารางที่ 3-9 ค่าน้ำหนักของปัจจัยการพยากรณ์ผลการเรียนด้วยเทคนิค Information Gain (IG)

attribute	weight
grade_T2	0.928247
grade_T3	0.917103
grade_T4	0.810689
grade_T1	0.775998
202507	0.729367
grade_T5	0.717749

ตารางที่ 3-9 (ต่อ)

attribute	weight
202506	0.68555
105131	0.655477
202101	0.648303
901101	0.624847
202211	0.596433
202207	0.572321
101501	0.524012
903301	0.502386
202505	0.463088
202102	0.431724
202307	0.419521
902301	0.418075
901201	0.382464
904301	0.365833
101201	0.360603
904201	0.329679
903101	0.3105
902101	0.309214
202702	0.263139
101301	0.26116
105142	0.24395
101101	0.242279
101401	0.171101
Fat_status	0.045969
Sex	0.012024
Mot_status	0.003567
Old_ed	0.001648

จากตารางที่ 3-9 เทคนิคการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนด้วยเทคนิค Information Gain พบว่า ผลการเรียนเฉลี่ยปี 1 เทอม 2 (grade_T2) มีความสำคัญสูงสุดโดยมีค่าน้ำหนักที่ 0.9282 และข้อมูลการจบการศึกษาเดิม (Old_ed) มีความสำคัญน้อยสุดมีค่าน้ำหนักที่ 0.0016 ซึ่งสังเกตได้ว่าไม่มีแอททริบิวต์ที่มีค่าน้ำหนักเป็น 0 ซึ่งหมายถึงทุกปัจจัยส่งผลต่อการออกกลางคันของนักศึกษาและสามารถนำไปสร้างแบบพยากรณ์ผลการเรียนได้ แต่เนื่องด้วยจำนวนของปัจจัยค่อนข้างเยอะ

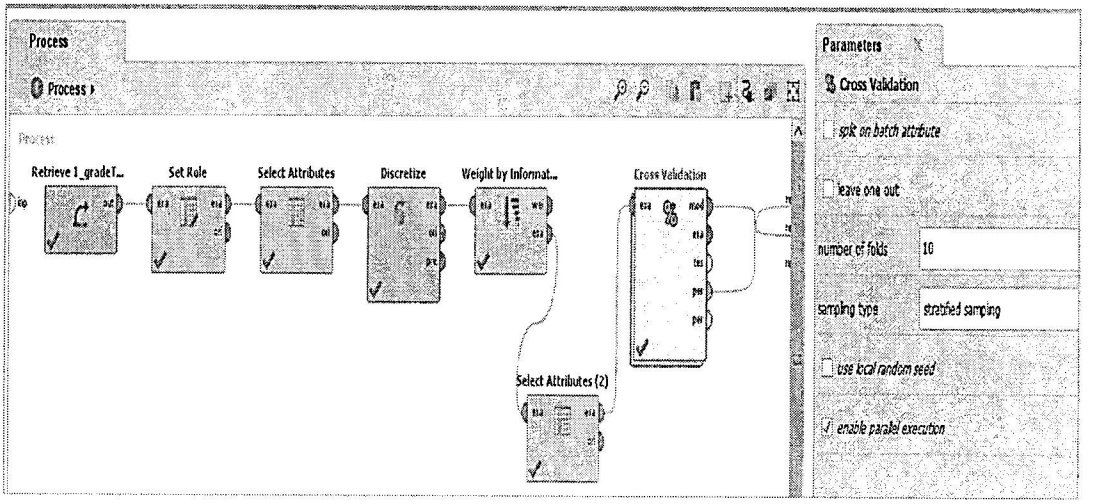
ดังนั้นผู้วิจัยจึงทำการตรวจสอบความถูกต้องสูงสุดของแต่ละแบบจำลองด้วยการตัดปัจจัยที่มีความสำคัญน้อยออกทีละ 1 ปัจจัยเพื่อคัดเลือกจำนวนปัจจัยที่ส่งผลต่อแบบจำลองการพยากรณ์ผลการเรียนที่ความถูกต้องสูงสุด

ข) ตรวจสอบคุณลักษณะสำคัญของปัจจัยการพยากรณ์ผลการเรียน สามารถคัดเลือกปัจจัยการพยากรณ์ได้จากการคำนวณค่าน้ำหนัก (Weight) ด้วยเทคนิค Information Gain โดยปัจจัยที่ถูกตัดออกจะเป็นปัจจัยที่ไม่ส่งผลต่อการพยากรณ์ผลการเรียน ซึ่งได้แก่ปัจจัยที่มีค่าน้ำหนักเท่ากับ 0 โดยจากผลการคำนวณค่าน้ำหนัก 33 ปัจจัยไม่พบปัจจัยที่มีค่าน้ำหนักเป็น 0

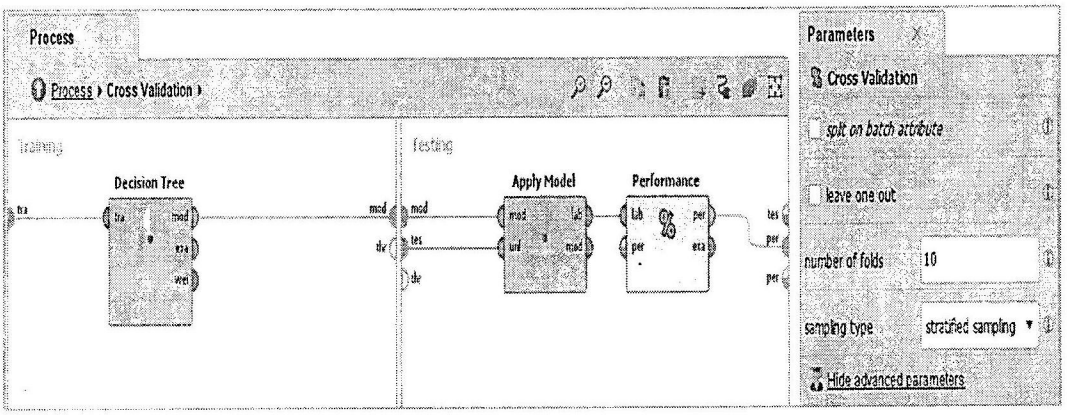
3.1.3.3.5 ตรวจสอบประสิทธิภาพของการพยากรณ์ผลการเรียนเป็นการตรวจสอบความถูกต้องการพยากรณ์ ผู้วิจัยดำเนินการตรวจสอบประสิทธิภาพการพยากรณ์ผลการเรียนโดยทำการเปรียบเทียบประสิทธิภาพของการพยากรณ์แบบไม่มีการคัดเลือกคุณลักษณะมีการเลือกคุณลักษณะด้วยเทคนิค Information Gain และมีการลดคุณลักษณะลงทีละ 1 ปัจจัย ด้วยเทคนิคเหมืองข้อมูล Decision Tree และเทคนิค Naïve Bayes

ก) ตรวจสอบประสิทธิภาพของเทคนิค Information Gain ด้วย Cross Validation แบบ 10-fold Cross Validation ด้วยเทคนิคเหมืองข้อมูล 2 เทคนิค ได้แก่ (1) เทคนิค Decision Tree และ (2) เทคนิค Naïve Bayes จากปัจจัยทั้งหมด 33 ปัจจัย ผลการตรวจสอบแสดงดังนี้

(1) เทคนิค Decision Tree วิธีการวัดประสิทธิภาพของเทคนิค Information Gain ด้วย Cross Validation แบบรูปแบบ 10-fold Cross Validation ของต้นไม้ตัดสินใจ (Decision Tree) ซึ่งประกอบด้วยชุดเครื่องมือ Apply Model และ Performance ทำหน้าที่สร้างแบบพยากรณ์ผลการเรียนมาทำการทดสอบความถูกต้อง ด้วยการสลับชุดข้อมูลย่อยเพื่อสร้างแบบพยากรณ์และชุดทดสอบการพยากรณ์ แสดงดังภาพที่ 3-9



ภาพที่ 3-9 การวัดประสิทธิภาพ แบบ 10-fold Cross Validation



ภาพที่ 3-10 การจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree)

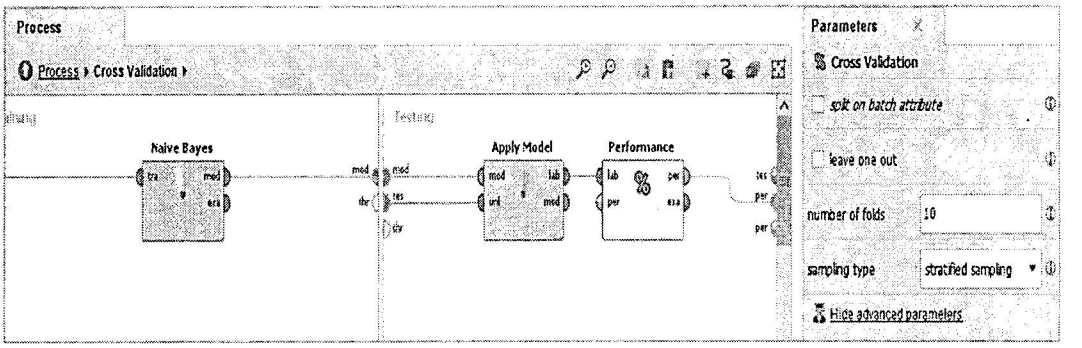
ประสิทธิภาพของตัวแบบพยากรณ์ผลการเรียนจากปัจจัยการพยากรณ์ผลการเรียนทั้งหมด 33 ปัจจัย ด้วยเทคนิค Decision Tree ผลการประเมินความถูกต้องเท่ากับร้อยละ 64.67 แสดงรายละเอียดดังภาพที่ 3-11

accuracy: 64.67% +/- 10.93% (micro average: 64.54%)

	true Middle	true Good	true High	true Low	class precision
pred. Middle	28	11	0	9	58.33%
pred. Good	9	19	7	0	54.29%
pred. High	0	6	31	0	83.78%
pred. Low	8	0	0	13	61.90%
class recall	62.22%	52.78%	81.58%	59.09%	

ภาพที่ 3-11 ประสิทธิภาพตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Decision Tree จากปัจจัย 33 ปัจจัย

(2) เทคนิค Naïve Bayes วิธีการวัดประสิทธิภาพของเทคนิค Information Gain ด้วย Cross Validation แบบรูปแบบ 10-fold Cross Validation ของนาอีฟเบย์ (Naïve Bayes) ซึ่งประกอบด้วยชุดเครื่องมือ Apply Model และ Performance ทำหน้าที่สร้างแบบพยากรณ์ผลการเรียนมาทำการทดสอบความถูกต้อง ด้วยการสลับชุดข้อมูลย่อย เพื่อสร้างแบบพยากรณ์ผลการเรียนและชุดทดสอบการพยากรณ์ผลการเรียน แสดงดังภาพที่ 3-12 และภาพที่ 3-13



ภาพที่ 3-12 การจำแนกข้อมูลด้วยเทคนิค Naïve Bayes

ประสิทธิภาพของตัวแบบพยากรณ์ผลการเรียนจากปัจจัย 33 ปัจจัยด้วยเทคนิค Naïve Bayes ผลการประเมินประสิทธิภาพความถูกต้องเท่ากับร้อยละ 78.10 แสดงรายละเอียดดังภาพ

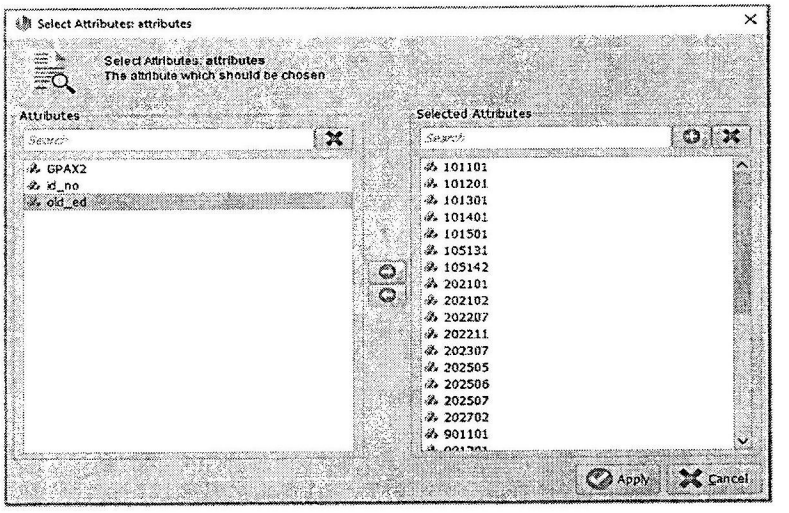
accuracy: 78.10% +/- 8.82% (micro average: 78.01%)

	true Middle	true Good	true High	true Low	class precision
pred. Middle	38	5	0	6	77.55%
pred. Good	2	26	7	1	72.22%
pred. High	0	5	31	0	86.11%
pred. Low	5	0	0	15	75.00%
class recall	84.44%	72.22%	81.58%	68.18%	

ภาพที่ 3-13 ประสิทธิภาพตัวแบบพยากรณ์ผลการเรียนด้วย Naïve Bayes จากปัจจัย 33 ปัจจัย

ข) ตรวจสอบประสิทธิภาพของเทคนิค Information Gain ด้วย Cross Validation แบบ 10-fold Cross Validation โดยทำการทดลองลดปัจจัยทีละหนึ่งปัจจัยเพื่อตรวจสอบค่าความถูกต้องแต่ละรอบและเลือกปัจจัยการพยากรณ์ผลการเรียนที่ให้ค่าความถูกต้องที่ดีที่สุดในการพยากรณ์ผลการเรียนด้วยเทคนิคเหมืองข้อมูล 2 เทคนิค ได้แก่ (1) เทคนิค Decision Tree และ (2) เทคนิค Naïve Bayes ผลการตรวจสอบแสดงดังนี้

(1) เทคนิค Decision Tree ผลการตรวจสอบค่าความถูกต้องของการพยากรณ์ผลการเรียน โดยปัจจัยข้อมูลการจบการศึกษาเดิม (Old_ed) มีความสำคัญน้อยสุดที่ค่าน้ำหนัก 0.0016 และปัจจัยผลการเรียนเฉลี่ยปี 1 เทอม 2 (grade_T2) มีความสำคัญสูงสุดที่ค่าน้ำหนัก 0.9282 ดังนั้นผู้วิจัยทำการตรวจสอบความถูกต้องของการพยากรณ์จากการลดปัจจัยที่มีน้ำหนักน้อยตัวถัดไปทีละรอบจนครบทุกปัจจัย สามารถแสดงวิธีการลดปัจจัยได้ดังภาพที่ 3-14 และผลการวัดประสิทธิภาพแสดงดังตารางที่ 3-10



ภาพที่ 3-14 การลดปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาครั้งละ 1 ปัจจัย

ตารางที่ 3-10 ผลการเปรียบเทียบค่าความถูกต้องของการพยากรณ์ผลการเรียนด้วยเทคนิค Decision Tree จากการลดปัจจัยครั้งละ1 ปัจจัย

Number of Attribute	Attribute Reduction	Accuracy
33	0	64.67
32	1	63.95
31	2	63.95
30	3	63.95
29	4	65.33
28	5	66.05
27	6	66.05
26	7	66.05
25	8	66.05
24	9	66.05
23	10	66.05
22	11	65.38
21	12	66.10
20	13	66.81
19	14	66.81
18	15	66.81
17	16	65.48
16	17	65.48

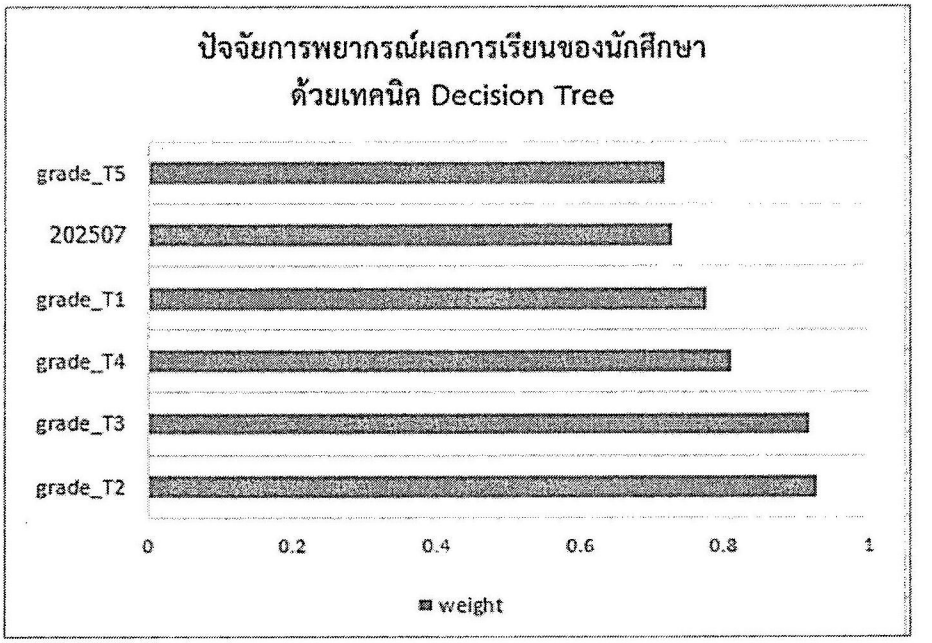
ตารางที่ 3-10 (ต่อ)

Number of Attribute	Attribute Reduction	Accuracy
15	18	66.14
14	19	66.14
13	20	70.43
12	21	71.10
11	22	72.52
10	23	69.62
9	24	69.62
8	25	68.86
7	26	70.29
6	27	**73.86***
5	28	71.76
4	29	72.48
3	30	65.38
2	31	65.33
1	32	64.62

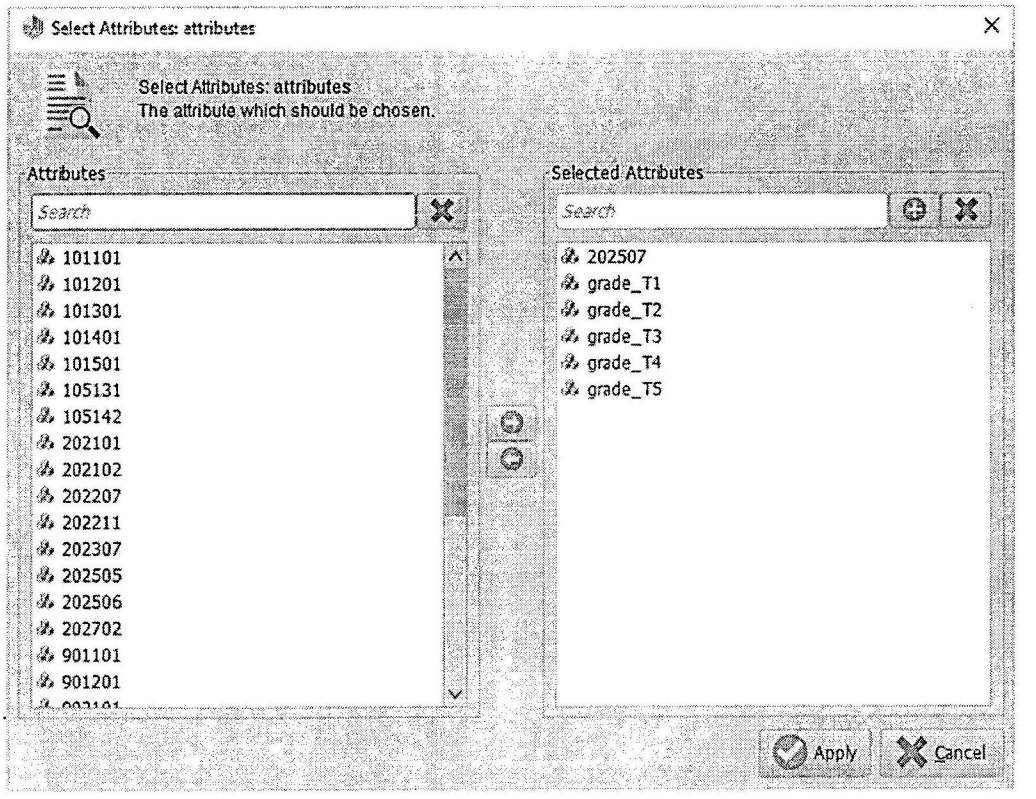
จากตารางที่ 3-10 ผลการวัดประสิทธิภาพการพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Decision Tree พบว่า จำนวนปัจจัยที่ค่าการพยากรณ์ผลการเรียนค่าความถูกต้องสูงสุดได้แก่ จำนวน 6 ปัจจัยให้ค่าความถูกต้องที่ร้อยละ 73.86 อันได้แก่ ผลการเรียนเฉลี่ยปี 1 เทอม 2 (grade_T2) ผลการเรียนเฉลี่ยปี 2 เทอม 1 (grade_T3) ผลการเรียนเฉลี่ยปี 2 เทอม 2 (grade_T4) ผลการเรียนเฉลี่ยปี 1 เทอม 1 (grade_T1) ผลการเรียนรายวิชา 202507 และผลการเรียนเฉลี่ยปี 3เทอม 1 (grade_T5) สามารถสรุปได้ดังตารางที่ 3-11 และภาพที่ 3-15

ตารางที่ 3-11 ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Decision Tree

attribute	weight
grade_T2	0.928247
grade_T3	0.917103
grade_T4	0.810689
grade_T1	0.775998
202507	0.729367
grade_T5	0.717749



ภาพที่ 3-15 ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Decision Tree



ภาพที่ 3-16 ปัจจัยการพยาการณ์ผลการเรียนจำนวน 6 ปัจจัย

accuracy: 73.86% +/- 10.75% (micro average: 73.76%)

	true Middle	true Good	true High	true Low	class precision
pred. Middle	34	10	0	8	65.38%
pred. Good	7	23	5	0	65.71%
pred. High	0	3	33	0	91.67%
pred. Low	4	0	0	14	77.78%
class recall	75.56%	63.89%	86.84%	63.64%	

ภาพที่ 3-17 ประสิทธิภาพการพยากรณ์ผลการเรียนด้วยเทคนิคต้นไม้ตัดสินใจจากปัจจัย 6 ปัจจัย

(2) เทคนิค Naïve Bayes การตรวจสอบค่าความถูกต้องของการพยากรณ์ผลการรเรียนด้วยเทคนิค Naïve Bayes ผู้วิจัยทำการตรวจสอบความถูกต้องของการพยากรณ์จากการลดปัจจัยที่มีค่าน้ำหนักน้อยทีละปัจจัย แสดงดังตารางที่ 3-12

ตารางที่ 3-12 ผลการเปรียบเทียบค่าความถูกต้องของการพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes จากการลดปัจจัยครั้งละ 1 ปัจจัย

Number of Attribute	Attribute Reduction	Accuracy
33	0	78.10
32	1	80.19
31	2	80.19
30	3	80.19
29	4	79.43
28	5	80.14
27	6	80.19
26	7	80.90
25	8	79.43
24	9	79.48
23	10	80.24
22	11	81.57
21	12	83.71
20	13	82.95
19	14	83.62
18	15	84.33*

ตารางที่ 3-12 (ต่อ)

Number of Attribute	Attribute Reduction	Accuracy
17	16	83.62
16	17	83.67
15	18	83.71
14	19	**84.33***
13	20	83.71
12	21	82.29
11	22	83.05
10	23	81.62
9	24	81.62
8	25	82.33
7	26	83.81
6	27	83.10
5	28	80.24
4	29	80.19
3	30	72.43
2	31	71.05
1	32	64.62

จากตารางที่ 3-12 ผลการวัดประสิทธิภาพการพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes พบว่า จำนวนปัจจัยที่ค่าการพยากรณ์ผลการเรียนมีค่าความถูกต้องสูงสุดได้แก่ จำนวน 14 ปัจจัย และ 18 ปัจจัย โดยให้ค่าความถูกต้องสูงสุดร้อยละ 84.33 ดังนั้นผู้วิจัยเลือกใช้จำนวนปัจจัยที่มีคุณลักษณะน้อยที่สุด ได้แก่ 14 ปัจจัยในการสร้างตัวแบบพยากรณ์ผลการเรียน ซึ่งได้แก่ ผลการเรียนเฉลี่ยปี 1 เทอม 2 (grade_T2) ผลการเรียนเฉลี่ยปี 2 เทอม 1 (grade_T3) ผลการเรียนเฉลี่ยปี 2 เทอม 2 (grade_T4) ผลการเรียนเฉลี่ยปี 1 เทอม 1 (grade_T1) ผลการเรียนรายวิชา 202507 ผลการเรียนเฉลี่ยปี 3 เทอม 1 (grade_T5) ผลการเรียนรายวิชา 202506 ผลการเรียนรายวิชา 105131 ผลการเรียนรายวิชา 202101 ผลการเรียนรายวิชา 901101 ผลการเรียนรายวิชา 202211 ผลการเรียนรายวิชา 202207 ผลการเรียนรายวิชา 101501 และผลการเรียนรายวิชา 903301 แสดงดังตารางที่ 3-13

ตารางที่ 3-13 ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Naïve Bayes

attribute	weight
grade_T2	0.928247
grade_T3	0.917103
grade_T4	0.810689
grade_T1	0.775998

ตารางที่ 3-13 (ต่อ)

attribute	weight
202507	0.729367
grade_T5	0.717749
202506	0.68555
105131	0.655477
202101	0.648303
901101	0.624847
202211	0.596433
202207	0.572321
101501	0.524012
903301	0.502386

accuracy: 84.33% +/- 9.45% (micro average: 84.40%)

	true Middle	true Good	true High	true Low	class precision
pred. Middle	40	4	0	4	83.33%
pred. Good	1	28	5	0	82.35%
pred. High	0	4	33	0	89.19%
pred. Low	4	0	0	18	81.82%
class recall	88.89%	77.78%	86.84%	81.82%	

ภาพที่ 3-18 ประสิทธิภาพการพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes จากปัจจัย 14 ปัจจัย

ค) เปรียบเทียบประสิทธิภาพของการพยากรณ์ผลการเรียนจากการคัดเลือกคุณลักษณะโดยใช้เทคนิค Information Gain ด้วย Cross Validation แบบ 10-fold Cross Validation ด้วยเทคนิคเหมืองข้อมูลจำนวน 2 เทคนิค ได้แก่ เทคนิค Decision Tree และเทคนิค Naïve Bayes ด้วยข้อมูลของผู้เข้าศึกษาประจำปีการศึกษา 2555-2557 และสำเร็จการศึกษาแล้วของนักศึกษาสาขาวิชาภาษาอังกฤษ จำนวน 141 คน ผลการวิจัยพบว่า ตัวแบบพยากรณ์ผลการเรียนที่ใช้ปัจจัยทั้งหมด 33 ปัจจัยจะมีค่าความถูกต้องของการพยากรณ์ต่ำกว่าการสร้างแบบพยากรณ์ที่มีการลดปัจจัยการพยากรณ์ ซึ่งเทคนิคการ Decision Tree การลดปัจจัยแล้วให้ค่าความถูกต้องของการพยากรณ์สูงสุด ได้แก่ ปัจจัยการพยากรณ์ 6 ปัจจัย ให้ค่าความถูกต้องร้อยละ 73.86 และเทคนิค Naïve Bayes การลดปัจจัยแล้วให้ค่าความถูกต้องของการพยากรณ์สูงสุด ได้แก่ ปัจจัยการพยากรณ์ผลการเรียนจำนวน 14 ปัจจัย ให้ค่าความถูกต้องร้อยละ 84.33 แสดงดังตารางที่ 3-14

ตารางที่ 3-14 การเปรียบเทียบประสิทธิภาพการพยากรณ์ผลการเรียนด้วยเทคนิค Decision Tree และเทคนิค Naïve Bayes

เทคนิคในการคัดเลือกคุณลักษณะ	เทคนิคการพยากรณ์	จำนวนปัจจัย	ค่าความถูกต้อง (Accuracy)
ไม่มีการคัดเลือกคุณลักษณะ	Decision Tree	33	64.67
	Naïve Bayes	33	78.10
Information Gain	Decision Tree	6	73.86
	Naïve Bayes	14	84.33

3.2 ระยะที่ 2 พัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.2.1 วัตถุประสงค์การวิจัยระยะที่ 2

3.2.1.1 พัฒนาตัวแบบการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.2.2 ขอบเขตการวิจัยระยะที่ 2

3.2.2.1 ประชากร ได้แก่ นักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี เข้าศึกษาในปีการศึกษา 2555-2557 ที่สำเร็จการศึกษา จำนวน 2,379 คน และนักศึกษาที่ลงทะเบียนเรียนในภาคการศึกษา 2/2563 จำนวน 3,579 คน

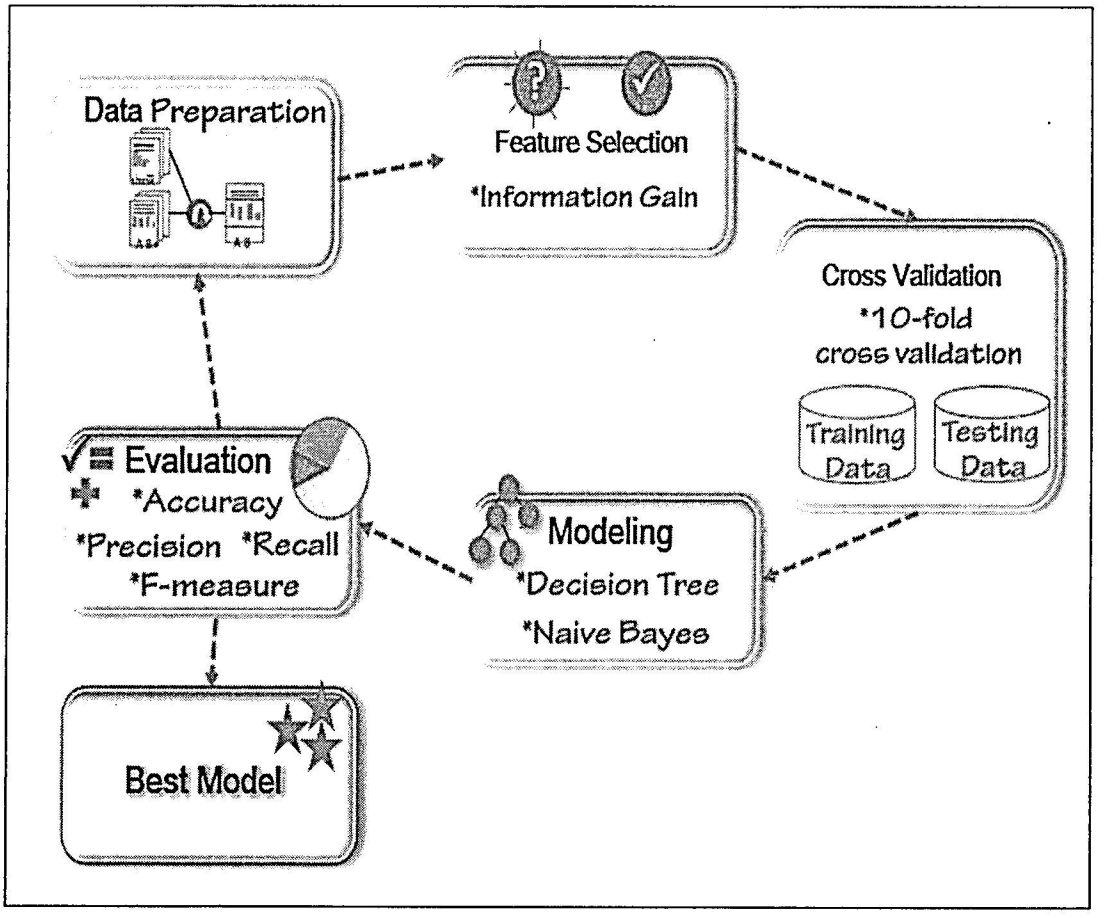
3.2.2.2 กลุ่มตัวอย่าง ได้แก่ นักศึกษาภาคปกติในสาขาวิชาภาษาอังกฤษ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่เข้าศึกษาในปีการศึกษา 2555-2557 ซึ่งจบการศึกษาแล้ว จำนวน 141 คน ได้จากวิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling)

3.2.3 เครื่องมือที่ใช้ในการวิจัยระยะที่ 2

3.2.3.1 ตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อการลดการออกกลางคันของนักศึกษาปริญญาตรี

3.2.4 วิธีการดำเนินการวิจัยระยะที่ 2

จากการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนในระยะที่ 1 นำมาพัฒนาตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิคเหมืองข้อมูล 2 เทคนิค ได้แก่ เทคนิค Decision Tree และ เทคนิค Naïve Bayes แสดงขั้นตอนการพยากรณ์ผลการเรียนดังภาพที่ 3-19



ภาพที่ 3-19 ขั้นตอนการสร้างตัวแบบพยากรณ์ผลการเรียน

3.2.4.1 วิเคราะห์ข้อมูลและจัดรูปแบบข้อมูลให้อยู่ในรูปแบบพร้อมใช้งาน เพื่อนำข้อมูลเข้าสู่การทำเหมืองข้อมูล จากผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษา พบว่าจำนวนปัจจัยที่ให้ค่าความถูกต้องสูงสุดของการพัฒนาตัวแบบพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Decision Tree จำนวน 6 ปัจจัย ได้แก่ ผลการเรียนเฉลี่ยปี 1 เทอม 2 (grade_T2) ผลการเรียนเฉลี่ยปี 2 เทอม 1 (grade_T3) ผลการเรียนเฉลี่ยปี 2 เทอม 2 (grade_T4) ผลการเรียนเฉลี่ยปี 1 เทอม 1 (grade_T1) ผลการเรียนรายวิชา 202507 และผลการเรียนเฉลี่ยปี 3 เทอม 1 (grade_T5) และปัจจัยที่ให้ค่าความถูกต้องสูงสุดของการพัฒนาตัวแบบพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Naïve Bayes จำนวน 14 ปัจจัย ได้แก่ ผลการเรียนเฉลี่ยปี 1 เทอม 2 (grade_T2) ผลการเรียนเฉลี่ยปี 2 เทอม 1 (grade_T3) ผลการเรียนเฉลี่ยปี 2 เทอม 2 (grade_T4) ผลการเรียนเฉลี่ยปี 1 เทอม 1 (grade_T1) ผลการเรียนรายวิชา 202507 ผลการเรียนเฉลี่ยปี 3 เทอม 1 (grade_T5) ผลการเรียนรายวิชา 202506 ผลการเรียนรายวิชา 105131 ผลการเรียนรายวิชา 202101 ผลการเรียนรายวิชา 901101 ผลการเรียนรายวิชา 202211 ผลการเรียนรายวิชา 202207 ผลการเรียนรายวิชา 101501 และผลการเรียนรายวิชา 903301 สรุปได้ดังตารางที่ 3-15

ตารางที่ 3-15 รายการข้อมูลสำหรับพัฒนาตัวแบบพยากรณ์ผลการเรียนของนักศึกษา

แอททริบิวต์	คำอธิบาย	รายละเอียด
grade_T1	ผลการเรียนเฉลี่ย ปี 1 เทอม 1	VH = 3.50-4.00 H = 3.00-3.49 M = 2.50-2.99 L = 2.00-2.49 VL = 0.00-1.99
grade_T2	ผลการเรียนเฉลี่ย ปี 1 เทอม 2	
grade_T3	ผลการเรียนเฉลี่ย ปี 2 เทอม 1	
grade_T4	ผลการเรียนเฉลี่ย ปี 2 เทอม 2	
grade_T5	ผลการเรียนเฉลี่ย ปี 3 เทอม 1	
ผลการเรียน 9 วิชา	ผลการเรียน ได้แก่ 101501 105131 202101 202207 202211 202506 202507 901101 903301	4.0 = A 3.5 = B+ 3.0 = B 2.5 = C+ 2.0 = C 1.5 = D+ 1.0 = D 0.0 = F
GPAX	เกรดเฉลี่ยจบ	Low=2.25-2.80 Middle=2.81-3.19 Good=3.20-3.53 High=3.54-4.00

3.2.4.2 คัดเลือกคุณลักษณะ (Feature Selection) ด้วย Information Gain เป็นการประเมินประสิทธิภาพคุณลักษณะของข้อมูลแต่ละแอททริบิวต์ว่ามีความเหมาะสมกับการวิเคราะห์ข้อมูลมากน้อยเพียงใดซึ่งผลการเรียงลำดับความสำคัญ (Ranking) ระดับความสำคัญ เทคนิคเหมืองข้อมูล 2 เทคนิค ได้แก่ เทคนิค Decision Tree และเทคนิค Naïve Bayes โดยปัจจัยสำหรับพยากรณ์ผลการเรียนด้วยเทคนิค Decision Tree ประกอบด้วย 6 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอม2 เกรดเฉลี่ยเทอม3 เกรดเฉลี่ยเทอม4 เกรดเฉลี่ยเทอม1 ผลการเรียนวิชา 202507 และเกรดเฉลี่ยเทอม1 เป็นต้น และเทคนิค Naïve Bayes ประกอบด้วย 14 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา 202101 ผลการเรียนวิชา 901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301 เป็นต้น

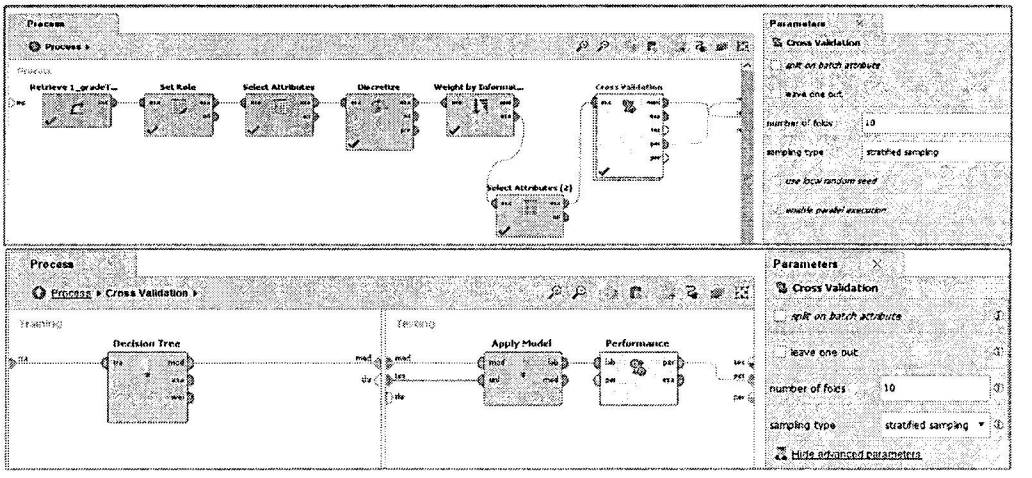
3.2.4.3 ทำการแบ่งส่วนข้อมูลออกเป็นหลายส่วนเท่า ๆ กันเพื่อทำการทดสอบ (Cross-Validation) ประกอบด้วยข้อมูลส่วน Training Set และ Testing Set ด้วยการใช้วิธี 10-fold Cross-validation ซึ่งแบ่งข้อมูลออกเป็น 10 ส่วนเท่า ๆ กัน แล้วนำข้อมูลที่แบ่งไว้ใช้ในการสร้างตัว

แบบพยากรณ์และการทดสอบตัวแบบพยากรณ์ เวียนกันไปจนครบทุกส่วนของข้อมูล ซึ่งกระบวนการทดสอบจะทำการให้ข้อมูลชุดที่ 1 เป็นข้อมูลทดสอบ (Testing Set) และข้อมูลชุดที่ 2-10 ชุดเทรนนิ่ง (Training Set) วนไปเรื่อย ๆ จนกระทั่งชุดที่ 10 เป็นชุดทดสอบ (Testing Set) และชุดที่ 1-9 เป็นชุดเทรนนิ่ง (Training Set) ดังภาพที่ 3-20

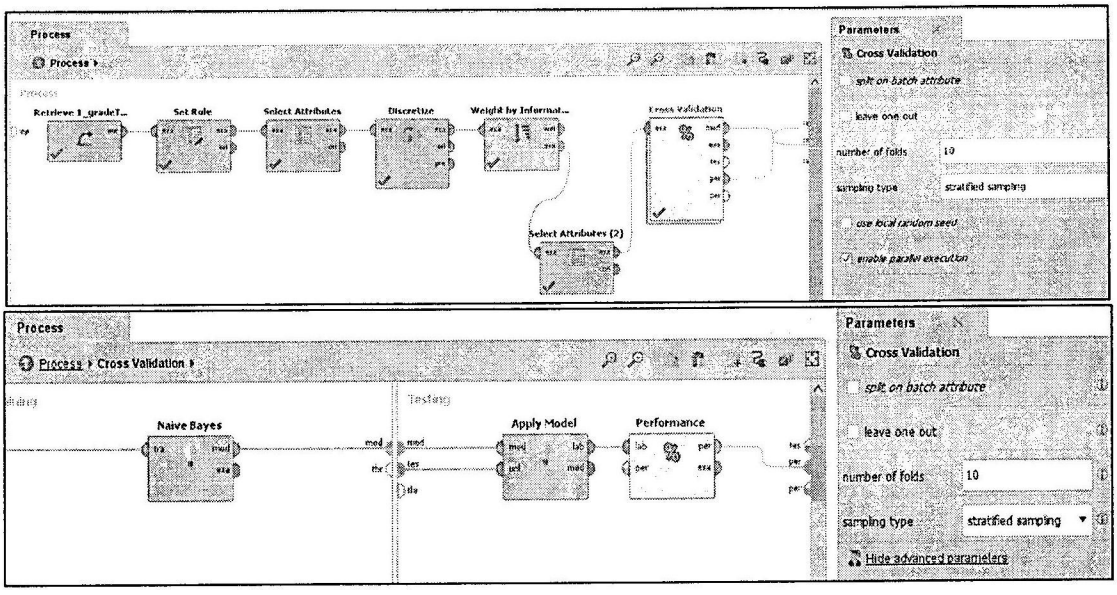
	Training Set									Testing Set
รอบที่1	2	3	4	5	6	7	8	9	10	1
รอบที่2	1	3	4	5	6	7	8	9	10	2
รอบที่3	1	2	4	5	6	7	8	9	10	3
รอบที่4	1	2	3	5	6	7	8	9	10	4
รอบที่5	1	2	3	4	6	7	8	9	10	5
รอบที่6	1	2	3	4	5	7	8	9	10	6
รอบที่7	1	2	3	4	5	6	8	9	10	7
รอบที่8	1	2	3	4	5	6	7	9	10	8
รอบที่9	1	2	3	4	5	6	7	8	10	9
รอบที่10	1	2	3	4	5	6	7	8	9	10

ภาพที่ 3-20 การแบ่งกลุ่มข้อมูลทดสอบ (10-fold Cross-validation)

3.2.4.4 การพัฒนาตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree) และเทคนิค Naïve Bayes เครื่องมือที่ใช้ทำเหมืองข้อมูลด้วยโปรแกรม RapidMiner พบว่าเทคนิค Decision Tree จากคุณลักษณะ 6 คุณลักษณะซึ่งให้ค่าความถูกต้องของการพยากรณ์ร้อยละ 73.86 และเทคนิค Naïve Bayes จากคุณลักษณะ 14 คุณลักษณะซึ่งให้ค่าความถูกต้องของการพยากรณ์ร้อยละ 84.33



ภาพที่ 3-21 การพัฒนาตัวแบบพยากรณ์ผลการเรียนโดยใช้เทคนิค Decision Tree



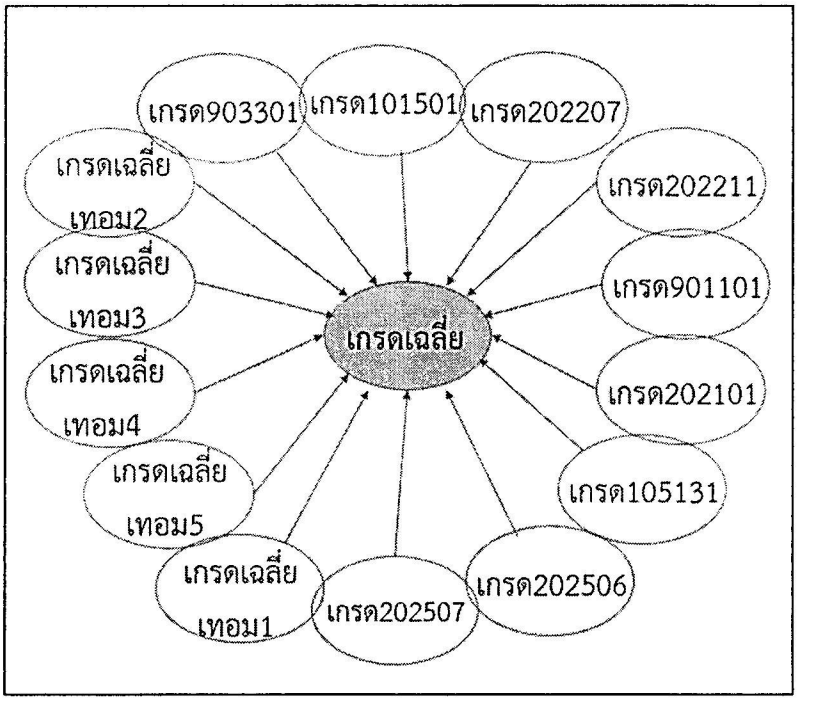
ภาพที่ 3-22 การพัฒนาตัวแบบพยากรณ์ผลการเรียนโดยใช้เทคนิค Naïve Bayes

3.2.4.5 พัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

เนื่องจากผลการเปรียบเทียบเทคนิคการพยากรณ์ผลการเรียนของนักศึกษา พบว่า ตัวแบบพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Naïve Bayes ให้ความถูกต้องสูงกว่าเทคนิคการพยากรณ์ผลการเรียนของนักศึกษาด้วยเทคนิค Decision Tree ดังนั้นผู้วิจัยใช้ตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes สำหรับการสร้างตัวพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี โดยกลุ่มตัวอย่างสำหรับการสร้างตัวแบบพยากรณ์ผลการเรียนอัจฉริยะ ได้แก่ นักศึกษาสาขาภาษาอังกฤษที่จบการศึกษาแล้ว จำนวน 141 คน และกลุ่มตัวอย่างสำหรับการทดสอบตัวแบบพยากรณ์ผลการเรียนอัจฉริยะ ได้แก่ นักศึกษาสาขาวิชาภาษาอังกฤษชั้นปีที่ 3 จำนวน 40 คน โดยแอททริบิวต์ที่สำคัญสำหรับการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษามีจำนวนทั้งหมด 14 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา 202101 ผลการเรียนวิชา 901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301

ผลการสร้างตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีสามารถแสดงสมการความน่าจะเป็นสำหรับการพยากรณ์ผลการเรียนอัจฉริยะด้วยเทคนิคนาอีฟเบย์ ดังสมการที่ (3-1)

$$P(C|A) = \frac{P(a_1|C) * P(a_2|C) * \ldots\ldots * P(a_{14}|C) * \mathrm{P(C)}}{P(A)} \qquad (3\text{-}1)$$



ภาพที่ 3-23 แอททริบิวต์สำหรับการพยากรณ์ผลการเรียนของนักศึกษา

ขั้นตอนการสร้างตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี มีขั้นตอนดังนี้

3.2.4.5.1 เทคนิคการจำแนกประเภทข้อมูลด้วยวิธี Naïve Bayes เป็นวิธีหนึ่งที่นิยมเนื่องจากการสร้างโมเดลไม่ซับซ้อน โดยเป็นการใช้ทฤษฎีความน่าจะเป็น (Probability) ในสมการที่ (3-2)

$$P(A|B) = \frac{p(A \cap B)}{P(B)} \tag{3-2}$$

กำหนดให้

$P(A|B)$ หมายถึง ความน่าจะเป็นที่จะเกิดเหตุการณ์ B ขึ้นก่อนแล้วจะมีเหตุการณ์ A ตามมา

$p(A \cap B)$ หมายถึง ความน่าจะเป็นที่เกิดเหตุการณ์ A และ B พร้อมกัน

$P(B)$ หมายถึง ความน่าจะเป็นที่เหตุการณ์ B เกิดขึ้น

$$P(B|A) = \frac{p(A \cap B)}{P(A)} \tag{3-3}$$

กำหนดให้

$P(B|A)$ หมายถึง ความน่าจะเป็นที่จะเกิดเหตุการณ์ A ขึ้นก่อนแล้วจะมีเหตุการณ์ B ตามมา

$p(A \cap B)$ หมายถึง ความน่าจะเป็นที่เกิดเหตุการณ์ A และ B พร้อมกัน

$P(A)$ หมายถึง ความน่าจะเป็นที่เหตุการณ์ A เกิดขึ้น

จากสมการที่ (3-2) และสมการที่ (3-3) มีค่าของ $p(A \cap B)$ ทั้งสองสมการ ดังนั้นสามารถเขียนสมการของ $p(A \cap B)$ ได้ดังนี้

$$p(A \cap B) = P(A|B) * P(B) = P(B|A) * P(A) \tag{3-4}$$

หรือ

$$P(B|A) = \frac{P(A|B) * P(B)}{P(A)} \tag{3-5}$$

จากสมการที่ (3-5) จะพบว่านี่คือสมการของ Bayes theorem หรือทฤษฎีของเบย์ ซึ่งการหาความน่าจะเป็นของข้อมูลสามารถเปลี่ยน A และ B ให้เป็น A และ C โดย A หมายถึง แอททริบิวต์ และ C หมายถึง คลาสในสมการที่ (3-6)

$$P(C|A) = \frac{P(A|C) * \mathrm{P(C)}}{P(A)} \tag{3-6}$$

กำหนดให้

$P(C|A)$ หมายถึง ความน่าจะเป็นที่จะมีแอททริบิวต์เป็น A จะมีคลาส C

$P(A|C)$ หมายถึง ความน่าจะเป็นของเทรนนิ่งดาต้าที่มีคลาส C และมีแอททริบิวต์ A โดย $A = a_1 \cap a_2 \cap \ldots \cap a_M$ และ M คือ จำนวนแอททริบิวต์ในเทรนนิ่งดาต้า

$P(C)$ หมายถึง ความน่าจะเป็นของคลาส C

$P(A)$ หมายถึง ค่าความน่าจะเป็นของ $A = a_1 \cap a_2 \cap \ldots \cap a_M$

จากสมการที่ (3-2) สามารถเขียนในอีกรูปแบบหนึ่งโดยแทนค่าของแอททริบิวต์ A ตามจำนวนของแอททริบิวต์ ซึ่งจะได้สมการที่ (3-7) หรือสมการที่ (3-8) ดังนี้

$$P(C|A) = \frac{P(a_1|C) * P(a_2|C) * \ldots\ldots * P(a_M|C) * \mathrm{P(C)}}{P(A)} \quad (3\text{-}7)$$

หรือ

$$P(C|A) = \frac{P(a_1|C) * P(a_2|C) * \ldots\ldots * P(a_{14}|C) * \mathrm{P(C)}}{P(A)} \quad (3\text{-}8)$$

จากสมการที่ (3-8) เป็นสมการความน่าจะเป็นสำหรับการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี โดยสามารถพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาของนักศึกษาปัจจุบันได้ด้วยวิธีการนำปัจจัยการพยากรณ์ที่เกี่ยวข้องทั้งหมด จำนวน 14 ปัจจัยแทนค่าในสมการความน่าจะเป็นดังกล่าว

3.2.4.5.2 ขั้นตอนการคำนวณหาค่าความน่าจะเป็นของแต่ละคลาสในสมการของ Bayes แสดงดังสมการที่ (3-7) มีขั้นตอนดังต่อไปนี้

ก) คำนวณความน่าจะเป็น P(C) ของคลาส GPAX

P(GPAX = High) = 38/141 = 0.269

P(GPAX = Good) = 36/141 = 0.255

P(GPAX = Middle) = 45/141 = 0.319

P(GPAX = Low) = 22/141 = 0.156

ข) คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2) เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2) เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2) เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2) เมื่อ GPAX = Low

ค) คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 1

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 1 (Grade_T3) เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 1 (Grade_T3) เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 1 (Grade_T3) เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 1 (Grade_T3) เมื่อ GPAX = Low

ง) คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 2

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 2 (Grade_T4) เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 2 (Grade_T4) เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 2 (Grade_T4) เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 2 (Grade_T4) เมื่อ GPAX = Low

จ) คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 1

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 1 (Grade_T1) เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 1 (Grade_T1) เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 1 (Grade_T1) เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 1 (Grade_T1) เมื่อ GPAX = Low

ฉ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202507

ความน่าจะเป็นของผลการเรียนรายวิชา 202507 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 202507 เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 202507 เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 202507 เมื่อ GPAX = Low

ช) คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 3 เทอม 1

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 3 เทอม 1 (Grade_T5) เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 3 เทอม 1 (Grade_T5) เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 3 เทอม 1 (Grade_T5) เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 3 เทอม 1 (Grade_T5) เมื่อ GPAX = Low

ซ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202506

ความน่าจะเป็นของผลการเรียนรายวิชา 202506 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 202506 เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 202506 เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 202506 เมื่อ GPAX = Low

ฌ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 105131

ความน่าจะเป็นของผลการเรียนรายวิชา 105131 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 105131 เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 105131 เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 105131 เมื่อ GPAX = Low

ญ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202101

ความน่าจะเป็นของผลการเรียนรายวิชา 202101 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 202101 เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 202101 เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 202101 เมื่อ GPAX = Low

ฎ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 901101

ความน่าจะเป็นของผลการเรียนรายวิชา 901101 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 901101 เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 901101 เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 901101 เมื่อ GPAX = Low

ฏ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202211

ความน่าจะเป็นของผลการเรียนรายวิชา 202211 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 202211 เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 202211 เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 202211 เมื่อ GPAX = Low

ฐ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202207

ความน่าจะเป็นของผลการเรียนรายวิชา 202207 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 202207 เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 202207 เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 202207 เมื่อ GPAX = Low

ฑ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 101501

ความน่าจะเป็นของผลการเรียนรายวิชา 101501 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 101501 เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 101501 เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 101501 เมื่อ GPAX = Low

ฒ) คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 903301

ความน่าจะเป็นของผลการเรียนรายวิชา 903301 เมื่อ GPAX = High

ความน่าจะเป็นของผลการเรียนรายวิชา 903301เมื่อ GPAX = Good

ความน่าจะเป็นของผลการเรียนรายวิชา 903301เมื่อ GPAX = Middle

ความน่าจะเป็นของผลการเรียนรายวิชา 903301เมื่อ GPAX = Low

จากการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะ จำนวน 14 ปัจจัย สามารถหาความน่าจะเป็นดังรายละเอียดในภาคผนวก ค และสามารถสรุปได้ดังตารางที่ 3-16 และตารางที่ 3-17 และเรียกตารางดังกล่าวว่าโมเดล Naïve Bayes ซึ่งเราสามารถนำโมเดลดังกล่าวไปพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี จากสมการของ Bayes ได้เลย

ตารางที่ 3-16 ความน่าจะเป็น P(C) ของคลาส GPAX

attribute	Class	P(C)
GPAX	High	0.269
	Good	0.255
	Middle	0.319
	Low	0.156

ตารางที่ 3-17 โมเดล Naïve Bayes

attribute	Parameter	High	Good	Middle	Low
101501	value=0.0	0.000	0.000	0.000	0.000
101501	value=1.0	0.000	0.000	0.000	0.000
101501	value=1.5	0.000	0.000	0.000	0.000
101501	value=2.0	0.000	0.000	0.022	0.045
101501	value=2.5	0.000	0.000	0.044	0.272
101501	value=3.0	0.026	0.111	0.310	0.498

ตารางที่ 3-17 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
101501	value=3.5	0.131	0.332	0.554	0.136
101501	value=4.0	0.840	0.554	0.066	0.045
105131	value=0.0	0.000	0.000	0.000	0.000
105131	value=1.0	0.000	0.027	0.177	0.453
105131	value=1.5	0.000	0.083	0.355	0.362
105131	value=2.0	0.052	0.249	0.310	0.136
105131	value=2.5	0.157	0.388	0.088	0.045
105131	value=3.0	0.341	0.194	0.066	0.000
105131	value=3.5	0.236	0.055	0.000	0.000
105131	value=4.0	0.210	0.000	0.000	0.000
202101	value=0.0	0.000	0.000	0.000	0.000
202101	value=1.0	0.000	0.000	0.000	0.000
202101	value=1.5	0.000	0.000	0.000	0.136
202101	value=2.0	0.000	0.000	0.044	0.453
202101	value=2.5	0.000	0.027	0.244	0.226
202101	value=3.0	0.052	0.222	0.333	0.045
202101	value=3.5	0.184	0.360	0.333	0.136
202101	value=4.0	0.762	0.388	0.044	0.000
202207	value=0.0	0.000	0.000	0.000	0.000
202207	value=1.0	0.000	0.000	0.000	0.000
202207	value=1.5	0.000	0.000	0.000	0.000
202207	value=2.0	0.000	0.027	0.088	0.544
202207	value=2.5	0.026	0.194	0.488	0.362
202207	value=3.0	0.315	0.443	0.333	0.090
202207	value=3.5	0.210	0.305	0.088	0.000
202207	value=4.0	0.446	0.027	0.000	0.000
202211	value=0.0	0.000	0.000	0.000	0.000
202211	value=1.0	0.000	0.000	0.000	0.090
202211	value=1.5	0.000	0.000	0.000	0.045
202211	value=2.0	0.000	0.000	0.022	0.317
202211	value=2.5	0.000	0.000	0.155	0.226
202211	value=3.0	0.052	0.194	0.488	0.317
202211	value=3.5	0.341	0.416	0.244	0.000
202211	value=4.0	0.604	0.388	0.088	0.000

ตารางที่ 3-17 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
202506	value=0.0	0.000	0.000	0.000	0.181
202506	value=1.0	0.000	0.055	0.133	0.136
202506	value=1.5	0.000	0.083	0.310	0.362
202506	value=2.0	0.000	0.222	0.266	0.272
202506	value=2.5	0.079	0.222	0.222	0.045
202506	value=3.0	0.367	0.305	0.066	0.000
202506	value=3.5	0.262	0.083	0.000	0.000
202506	value=4.0	0.289	0.027	0.000	0.000
202507	value=0.0	0.000	0.027	0.066	0.226
202507	value=1.0	0.000	0.055	0.333	0.362
202507	value=1.5	0.000	0.083	0.222	0.045
202507	value=2.0	0.026	0.277	0.199	0.317
202507	value=2.5	0.079	0.222	0.177	0.045
202507	value=3.0	0.184	0.194	0.000	0.000
202507	value=3.5	0.289	0.027	0.000	0.000
202507	value=4.0	0.420	0.111	0.000	0.000
901101	value=0.0	0.000	0.000	0.000	0.000
901101	value=1.0	0.000	0.000	0.000	0.045
901101	value=1.5	0.000	0.000	0.000	0.226
901101	value=2.0	0.000	0.027	0.111	0.136
901101	value=2.5	0.000	0.111	0.377	0.498
901101	value=3.0	0.210	0.194	0.399	0.090
901101	value=3.5	0.420	0.527	0.111	0.000
901101	value=4.0	0.367	0.138	0.000	0.000
903301	value=0.0	0.000	0.000	0.000	0.000
903301	value=1.0	0.000	0.000	0.000	0.045
903301	value=1.5	0.000	0.000	0.000	0.000
903301	value=2.0	0.000	0.000	0.000	0.090
903301	value=2.5	0.000	0.027	0.044	0.181
903301	value=3.0	0.026	0.194	0.399	0.362
903301	value=3.5	0.210	0.305	0.532	0.317
903301	value=4.0	0.762	0.471	0.022	0.000

ตารางที่ 3-17 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
grade_T1	value=VL	0.000	0.000	0.000	0.181
grade_T1	value=L	0.000	0.027	0.088	0.408
grade_T1	value=M	0.000	0.055	0.577	0.317
grade_T1	value=H	0.105	0.610	0.266	0.091
grade_T1	value=VH	0.893	0.305	0.066	0.000
grade_T2	value=VL	0.000	0.000	0.000	0.000
grade_T2	value=L	0.000	0.000	0.022	0.091
grade_T2	value=M	0.000	0.000	0.155	0.726
grade_T2	value=H	0.000	0.471	0.821	0.181
grade_T2	value=VH	0.999	0.527	0.000	0.000
grade_T3	value=VL	0.000	0.000	0.000	0.045
grade_T3	value=L	0.000	0.000	0.066	0.544
grade_T3	value=M	0.000	0.083	0.666	0.363
grade_T3	value=H	0.210	0.804	0.266	0.045
grade_T3	value=VH	0.788	0.111	0.000	0.000
grade_T4	value=VL	0.000	0.000	0.000	0.000
grade_T4	value=L	0.000	0.000	0.000	0.317
grade_T4	value=M	0.000	0.055	0.510	0.635
grade_T4	value=H	0.105	0.610	0.466	0.045
grade_T4	value=VH	0.893	0.333	0.022	0.000
grade_T5	value=VL	0.000	0.000	0.000	0.000
grade_T5	value=L	0.000	0.000	0.000	0.181
grade_T5	value=M	0.000	0.000	0.310	0.726
grade_T5	value=H	0.131	0.499	0.621	0.091
grade_T5	value=VH	0.867	0.499	0.066	0.000

3.2.4.5.3 การพยากรณ์ผลการเรียน (GPAX) ของนักศึกษาที่ไม่ทราบคลาส สามารถคำนวณความน่าจะเป็นของข้อมูลใหม่ของนักศึกษาด้วยเทคนิค Naïve Bayes โดยประกอบด้วยข้อมูลพื้นฐาน 14 แอททริบิวต์ที่มีผลต่อการทำนาย ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา 202101 ผลการเรียนวิชา901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301 โดยข้อมูลที่ต้องการพยากรณ์ผลการเรียน (GPAX) แสดงดังตารางที่ 3-18

ตารางที่ 3-18 ข้อมูลนักศึกษาที่ยังไม่ทราบคลาสคำตอบ

Attribute/No	1	2	3
101501	2.5	3.5	4.0
105131	2.0	2.5	3.0
202101	3.0	3.0	3.5
202207	3.0	3.5	4.0
202211	2.5	2.0	4.0
202506	3.5	3.5	4.0
202507	2.0	3.0	4.0
901101	4.0	4.0	3.5
903301	3.5	3.5	3.5
Grade_T1	M	H	VH
Grade_T2	H	H	VH
Grade_T3	M	M	VH
Grade_T4	VH	VH	VH
Grade_T5	H	H	VH
GPAX	?	?	?

จากตารางที่ 3-18 โมเดล Naïve Bayes สังเกตพบว่าค่าความน่าจะเป็นบางค่าเป็น 0 นั่นคือไม่มีรูปแบบของแอททริบิวต์เกิดขึ้นในข้อมูลเทรนนิ่งดาต้า (Training Data) ดังนั้นการใช้งานโมเดล Naïve Bayes ที่มีความน่าจะเป็นเท่ากับ 0 จะทำให้ค่าที่พยากรณ์เกิดปัญหาเนื่องจากผลลัพธ์ที่ได้จะเท่ากับ 0 ตัวอย่างเช่น P(101501=2.5|GPAX=High) = 0/38 ผลลัพธ์ที่ได้จะเท่ากับ 0 ซึ่งปัญหาดังกล่าวสามารถแก้ไขได้โดยใช้เทคนิควิธีที่เรียกว่า Laplace Smoothing ซึ่งเป็นเทคนิคการปรับให้เรียบใช้กับปัญหาความน่าจะเป็นที่เป็นศูนย์ในเทคนิค Naïve Bayes ด้วยการเพิ่มความถี่ของข้อมูลเข้าไปอีก 1 ค่า ดังสมการ

$$P(C|A) = \frac{number\ of\ A + \propto}{N + \propto * K} \quad (3\text{-}9)$$

เมื่อ

$\propto$ หมายถึง ค่าที่เพิ่มขึ้น ($\propto$ = 1)

A หมายถึง จำนวนแอททริบิวต์ของ a_x ในคลาส C

K หมายถึง ค่าความเป็นไปได้ทั้งหมดของ A

N หมายถึง จำนวนแอทริบิวต์ทั้งหมดของ A ในคลาส C

แสดงตัวอย่างการปรับค่า Laplace smoothing ของ P(101501=2.5|GPAX=High) ด้วยการเพิ่มค่าความถี่ข้อมูลอีก 1 จะได้เป็น P(101501=2.5|GPAX=High) = 0+1/(38+(1*8)) ซึ่งจะพบว่าค่าความน่าจะเป็นเท่ากับ 0.0217

$$P(101501=0.0|GPAX=High) = \frac{0+1}{38+1*8} = 0.0217$$

$$P(101501=1.0|GPAX=High) = \frac{0+1}{38+1*8} = 0.0217$$

$$P(101501=1.5|GPAX=High) = \frac{0+1}{38+1*8} = 0.0217$$

$$P(101501=2.0|GPAX=High) = \frac{0+1}{38+1*8} = 0.0217$$

$$P(101501=2.5|GPAX=High) = \frac{0+1}{38+1*8} = 0.0217$$

$$P(101501=3.0|GPAX=High) = \frac{1+1}{38+1*8} = 0.0434$$

$$P(101501=3.4|GPAX=High) = \frac{5+1}{38+1*8} = 0.1304$$

$$P(101501=4.0|GPAX=High) = \frac{32+1}{38+1*8} = 0.7174$$

Naïve Bayes เป็นการจำแนกประเภทข้อมูลวิธีหนึ่งที่มีประสิทธิภาพและมีการทำงานที่ไม่ซับซ้อน ปัจจุบันมีการประยุกต์ใช้ Naïve Bayes อย่างแพร่หลาย และสามารถแสดงผลการปรับเรียบด้วยวิธี Laplace Smoothing แสดงดังตารางที่ 3-19 และตารางที่ 3-20

ตารางที่ 3-19 ความน่าจะเป็น P(C) ของคลาส GPAX หลังปรับ Laplace smoothing

attribute	Class	P(C)
GPAX	High	0.266
	Good	0.254
	Middle	0.307
	Low	0.171

ตารางที่ 3-20 โมเดล Naïve Bayes จากการปรับ Laplace smoothing

attribute	Parameter	High	Good	Middle	Low
101501	value=0.0	0.021	0.022	0.018	0.033
101501	value=1.0	0.021	0.022	0.018	0.033
101501	value=1.5	0.021	0.022	0.018	0.033
101501	value=2.0	0.021	0.022	0.037	0.066
101501	value=2.5	0.021	0.022	0.056	0.233
101501	value=3.0	0.043	0.113	0.283	0.400
101501	value=3.5	0.130	0.295	0.490	0.133

ตารางที่ 3-20 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
101501	value=4.0	0.717	0.477	0.075	0.066
105131	value=0.0	0.021	0.022	0.018	0.033
105131	value=1.0	0.021	0.045	0.169	0.366
105131	value=1.5	0.021	0.090	0.32	0.300
105131	value=2.0	0.065	0.227	0.283	0.133
105131	value=2.5	0.152	0.34	0.094	0.066
105131	value=3.0	0.304	0.181	0.075	0.033
105131	value=3.5	0.217	0.068	0.018	0.033
105131	value=4.0	0.195	0.022	0.018	0.033
202101	value=0.0	0.021	0.022	0.018	0.033
202101	value=1.0	0.021	0.022	0.018	0.033
202101	value=1.5	0.021	0.022	0.018	0.133
202101	value=2.0	0.021	0.022	0.056	0.366
202101	value=2.5	0.021	0.045	0.226	0.200
202101	value=3.0	0.065	0.204	0.301	0.066
202101	value=3.5	0.173	0.318	0.301	0.133
202101	value=4.0	0.652	0.34	0.056	0.033
202207	value=0.0	0.021	0.022	0.018	0.033
202207	value=1.0	0.021	0.022	0.018	0.033
202207	value=1.5	0.021	0.022	0.018	0.033
202207	value=2.0	0.021	0.045	0.094	0.433
202207	value=2.5	0.043	0.181	0.433	0.300
202207	value=3.0	0.282	0.386	0.301	0.100
202207	value=3.5	0.195	0.272	0.094	0.033
202207	value=4.0	0.391	0.045	0.018	0.033
202211	value=0.0	0.021	0.022	0.018	0.033
202211	value=1.0	0.021	0.022	0.018	0.100
202211	value=1.5	0.021	0.022	0.018	0.066
202211	value=2.0	0.021	0.022	0.037	0.266
202211	value=2.5	0.021	0.022	0.150	0.200
202211	value=3.0	0.065	0.181	0.433	0.266
202211	value=3.5	0.304	0.363	0.226	0.033
202211	value=4.0	0.521	0.340	0.094	0.033

ตารางที่ 3-20 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
202506	value=0.0	0.021	0.022	0.018	0.166
202506	value=1.0	0.021	0.068	0.132	0.133
202506	value=1.5	0.021	0.09	0.283	0.300
202506	value=2.0	0.021	0.204	0.245	0.233
202506	value=2.5	0.086	0.204	0.207	0.066
202506	value=3.0	0.326	0.272	0.075	0.033
202506	value=3.5	0.239	0.090	0.018	0.033
202506	value=4.0	0.260	0.045	0.018	0.033
202507	value=0.0	0.021	0.045	0.075	0.200
202507	value=1.0	0.021	0.068	0.301	0.300
202507	value=1.5	0.021	0.090	0.207	0.066
202507	value=2.0	0.043	0.250	0.188	0.266
202507	value=2.5	0.086	0.204	0.169	0.066
202507	value=3.0	0.173	0.181	0.018	0.033
202507	value=3.5	0.260	0.045	0.018	0.033
202507	value=4.0	0.369	0.113	0.018	0.033
901101	value=0.0	0.021	0.022	0.018	0.033
901101	value=1.0	0.021	0.022	0.018	0.066
901101	value=1.5	0.021	0.022	0.018	0.200
901101	value=2.0	0.021	0.045	0.113	0.133
901101	value=2.5	0.021	0.113	0.339	0.400
901101	value=3.0	0.195	0.181	0.358	0.100
901101	value=3.5	0.369	0.454	0.113	0.033
901101	value=4.0	0.326	0.136	0.018	0.033
903301	value=0.0	0.021	0.022	0.018	0.033
903301	value=1.0	0.021	0.022	0.018	0.066
903301	value=1.5	0.021	0.022	0.018	0.033
903301	value=2.0	0.021	0.022	0.018	0.100
903301	value=2.5	0.021	0.045	0.056	0.166
903301	value=3.0	0.043	0.181	0.358	0.300
903301	value=3.5	0.195	0.272	0.471	0.266
903301	value=4.0	0.652	0.409	0.037	0.033

ตารางที่ 3-20 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
grade_T1	value=VL	0.023	0.024	0.020	0.185
grade_T1	value=L	0.023	0.048	0.100	0.370
grade_T1	value=M	0.023	0.073	0.540	0.296
grade_T1	value=H	0.116	0.560	0.260	0.111
grade_T1	value=VH	0.813	0.292	0.080	0.037
grade_T2	value=VL	0.023	0.024	0.020	0.037
grade_T2	value=L	0.023	0.024	0.040	0.111
grade_T2	value=M	0.023	0.024	0.160	0.629
grade_T2	value=H	0.023	0.439	0.760	0.185
grade_T2	value=VH	0.906	0.487	0.020	0.037
grade_T3	value=VL	0.023	0.024	0.020	0.074
grade_T3	value=L	0.023	0.024	0.080	0.481
grade_T3	value=M	0.023	0.097	0.620	0.333
grade_T3	value=H	0.209	0.731	0.260	0.074
grade_T3	value=VH	0.72	0.121	0.020	0.037
grade_T4	value=VL	0.023	0.024	0.020	0.037
grade_T4	value=L	0.023	0.024	0.020	0.296
grade_T4	value=M	0.023	0.073	0.480	0.555
grade_T4	value=H	0.116	0.560	0.440	0.074
grade_T4	value=VH	0.813	0.317	0.040	0.037
grade_T5	value=VL	0.023	0.024	0.020	0.037
grade_T5	value=L	0.023	0.024	0.020	0.185
grade_T5	value=M	0.023	0.024	0.300	0.629
grade_T5	value=H	0.139	0.463	0.580	0.111
grade_T5	value=VH	0.790	0.463	0.080	0.037

จากตารางที่ 3-20 โมเดล Naïve Bayes จากการปรับค่า Laplace Smoothing ซึ่งเป็นการทำให้แอททริบิวต์ที่เกิดขึ้นในเทรนนิ่งดาต้า (Training Data) ไม่เป็น 0 ด้วยการเพิ่มความถี่ของข้อมูลเข้าไปอีก 1 ค่า ดังนั้นการใช้โมเดลพยากรณ์ผลการเรียนเฉลี่ยนักศึกษาที่ไม่ทราบคลาสคำตอบ ดังตารางที่ 3-18 สามารถคำนวณหาค่าความน่าจะเป็นผลการเรียนเฉลี่ยด้วยสมการ Bayes ดังสมการที่ (3-8)

ตารางที่ 3-21 โมเดลพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.1

attribute	Parameter	High	Good	Middle	Low
101501	value=2.5	0.021	0.022	0.056	0.233
105131	value=2.0	0.065	0.227	0.283	0.133
202101	value=3.0	0.065	0.204	0.301	0.066
202207	value=3.0	0.282	0.386	0.301	0.100
202211	value=2.5	0.021	0.022	0.150	0.200
202506	value=3.5	0.239	0.090	0.018	0.033
202507	value=2.0	0.043	0.250	0.188	0.266
901101	value=4.0	0.326	0.136	0.018	0.033
903301	value=3.5	0.195	0.272	0.471	0.266
grade_T1	value=M	0.023	0.073	0.540	0.296
grade_T2	value=H	0.023	0.439	0.760	0.185
grade_T3	value=M	0.023	0.097	0.620	0.333
grade_T4	value=VH	0.813	0.317	0.040	0.037
grade_T5	value=H	0.139	0.463	0.580	0.111

ก) การพยากรณ์ผลการเรียนเฉลี่ยของนักศึกษา No1 จากการแทนค่าในสมการของ Bayes

P(GPAX=High|A=No1) = (P(101501=2.5|GPAX=High) * P(105131=2.0|GPAX=High) * P(202101=3.0|GPAX=High) * P(202207=3.0|GPAX=High) * P(202211=2.5|GPAX=High) * P(202506=3.5|GPAX=High) * P(202507=2.0|GPAX=High) * P(901101=4.0|GPAX=High) * P(903301=3.5|GPAX=High) * P(Grade_T1=M|GPAX=High) * P(Grade_T2=H|GPAX=High) * P(Grade_T3=M|GPAX=High) * P(Grade_T4=VH|GPAX=High) * P(Grade_T5=H|GPAX=High) * P(GPAX=High)) / P(A)

= (0.021 * 0.065 * 0.065 * 0.282 * 0.021 * 0.239 * 0.043 * 0.326 * 0.195 * 0.023 * 0.023 * 0.023 * 0.813 * 0.139 * 0.266) / 1.2732E-11

= 1.03988E-05

P(GPAX=Good|A=No1)= (P(101501=2.5|GPAX= Good) * P(105131=2.0|GPAX= Good) * P(202101=3.0|GPAX= Good) * P(202207=3.0|GPAX= Good) * P(202211=2.5|GPAX= Good) * P(202506=3.5|GPAX= Good) *

P(202507=2.0|GPAX= Good) * P(901101=4.0|GPAX= Good) * P(903301=3.5|GPAX= Good) * P(Grade_T1=M|GPAX= Good) * P(Grade_T2=H|GPAX= Good) * P(Grade_T3=M|GPAX= Good) * P(Grade_T4=VH|GPAX= Good) * P(Grade_T5=H|GPAX= Good) * P(GPAX= Good)) / P(A)

= (0.022 * 0.227 * 0.204 * 0.386 * 0.022 * 0.09 * 0.250 * 0.136 * 0.272 * 0.073 * 0.439 * 0.097 * 0.317 * 0.463 * 0.254) / 1.20732E-11

= 0.069

P(GPAX=Middle|A=No1) = (P(101501=2.5|GPAX= Middle) * P(105131=2.0|GPAX= Middle) * P(202101=3.0|GPAX= Middle) * P(202207=3.0|GPAX= Middle) * P(202211=2.5|GPAX= Middle) * P(202506=3.5|GPAX= Middle) * P(202507=2.0|GPAX= Middle) * P(901101=4.0|GPAX= Middle) * P(903301=3.5|GPAX= Middle) * P(Grade_T1=M|GPAX= Middle) * P(Grade_T2=H|GPAX= Middle) * P(Grade_T3=M|GPAX= Middle) * P(Grade_T4=VH|GPAX= Middle) * P(Grade_T5=H|GPAX= Middle) * P(GPAX= Middle)) / P(A)

= (0.056 * 0.283 * 0.301 * 0.301 * 0.15 * 0.018 * 0.188 * 0.018 * 0.471 * 0.54 * 0.76 * 0.62 * 0.04 * 0.58 * 0.307) / 1.20732E-11

= 0.927

P(GPAX=Low | A=No1) = (P(101501=2.5|GPAX= Low) * P(105131=2.0|GPAX= Low) * P(202101=3.0|GPAX= Low) * P(202207=3.0|GPAX= Low) * P(202211=2.5|GPAX= Low) * P(202506=3.5|GPAX= Low) * P(202507=2.0|GPAX= Low) * P(901101=4.0|GPAX= Low) * P(903301=3.5|GPAX= Low) * P(Grade_T1=M|GPAX= Low) * P(Grade_T2=H|GPAX= Low) * P(Grade_T3=M|GPAX= Low) * P(Grade_T4=VH|GPAX= Low) * P(Grade_T5=H|GPAX= Low) * P(GPAX= Low)) / P(A)

= (0.233 * 0.133 * 0.066 * 0.100 * 0.200 * 0.033 * 0.266
* 0.033 * 0.266 * 0.296 * 0.185 * 0.333 * 0.037 * 0.111
* 0.171) / 1.20732E-11
= 0.003

ตารางที่ 3-22 ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.1

No	Prediction (GPAX)	Confidence (High)	Confidence (Good)	Confidence (Middle)	Confidence (Low)
1	Middle	1.03988E-05	0.069	0.927	0.003

จากตารางที่ 3-22 ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.1 พบว่า นักศึกษา No.1 มีผลการเรียน GPAX=Middle เนื่องจากค่าความน่าจะเป็นของ P(GPAX=Middle |A=No1) มีค่าสูงสุด และมีค่าความเชื่อมั่นระดับสูง (Confidence=0.927) หมายถึง ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตของนักศึกษาเมื่อจบการศึกษาจะมีผลการเรียนเฉลี่ยอยู่ในระดับกลาง (Middle) ซึ่งได้แก่ ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตอยู่ระหว่าง 2.81-3.19

ตารางที่ 3-23 โมเดลพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.2

attribute	Parameter	High	Good	Middle	Low
101501	value=3.5	0.13	0.295	0.49	0.133
105131	value=2.5	0.152	0.34	0.094	0.066
202101	value=3.0	0.065	0.204	0.301	0.066
202207	value=3.5	0.195	0.272	0.094	0.033
202211	value=2.0	0.021	0.022	0.037	0.266
202506	value=3.5	0.239	0.09	0.018	0.033
202507	value=3.0	0.173	0.181	0.018	0.033
901101	value=4.0	0.326	0.136	0.018	0.033
903301	value=3.5	0.195	0.272	0.471	0.266
grade_T1	value=H	0.116	0.56	0.26	0.111
grade_T2	value=H	0.023	0.439	0.76	0.185
grade_T3	value=M	0.023	0.097	0.62	0.333
grade_T4	value=VH	0.813	0.317	0.04	0.037
grade_T5	value=H	0.139	0.463	0.58	0.111

ข) การพยากรณ์ผลการเรียนเฉลี่ยของนักศึกษา No2 จากการแทนค่าในสมการของ Bayes

P(GPAX=High | A=No2) = (P(101501=3.5|GPAX=High) * P(105131=2.5|GPAX=High) * P(202101=3.0|GPAX=High) * P(202207=3.5|GPAX=High) * P(202211=2.0|GPAX=High) * P(202506=3.5|GPAX=High) * P(202507=3.0|GPAX=High) * P(901101=4.0|GPAX=High) * P(903301=3.5|GPAX=High) * P(Grade_T1=H|GPAX=High) * P(Grade_T2=H|GPAX=High) * P(Grade_T3=M|GPAX=High) * P(Grade_T4=VH|GPAX=High) * P(Grade_T5=H|GPAX=High) * P(GPAX=High)) / P(A)

= (0.13 * 0.152 * 0.065 * 0.195 * 0.021 * 0.239 * 0.173 * 0.326 * 0.195 * 0.116 * 0.023 * 0.023 * 0.813 * 0.139 * 0.266) / 6.57332E-11

= 0.000

P(GPAX=Good | A=No2)= (P(101501=3.5|GPAX= Good) * P(105131=2.5|GPAX= Good) * P(202101=3.0|GPAX= Good) * P(202207=3.5|GPAX= Good) * P(202211=2.0|GPAX= Good) * P(202506=3.5|GPAX= Good) * P(202507=3.0|GPAX= Good) * P(901101=4.0|GPAX= Good) * P(903301=3.5|GPAX= Good) * P(Grade_T1=H|GPAX= Good) * P(Grade_T2=H|GPAX= Good) * P(Grade_T3=M|GPAX= Good) * P(Grade_T4=VH|GPAX= Good) * P(Grade_T5=H|GPAX= Good) * P(GPAX= Good)) / P(A)

= (0.295 * 0.34 * 0.204 * 0.272 * 0.022 * 0.09 * 0.181 * 0.136 * 0.272 * 0.56 * 0.439 * 0.097 * 0.317 * 0.463 * 0.254) / 6.57332E-11

= 0.997

P(GPAX=Middle|A=No2)= (P(101501=3.5|GPAX= Middle) * P(105131=2.5|GPAX= Middle) * P(202101=3.0|GPAX= Middle) * P(202207=3.5|GPAX= Middle) * P(202211=2.0|GPAX= Middle) * P(202506=3.5|GPAX= Middle) * P(202507=3.0|GPAX= Middle) * P(901101=4.0|GPAX=

Middle) * P(903301=3.5|GPAX= Middle) *
P(Grade_T1=H|GPAX= Middle) * P(Grade_T2=H|GPAX=
Middle) * P(Grade_T3=M|GPAX= Middle) *
P(Grade_T4=VH|GPAX= Middle) *
P(Grade_T5=H|GPAX= Middle) * P(GPAX= Middle)) /
P(A)

= (0.49 * 0.094 * 0.301 * 0.094 * 0.037 * 0.018 * 0.018 *
0.018 * 0.471 * 0.26 * 0.76 * 0.62 * 0.04 * 0.58 *
0.307) / 6.57332E-11

= 0.001

P(GPAX=Low | A=No2) = (P(101501=3.5|GPAX= Low) * P(105131=2.5|GPAX Low
Low) * P(202101=3.0|GPAX= Low) *
P(202207=3.5|GPAX= Low) * P(202211=2.0|GPAX=
Low) * P(202506=3.5|GPAX= Low) *
P(202507=3.0|GPAX= Low) * P(901101=4.0|GPAX=
Low) * P(903301=3.5|GPAX= Low) *
P(Grade_T1=H|GPAX= Low) * P(Grade_T2=H|GPAX=
Low) * P(Grade_T3=M|GPAX= Low) *
P(Grade_T4=VH|GPAX= Low) * P(Grade_T5=H|GPAX=
Low) * P(GPAX= Low)) / P(A)

= (0.133 * 0.066 * 0.066 * 0.033 * 0.266 * 0.033 * 0.033
* 0.033 * 0.266 * 0.111 * 0.185 * 0.333 * 0.037 *
0.111) / 6.57332E-11

= 3.55168E-06

ตารางที่ 3-24 ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.2

No	Prediction (GPAX)	Confidence (High)	Confidence (Good)	Confidence (Middle)	Confidence (Low)
2	Good	0.000	0.997	0.001	3.55168E-06

จากตารางที่ 3-24 ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.2 พบว่า นักศึกษา No.2 มีผลการเรียนGPAX=Good เนื่องจากค่าความน่าจะเป็นของ P(GPAX= Good |A=No2) มีค่าสูงสุดและมีค่าความเชื่อมั่นระดับสูง (Confidence=0.997) หมายถึง ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตของนักศึกษาเมื่อจบการศึกษาจะมีผลการเรียนเฉลี่ยอยู่ในระดับดี (Good) ซึ่งได้แก่ ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตอยู่ระหว่าง 3.20-3.53

ตารางที่ 3-25 โมเดลพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.3

attribute	Parameter	High	Good	Middle	Low
101501	value=4.0	0.717	0.477	0.075	0.066
105131	value=3.0	0.304	0.181	0.075	0.033
202101	value=3.5	0.173	0.318	0.301	0.133
202207	value=4.0	0.391	0.045	0.018	0.033
202211	value=4.0	0.521	0.34	0.094	0.033
202506	value=4.0	0.26	0.045	0.018	0.033
202507	value=4.0	0.369	0.113	0.018	0.033
901101	value=3.5	0.369	0.454	0.113	0.033
903301	value=3.5	0.195	0.272	0.471	0.266
grade_T1	value=VH	0.813	0.292	0.08	0.037
grade_T2	value=VH	0.906	0.487	0.02	0.037
grade_T3	value=VH	0.72	0.121	0.02	0.037
grade_T4	value=VH	0.813	0.317	0.04	0.037
grade_T5	value=VH	0.79	0.463	0.08	0.037

ค) การพยากรณ์ผลการเรียนเฉลี่ยของนักศึกษา No3 จากการแทนค่าในสมการของ Bayes

P(GPAX=High | A=No3) = (P(101501=4.0|GPAX=High) *
P(105131=3.0|GPAX=High) * P(202101=3.5|GPAX=High)
* P(202207=4.0|GPAX=High) *
P(202211=4.0|GPAX=High) * P(202506=4.0|GPAX=High)
* P(202507=4.0|GPAX=High) *
P(901101=3.5|GPAX=High) * P(903301=3.5|GPAX=High)
* P(Grade_T1=VH|GPAX=High) *
P(Grade_T2=VH|GPAX=High) *
P(Grade_T3=VH|GPAX=High) *
P(Grade_T4=VH|GPAX=High) *
P(Grade_T5=VH|GPAX=High) * P(GPAX=High)) / P(A)

= (0.717 * 0.304 * 0.173 * 0.391 * 0.521 * 0.26 * 0.369 *
0.369 * 0.195 * 0.813 * 0.906 * 0.72 * 0.813 * 0.79 *
0.266) / 4.80485E-06

= 0.999

P(GPAX=Good | A=No3)= (P(101501=4.0|GPAX= Good) * P(105131=3.0|GPAX= Good) * P(202101=3.5|GPAX= Good) * P(202207=4.0|GPAX= Good) * P(202211=4.0|GPAX= Good) * P(202506=4.0|GPAX= Good) * P(202507=4.0|GPAX= Good) * P(901101=3.5|GPAX= Good) * P(903301=3.5|GPAX= Good) * P(Grade_T1=VH|GPAX= Good) * P(Grade_T2=VH|GPAX=Good) * P(Grade_T3=VH|GPAX=Good) * P(Grade_T4=VH|GPAX=Good) * P(Grade_T5=VH|GPAX=Good) * P(GPAX= Good)) / P(A)

= (0.477 * 0.181 * 0.318 * 0.045 * 0.34 * 0.045 * 0.113 * 0.454 * 0.272 * 0.292 * 0.487 * 0.121 * 0.317 * 0.463 * 0.254) / 4.80485E-06

= 3.52146E-05

P(GPAX=Middle|A=No3)= (P(101501=4.0|GPAX= Middle) * P(105131=3.0|GPAX= Middle) * P(202101=3.5|GPAX= Middle) * P(202207=4.0|GPAX=Middle) * P(202211=4.0|GPAX= Middle) * P(202506=4.0|GPAX=Middle) * P(202507=4.0|GPAX= Middle) * P(901101=3.5|GPAX=Middle) * P(903301=3.5|GPAX= Middle) * P(Grade_T1=VH|GPAX=Middle) * P(Grade_T2=VH|GPAX= Middle) * P(Grade_T3=VH|GPAX= Middle) * P(Grade_T4=VH|GPAX= Middle) * P(Grade_T5=VH|GPAX= Middle) * P(GPAX= Middle)) / P(A)

= (0.075 * 0.075 * 0.301 * 0.018 * 0.094 * 0.018 * 0.018 * 0.113 * 0.471 * 0.08 * 0.02 * 0.02 * 0.04 * 0.08 * 0.307) / 4.80485E-06

= 3.23216E-13

P(GPAX=Low | A=No3) = (P(101501=4.0|GPAX= Low) * P(105131=3.0|GPAX= Low) * P(202101=3.5|GPAX= Low) * P(202207=4.0|GPAX= Low) * P(202211=4.0|GPAX= Low) * P(202506=4.0|GPAX= Low) * P(202507=4.0|GPAX= Low) * P(901101=3.5|GPAX=

Low) * P(903301=3.5|GPAX= Low) *
P(Grade_T1=VH|GPAX= Low) * P(Grade_T2=VH|GPAX=
Low) * P(Grade_T3=VH|GPAX= Low) *
P(Grade_T4=VH|GPAX= Low) * P(Grade_T5=VH|GPAX=
Low) * P(GPAX= Low)) / P(A)

= (0.066 * 0.033 * 0.133 * 0.033 * 0.033 * 0.033 * 0.033
* 0.033 * 0.266 * 0.037 * 0.037 * 0.037 * 0.037 * 0.037
* 0.171) / 4.80485E-06

= 7.44193E-15

ตารางที่ 3-26 ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.3

No	Prediction (GPAX)	Confidence (High)	Confidence (Good)	Confidence (Middle)	Confidence (Low)
3	High	0.999	3.52146E-05	3.23216E-13	7.44193E-15

ตารางที่ 3-26 ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา No.3 พบว่า นักศึกษา No.3 มีผลการเรียน GPAX=High เนื่องจากค่าความน่าจะเป็นของ P(GPAX=Good |A=No2) มีค่าสูงสุดและมีค่าความเชื่อมั่นระดับสูง (Confidence=0.999) หมายถึง ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตของนักศึกษาเมื่อจบการศึกษาจะมีผลการเรียนเฉลี่ยอยู่ในระดับสูง (High) ซึ่งได้แก่ ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตอยู่ระหว่าง 3.54-4.00

จากการคำนวณหาค่าความน่าจะเป็นด้วยสมการของ Bayes เพื่อการพยากรณ์ผลการเรียนเฉลี่ยของนักศึกษาที่ไม่ทราบคลาสคำตอบ สามารถสรุปได้ดังนี้

ตารางที่ 3-27 ผลการพยากรณ์ผลการเรียน (GPAX) ในอนาคตเมื่อจบการศึกษา

No	Prediction (GPAX)	Confidence (High)	Confidence (Good)	Confidence (Middle)	Confidence (Low)
1	Middle	1.03988E-05	0.069	0.927	0.003
2	Good	0.000	0.997	0.001	3.55168E-06
3	High	0.999	3.52146E-05	3.23216E-13	7.44193E-15

ตารางที่ 3-7 ผลการพยากรณ์ผลการเรียน (GPAX) ในอนาคตเมื่อจบการศึกษา พบว่า นักศึกษา No1 การพยากรณ์ผลการเรียนเฉลี่ยในอนาคตของนักศึกษาเมื่อจบการศึกษาจะมีผลการเรียนเฉลี่ยอยู่ในระดับกลาง (Middle) ได้แก่ ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตอยู่ระหว่าง 2.81-3.19 มีค่าความเชื่อมั่นระดับสูง (Confidence=0.927) นักศึกษาคนที่ 2 การพยากรณ์ผลการเรียนเฉลี่ยในอนาคตของนักศึกษาเมื่อจบการศึกษาจะมีผลการเรียนเฉลี่ยอยู่ในระดับดี (Good) ได้แก่ ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตอยู่ระหว่าง 3.20-3.53 มีค่าความเชื่อมั่นระดับสูง (Confidence=0.997)

และนักศึกษา No3 การพยากรณ์ผลการเรียนเฉลี่ยในอนาคตของนักศึกษาเมื่อจบการศึกษาจะมีผลการเรียนเฉลี่ยอยู่ในระดับสูง (High) ได้แก่ ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตอยู่ระหว่าง 3.54-4.00 มีค่าความเชื่อมั่นระดับสูง (Confidence=0.999) โดยแสดงดังตารางที่ 3-28

ตารางที่ 3-28 ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา

No	1	2	3
101501	2.5	3.5	4.0
105131	2.0	2.5	3.0
202101	3.0	3.0	3.5
202207	3.0	3.5	4.0
202211	2.5	2.0	4.0
202506	3.5	3.5	4.0
202507	2.0	3.0	4.0
901101	4.0	4.0	3.5
903301	3.5	3.5	3.5
Grade_T1	M	H	VH
Grade_T2	H	H	VH
Grade_T3	M	M	VH
Grade_T4	VH	VH	VH
Grade_T5	H	H	VH
GPAX	Middle Confident(0.927)	Good Confident(0.997)	High Confident(0.999)

ตารางที่ 3-28 ผลการพยากรณ์ผลการเรียน (GPAX) ของนักศึกษา พบว่า ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตของนักศึกษาเมื่อจบการศึกษาจะมีผลการเรียนเฉลี่ยของนักศึกษา No1 มีผลการเรียนอยู่ในระดับกลาง (Middle) ซึ่งได้แก่ ผลการอยู่ระหว่าง 2.81-3.19 มีค่าความเชื่อมั่นระดับสูง (Confidence=0.927) นักศึกษาคนที่ 2 มีผลการเรียนอยู่ในระดับดี (Good) ได้แก่ ผลการเรียนอยู่ระหว่าง 3.20-3.53 มีค่าความเชื่อมั่นระดับสูง (Confidence=0.997) และนักศึกษา No3 มีผลการเรียนอยู่ในระดับสูง (High) ได้แก่ ผลการเรียนอยู่ระหว่าง 3.54-4.00 มีค่าความเชื่อมั่นระดับสูง (Confidence=0.999)

3.3 ระยะที่ 3 ออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.3.1 วัตถุประสงค์การวิจัยระยะที่ 3

3.3.1.1 เพื่อออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.3.2 วิธีการดำเนินการวิจัยระยะที่ 3

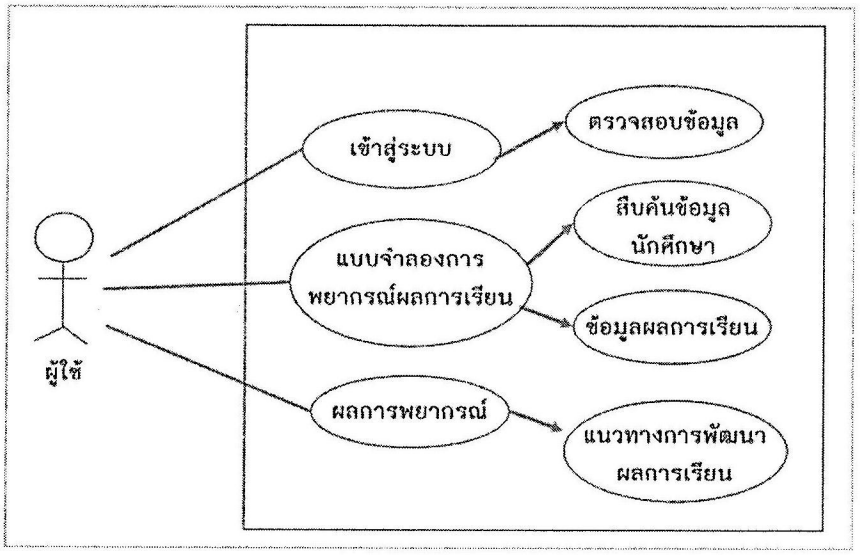
3.3.2.1 ศึกษาเครื่องมือการพัฒนาระบบที่เหมาะสมตามโครงการการออกแบบและความต้องการดังนี้

3.3.2.1.1 การพัฒนาระบบผ่านแอพพลิเคชั่นของ Appsheet ซึ่งเป็นเครื่องมือที่ช่วยในการสร้าง Mobile Application สำหรับใช้ในองค์กรโดยแทบไม่ต้องเขียนโค้ด ซึ่งมีฟังก์ชันการทำงานได้เทียบเท่าซอฟแวร์สำนักงานพวก Enterprise Resource Planning (ERP) ตั้งแต่การทำเอกสาร การทำบัญชี การทำระบบจองห้องประชุม การเก็บข้อมูลความพึงพอใจพนักงาน การตรวจวัดสต็อกสินค้า การติดตามยอดขายของฝ่ายแผนกการตลาด โดยสามารถเปิดใช้งานผ่าน Web Browser ได้ทั้งในสมาร์ทโฟน และคอมพิวเตอร์ตั้งโต๊ะ โดย Appsheet สามารถเชื่อมต่อฐานข้อมูลโดยใช้ Spread Sheet ได้ซึ่งแตกต่างจากการพัฒนาแอปพลิเคชั่นแบบเดิมที่ต้องสร้างฐานข้อมูลเฉพาะ

3.3.2.1.2 การสร้างแดชบอร์ดด้วย Google Data Studio โดย Dashboard เป็นหน้ากระดานที่ใช้ในการสรุปข้อมูลแบบ Executive ในมุมมองต่าง ๆ เพื่อให้สามารถดูได้ง่าย ใช้เวลาในการตีความสั้น และสามารถตอบโจทย์ได้ ใช้ในการติดตามเรื่องที่สนใจ เพื่อเห็นการเปลี่ยนแปลงของข้อมูลตลอดเวลา Google Data Studio เป็นเครื่องมือหนึ่งที่ฟรีใช้งานง่าย และสามารถต่อเข้ากับข้อมูลได้หลากหลายรูปแบบด้วยขั้นตอนที่ง่ายโดยเฉพาะคนที่ใช้เครื่องมือของ Google อยู่แล้ว ไม่ว่าจะเป็น Analytics, Google Ads, หรือ Google Search Console ยิ่งง่ายและสะดวกมาก ซึ่งการใช้ Data Studio นั้นช่วยลดเวลาการเตรียมข้อมูล และการทำพรีเซนต์ไปได้มาก

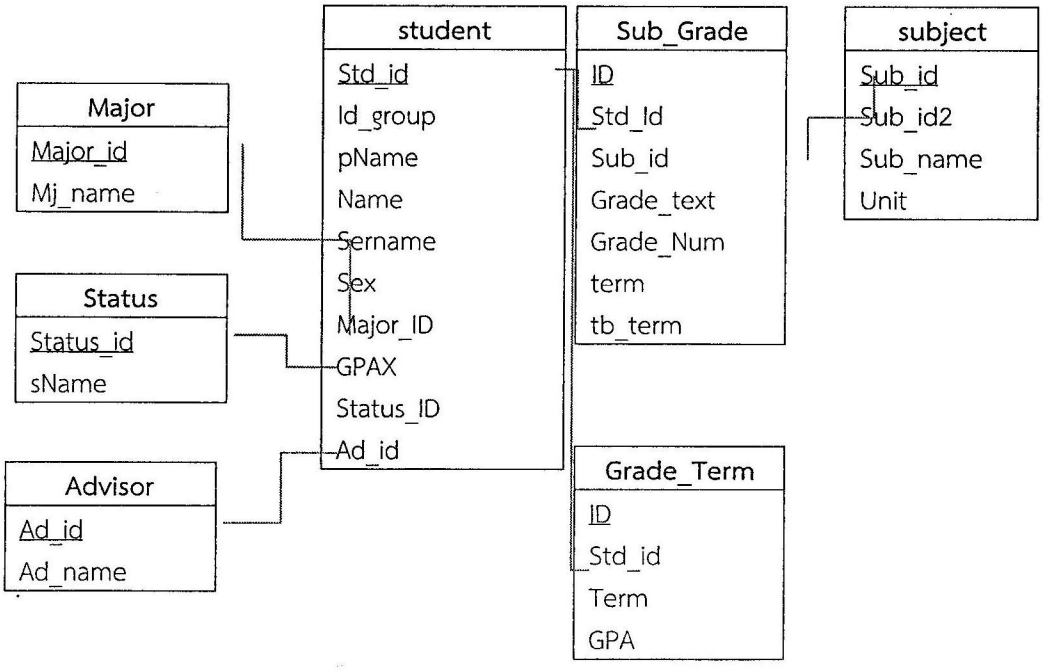
3.3.2.1.3 การจัดการฐานข้อมูลด้วย Google Sheet ซึ่งเป็น Apps ในกลุ่มของ Google Drive มีลักษณะการทำงานคล้าย ๆ กับ Excel มีการสร้าง Column Row สามารถใส่ข้อมูลต่าง ๆ ลงไปใน Cell ได้ คำนวณสูตรต่าง ๆ ได้ แต่วิธีการใช้สูตรคำนวณจะแตกต่างจาก Excel ไม่ต้องติดตั้งที่เครื่อง สามารถใช้งานบน Web ได้ โดยไฟล์จะถูกบันทึกไว้ที่ Server ของ Google ทำให้สามารถเปิดใช้งานได้ ไม่ว่าจะอยู่ที่ใด เพียงมี Web browser และอินเตอร์เน็ต สามารถแชร์ไฟล์ให้ผู้อื่นร่วมใช้งานได้ และมีระบบ Real Time Save อัตโนมัติ นอกจากนี้ยังสามารถ Save หรือ Export ออกมาใช้งานกับ Excel ที่เครื่องของเราได้อีกด้วย ทำให้การทำงานสะดวกสบายมากยิ่งขึ้น โดยการล็อกอินเข้าใช้งานในเว็บไซต์ Google ด้วย Google Account หรือ gmail ก็สามารถเข้าไปทำงาน

3.3.2.2 ออกแบบระบบการพยากณ์ผลการเรียน โดยแสดงแผนภาพการทำงานของผู้ใช้ระบบ (Use Case Diagram) และแผนความสัมพันธ์ข้อมูลเชิงสัมพันธ์ (Relation Database Model)



ภาพที่ 3-24 แผนภาพ Use Case Diagram ของระบบ

แผนภาพความสัมพันธ์ของข้อมูลในขั้นตอนการออกแบบระบบกายภาพ นำเสนอความสัมพันธ์ระหว่างข้อมูลทั้งระบบ ด้วยรูปภาพและสัญลักษณ์ในแนวทางเชิงโครงสร้าง เรียนกว่าแบบจำลองข้อมูลเชิงสัมพันธ์ (Entity Relation Diagram : E-R Diagram)



ภาพที่ 3-25 แผนผังแสดงความสัมพันธ์ของข้อมูล

พจนานุกรมข้อมูล (Data Dictionary) เป็นเครื่องมือสำหรับอธิบายการออกแบบโครงสร้างฐานข้อมูลในระดับตรรกะ (Logical Database Design) ที่ใช้ในการวิจัยพจนานุกรมข้อมูลประกอบด้วยตารางดังนี้

ตารางที่ 3-29 ข้อมูลนักศึกษา (Student)

ลำดับ	ชื่อฟิลด์	คำอธิบาย
1	Std_id	รหัสนักศึกษา
2	Id_group	รหัสหมู่เรียน
3	pName	คำนำหน้าชื่อ
4	Name	ชื่อ
5	Sername	นามสกุล
6	Sex	เพศ
7	Major_ID	รหัสสาขา
8	GPAX	เกรดเฉลี่ย
9	Status_ID	รหัสสถานะ
10	Ad_id	รหัสอาจารย์ที่ปรึกษา

ตารางที่ 3-30 ข้อมูลรายวิชา (subject)

ลำดับ	ชื่อฟิลด์	คำอธิบาย
1	Sub_id	รหัสวิชา
2	Sub_id2	รหัสวิชา2
3	Sub_name	ชื่อวิชา
4	Unit	หน่วยกิต

ตารางที่ 3-31 ข้อมูลรายวิชา (Subject)

ลำดับ	ชื่อฟิลด์	คำอธิบาย
1	ID	รหัสตาราง
2	Std_Id	รหัสนักศึกษา
3	Sub_id	รหัสวิชา
4	Grade_text	เกรดรายวิชา(ตัวอักษร)
5	Grade_Num	เกรดรายวิชา(ตัวเลข)
6	term	เทอม
7	tb_term	เทอม/ปีการศึกษา

ตารางที่ 3-32 ข้อมูลสาขาวิชา (Major)

ลำดับ	ชื่อฟิลด์	คำอธิบาย
1	Major_id	รหัสสาขา
2	Mj_name	ชื่อสาขา

ตารางที่ 3-33 ข้อมูลผลการเรียนแต่ละเทอม (Grade_Term)

ลำดับ	ชื่อฟิลด์	คำอธิบาย
1	ID	รหัสตาราง
2	Std_id	รหัสนักศึกษา
3	Term	เทอม
4	GPA	เกรดเฉลี่ยรายเทอม

ตารางที่ 3-34 ข้อมูลสถานะ (Status)

ลำดับ	ชื่อฟิลด์	คำอธิบาย
1	Status_id	รหัสสถานะ
2	sName	ชื่อสถานะ(0=กำลังศึกษา ,1=จบการศึกษา)

ตารางที่ 3-35 ข้อมูลอาจารย์ที่ปรึกษา (Advisor)

ลำดับ	ชื่อฟิลด์	คำอธิบาย
1	Ad_id	รหัสที่ปรึกษา
2	Ad_name	ชื่อที่ปรึกษา

3.4 ระยะที่ 4 พัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.4.1 วัตถุประสงค์การวิจัยระยะที่ 4

3.4.1.1 เพื่อพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.4.2 ขอบเขตการวิจัยระยะที่ 4

3.4.2.1 ประชากร ได้แก่ ผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศ ด้านเทคนิคการทำเหมืองข้อมูลหรือด้านการออกแบบแผนการเรียน

3.4.2.2 กลุ่มตัวอย่าง ได้แก่ ผู้เชี่ยวชาญด้านเทคโนโลยีสารสนเทศ ด้านเทคนิคการทำเหมืองข้อมูล และด้านการออกแบบแผนการเรียน จำนวน 9 ท่าน ได้จากการเลือกแบบสุ่มอย่างง่าย (Simple Random Sampling) โดยเป็นผู้ที่มีประสบการณ์ในด้านที่เกี่ยวข้องอย่างน้อย 3 ปี

3.4.3 เครื่องมือที่ใช้ในการวิจัยระยะที่ 4

3.4.3.1 ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.4.3.2 แบบประเมินประสิทธิภาพจากผู้เชี่ยวชาญที่มีต่อระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.4.4 ขั้นตอนการวิจัยระยะที่ 4

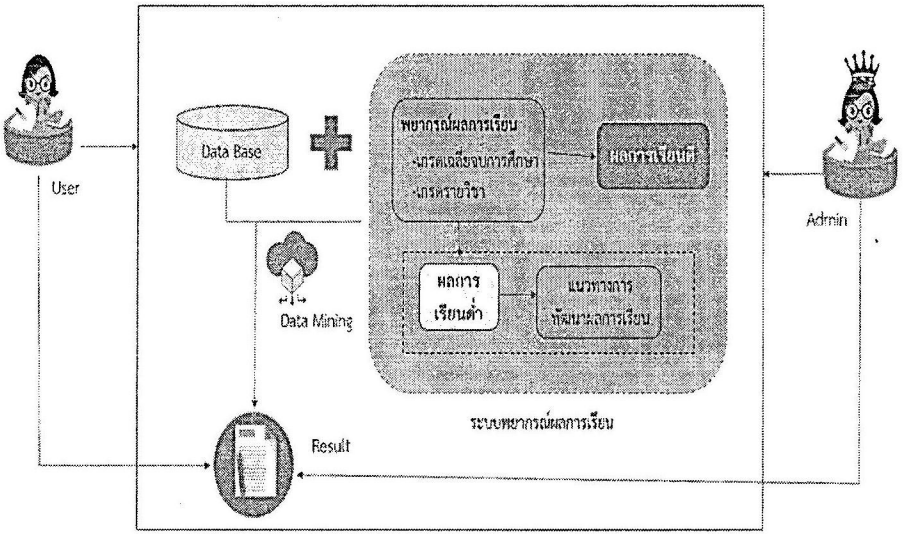
ขั้นตอนนี้ผู้วิจัยได้แบ่งขั้นตอนการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ตามกรอบแนวคิด การวิเคราะห์และสังเคราะห์ข้อมูลต่อจากขั้นตอนของการวิจัยระยะที่ 3 โดยแต่ละขั้นตอนมีรายละเอียดดังนี้

3.4.4.1 วิเคราะห์ความต้องการโดยศึกษาแนวคิด ทฤษฎี จากเอกสารและงานวิจัยที่เกี่ยวข้อง พร้อมทั้งรวบรวมรายละเอียดข้อมูลนักศึกษาจากฐานข้อมูล ความต้องการของผู้ใช้งานซึ่ง ได้แก่นักศึกษา และความเป็นไปได้ของการพัฒนาระบบ เพื่อให้ได้ข้อสรุปที่ชัดเจนและตรงจุดประสงค์ของการพัฒนาระบบ

3.4.4.2 การพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี โดยทำการเลือกเครื่องมือที่เหมาะสมตามโครงสร้างที่ออกแบบและตรงความต้องการได้ดังนี้

3.4.4.2.1 ซอฟต์แวร์สำหรับการพัฒนาระบบ มีการพัฒนาระบบผ่านแอพพลิเคชั่นของ Appsheet โดยทำการพัฒนาระบบให้ผู้ใช้งานสามารถล็อกอินเข้าสู่ระบบเพื่อเป็นการรักษาความปลอดภัยของข้อมูล และทำการพัฒนาระบบในส่วนของการพยากรณ์ผลการเรียนโดยขั้นตอนนี้ผู้ใช้งานจะต้องทำการกรอกข้อมูลส่วนตัวหลังจากนั้นระบบจะทำการพยากรณ์ผลการเรียนในอนาคตเมื่อนักศึกษาจบการศึกษา และสำหรับนักศึกษาที่มีผลการเรียนอยู่ในระดับต่ำ (Low) ระบบพยากรณ์ผลการเรียนจะทำการเสนอแผนการเรียนและผลการเรียนในปีการศึกษาถัดไปเพื่อเป็นการเสนอแนะแนวทางการพัฒนาผลการเรียนเพื่อลดการออกกลางค้นของนักศึกษาในกรณีที่นักศึกษามีผลการเรียนอยู่ในกลุ่มต่ำ (Low) ได้แก่ การพยากรณ์ผลการเรียนที่ลงทะเบียนในรายวิชาของเทอมถัดไปการพยากรณ์แผนการเรียนสำหรับนักศึกษา และแหล่งเรียบรู้เพิ่มเติมเพื่อเป็นการพัฒนาผลการเรียนของนักศึกษาให้สูงขึ้น

3.4.4.2.2 การจัดการฐานข้อมูล ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี จะทำการจัดเก็บข้อมูลของนักศึกษาสำหรับการพยากรณ์ผลการเรียนและพยากรณ์แผนการเรียนด้วยโปรแกรม Google Sheet ซึ่งข้อมูลที่ถูกจัดเก็บไว้จะทำการเชื่อมต่อไปยัง Google Data Studio เพื่อแสดงรายงานผลการพยากรณ์ผลการเรียนและผลการพยากรณ์แผนการเรียนของนักศึกษาต่อไป



ภาพที่ 3-26 แผนภาพการทำงานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคัน ของนักศึกษาระดับปริญญาตรี

3.4.4.3 การประเมินประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะของนักศึกษาระดับปริญญาตรี โดยผู้เชี่ยวชาญจำนวน 9 ท่าน โดยประเด็นการประเมินแบ่งออกเป็น 4 ด้าน ได้แก่ ด้านความสามารถของระบบตรงความต้องการ ด้านการประเมินประสิทธิภาพการใช้งานของระบบ ด้านผลลัพธ์ที่ได้จากระบบ ด้านความปลอดภัย เพื่อวัดประสิทธิภาพของระบบพยากรณ์ที่ได้ถูกพัฒนาขึ้น โดยผู้วิจัยได้ดำเนินปรับปรุงแก้ไขระบบตามข้อเสนอแนะของผู้เชี่ยวชาญก่อนทำการทดลองใช้กับกลุ่มตัวอย่างต่อไป

3.4.4 สถิติที่ใช้ในการวิจัยระยะที่ 4

3.4.4.1 การประเมินความคิดเห็นจากผู้เชี่ยวชาญที่มีต่อประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี โดยการหาค่าเฉลี่ย ($\bar{x}$) และส่วนเบี่ยงเบนมาตรฐาน (S.D)

โดยกำหนดเกณฑ์มาตราส่วนประมาณค่า (Rating Scale) ตามมาตรฐานวัดของลิเคอร์ท (Likert) 5 ระดับ ดังนี้

5 หมายถึง มีประสิทธิภาพระดับมากที่สุด
4 หมายถึง มีประสิทธิภาพระดับมาก
3 หมายถึง มีประสิทธิภาพระดับปานกลาง
2 หมายถึง มีประสิทธิภาพระดับน้อย
1 หมายถึง มีประสิทธิภาพระดับน้อยที่สุด

และกำหนดเกณฑ์การแปลผลความหมายของค่าเฉลี่ยตามแนวคิดดังนี้

4.50-5.00 หมายถึง มีประสิทธิภาพระดับมากที่สุด

3.50-4.49 หมายถึง มีประสิทธิภาพระดับมาก

2.50-3.49 หมายถึง มีประสิทธิภาพระดับปานกลาง

1.50-2.49 หมายถึง มีประสิทธิภาพระดับน้อย

1.00-1.19 หมายถึง มีประสิทธิภาพระดับน้อยที่สุด

3.5 ระยะที่ 5 ศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.5.1 วัตถุประสงค์การวิจัยระยะที่ 5

3.5.1.1 เพื่อศึกษาประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.5.1.2 เพื่อศึกษาผลการพยากรณ์ผลการเรียน และการพยากรณ์แผนการเรียนของนักศึกษา

3.5.2 ขอบเขตการวิจัยระยะที่ 5

3.5.2.1 ประชากร ได้แก่นักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่ลงทะเบียนเรียนในภาคการศึกษา 2/2563 จำนวน 3,579 คน

3.5.2.2 กลุ่มตัวอย่าง ได้แก่ นักศึกษาภาคปกติสาขาวิชาภาษาอังกฤษ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่ลงทะเบียนเรียนภาคการศึกษา 2/2563 จำนวน 40 คน ได้จากวิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling)

3.5.3 เครื่องมือที่ใช้ในการวิจัยระยะที่ 5

3.5.3.1 แบบประเมินความคิดเห็นจากผู้ใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

3.5.3.2 การประเมินค่าความเชื่อมั่นของการพยากรณ์ผลการเรียนและแผนการเรียนของนักศึกษา

3.5.4 วิธีการดำเนินการวิจัยระยะที่ 5

หลังจากดำเนินการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเรียบร้อยแล้วผู้วิจัยดำเนินการศึกษาประสิทธิภาพของระบบจากผู้เชี่ยวชาญ จำนวน 9 ท่าน และปรับปรุงตามข้อเสนอแนะเรียบร้อยแล้วนั้น จะดำเนินการศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี กับกลุ่มตัวอย่างจำนวน 40 คน รายละเอียดการทดลองมีดังนี้

3.5.4.1 การทดสอบระบบของผู้ใช้งาน จากกลุ่มตัวอย่าง จำนวน 40 คน มีดังนี้

3.5.4.1.1 ผู้ใช้งานกรอกข้อมูลประวัติ และผลการเรียนของนักศึกษาแต่ละรายวิชาสำหรับการนำข้อมูลดังกล่าวไปพยากรณ์ผลการเรียนเฉลี่ย

3.5.4.1.2 ตรวจสอบการพยากรณ์ผลการเรียน โดยระบบจะทำการบันทึกฐานข้อมูล (Database) ของนักศึกษาที่ Google Sheet เพื่อดำเนินการเก็บรวบรวมข้อมูลนักศึกษาสำหรับนำไปพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อนักศึกษาจบการศึกษาและระบบพยากรณ์ผล

การเรียนจะดำเนินการพยากรณ์ผลการเรียนอัจฉริยะโดยดำเนินการพยากรณ์ผลการเรียนซึ่งแบ่งออกเป็น 4 กลุ่ม ได้แก่ กลุ่มผลการเรียนสูง กลุ่มผลการเรียนดี กลุ่มผลการเรียนปานกลาง และระดับผลการเรียนต่ำ ตามลำดับ

3.5.4.1.3 ตรวจสอบการพยากรณ์แผนการเรียนและผลการเรียนในรายวิชาของปีการศึกษาถัดไป โดยระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันนักศึกษาระดับปริญญาตรีจะเสนอแผนการเรียนและผลการเรียนรายวิชาของเทอมให้กับนักศึกษาที่เหมาะสมกับนักศึกษาแต่ละบุคคล เพื่อเป็นแนวทางในการพัฒนาผลการเรียนที่สูงขึ้นและสามาถลดอัตราการออกกลางคันของนักศึกษาได้

3.5.4.1.4 รายงานผล (Result) ผลการพยากรณ์ผลการเรียน โดยระบบทำการรายงายและสถิติข้อมูลนักศึกษา ผลการพยากรณ์ผลเรียนเฉลี่ยในอนาคตเมื่อนักศึกษาจบการศึกษา รายงานแผนการเรียน รายงานผลการพยากรณ์ผลการเรียนในรายวิชาของเทอมถัดไป

3.5.4.2 ประเมินความคิดเห็นจากผู้ใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากกลุ่มตัวอย่างนักศึกษาสาขาวิชาภาษาอังกฤษที่ลงทะเบียนในประจำปีการศึกษา 2/2563 จำนวน 40 คน

3.5.5 สถิติที่ใช้ในการวิจัยระยะที่ 5

3.5.5.1 แบบประเมินความคิดเห็นจากผู้ใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี โดยการหาค่าเฉลี่ย ($\bar{X}$) และส่วนเบี่ยงเบนมาตรฐาน (S.D) โดยกำหนดเกณฑ์มาตราส่วนประมาณค่า (Rating Scale) ตามมาตรฐานวัดของลิเคอร์ท (Likert) 5 ระดับ ดังนี้

5 หมายถึง เห็นด้วยระดับมากที่สุด
4 หมายถึง เห็นด้วยระดับมาก
3 หมายถึง เห็นด้วยระดับปานกลาง
2 หมายถึง เห็นด้วยระดับน้อย
1 หมายถึง เห็นด้วยระดับน้อยที่สุด

และกำหนดเกณฑ์การแปลผลความหมายของค่าเฉลี่ยตามแนวคิดดังนี้

4.50-5.00 หมายถึง เห็นด้วยระดับมากที่สุด
3.50-4.49 หมายถึง เห็นด้วยระดับมาก
2.50-3.49 หมายถึง เห็นด้วยระดับปานกลาง
1.50-2.49 หมายถึง เห็นด้วยระดับน้อย
1.00-1.19 หมายถึง เห็นด้วยระดับน้อยที่สุด

3.5.5.2 การประเมินค่าความเชื่อมั่น (Confidence) ของการพยากรณ์ผลการเรียนและแผนการเรียนของนักศึกษาระดับปริญญาตรี โดยใช้ทฤษฎีความน่าจะเป็น (Probability) ดังสมการดังนี้

$$P(C|A) = \frac{P(a_1|C) * P(a_2|C) * \ldots\ldots * P(a_M|C) * \mathrm{P(C)}}{P(A)} \qquad (3\text{-}10)$$

กำหนดให้

P(C|A) หมายถึง ความน่าจะเป็นที่จะมีแอททริบิวต์เป็น A จะมีคลาส C

P(A|C) หมายถึง ความน่าจะเป็นของเทรนนิ่งดาต้าที่มีคลาส C และมีแอททริบิวต์ A โดย $A = a_1 \cap a_2 \cap \ldots \cap a_M$ และ M คือ จำนวนแอททริบิวต์ในเทรนนิ่งดาต้า

P(C) หมายถึง ความน่าจะเป็นของคลาส C

P(A) หมายถึง ค่าความน่าจะเป็นของ $A = a_1 \cap a_2 \cap \ldots \cap a_M$

และเกณฑ์การแปลผลความหมายของค่าความเชื่อมั่น ดังนี้

0.71-1.00 หมายถึง ความเชื่อมั่นระดับสูง

0.41-0.70 หมายถึง ความเชื่อมั่นระดับปานกลาง

0.21-0.40 หมายถึง ความเชื่อมั่นระดับต่ำ

0.00-0.20 หมายถึง ความเชื่อมั่นระดับต่ำมากหรือไม่มีเลย

บทที่ 4
ผลการวิจัย

จากการวิจัยเพื่อพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี เป็นการวิจัยและพัฒนา (Research and Development) โดยดำเนินการวิจัยตามจุดประสงค์ของงานวิจัย ได้แก่ 1) เพื่อวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี 2) เพื่อพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี 3) เพื่อออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี 4) เพื่อพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีและ 5) เพื่อศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ซึ่งได้สามารถผลการวิจัยแบ่งออกเป็น 5 ระยะ ดังนี้

4.1 ระยะที่ 1 ผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.2 ระยะที่ 2 ผลการพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.3 ระยะที่ 3 ผลการออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.4 ระยะที่ 4 ผลการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.5 ระยะที่ 5 ผลการศึกษาประสิทธิภาพการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.1 ระยะที่ 1 ผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.1.1 ผลการวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษา

ผลการวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษาด้วยเทคนิค Information Gain เป็นการคัดเลือกคุณลักษณะของปัจจัยสำคัญที่ส่งผลต่อการออกกลางคันของนักศึกษา จากการวิเคราะห์ปัจจัยทั้งหมด 14 ปัจจัย ได้แก่ ข้อมูลภูมิหลังของนักศึกษา ข้อมูลผู้ปกครอง ข้อมูลเกรดเฉลี่ย และสถานะของนักศึกษาที่เข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 3,913 คน ประกอบด้วยข้อมูลนักศึกษาที่ออกกลางคัน จำนวน 1,291 คน และไม่ออกกลางคัน จำนวน 2,622 คน พบว่า ปัจจัยที่ส่งผลต่อการออกลางคันของนักศึกษา 5 ปัจจัยแรก ได้แก่ เกรดเฉลี่ย (GPAX) สาขาวิชา(Id_program) คณะที่สังกัด (Id_faculty) และเพศ (Sex) ซึ่งข้อมูลภูมิหลังของนักศึกษา ข้อมูลผู้ปกครองไม่ส่งผลต่อ

การออกกลางคันของนักศึกษาแต่อย่างใด โดยแสดงรายละเอียดปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษาในตารางที่ 4-1 จากการสังเกตพบว่าปัจจัยสำคัญที่สุด ได้แก่ เกรดเฉลี่ย (GPAX) ส่งผลต่อการออกกลางคันถึงร้อยละ 78.8 ซึ่งทำให้การวิเคราะห์ปัจจัยที่ส่งผลต่อผลการเรียนของนักศึกษาจึงเป็นปัจจัยที่ต้องทำการแก้ไขอย่างเร่งด่วน ดังนั้นในขั้นตอนถัดไปผู้วิจัยจึงทำการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษา เพื่อเป็นการหาแนวทางการป้องกัน แก้ปัญหา และส่งเสริมผลการเรียนของนักศึกษาให้นักศึกษามีผลการเรียนสูงขึ้นและจะสามารถลดอัตราการออกกลางคันของนักศึกษาต่อไปได้

ตารางที่ 4-1 ปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา

รายการปัจจัย	ร้อยละของค่าน้ำหนัก
ผลการเรียนเฉลี่ย (GPAX)	78.8
สาขาวิชา (Id_program)	4.1
คณะที่สังกัด (Id_faculty)	2.4
ข้อมูลการจบการศึกษาเดิม (Old_ed)	1.6
เพศ (Sex)	1.2

4.1.2 ผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

ผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนด้วยเทคนิคการคัดเลือกคุณลักษณะสำคัญ (Feature Select)กับการพยากรณ์เหมืองข้อมูลด้วยเทคนิค Information Gain กับเทคนิค Decision Tree ผลปรากฏว่าปัจจัยการพยากรณ์ผลการเรียนของนักศึกษา ประกอบด้วย 6 ปัจจัย ที่ให้ค่าความถูกต้องร้อยละ 73.86 ประกอบด้วย 6 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 และเกรดเฉลี่ยเทอม 1 เป็นต้น

ผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนด้วยเทคนิคการคัดเลือกคุณลักษณะสำคัญ (Feature Select) กับการพยากรณ์เหมืองข้อมูลด้วยเทคนิค Information Gain กับเทคนิค เทคนิค Naïve Bayes ประกอบด้วย 14 ปัจจัย ที่ให้ค่าความถูกต้องร้อยละ 84.33 ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา 202101 ผลการเรียนวิชา 901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301 เป็นต้น แสดงรายละเอียดดังตารางที่ 4-2

ตารางที่ 4-2 รายการปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

รายการปัจจัย	เทคนิค Decision Tree	เทคนิค Naïve Bayes
เกรดเฉลี่ยเทอม 2 (a_1)	✓	✓
เกรดเฉลี่ยเทอม 3 (a_2)	✓	✓
เกรดเฉลี่ยเทอม 4 (a_3)	✓	✓
เกรดเฉลี่ยเทอม 1 (a_4)	✓	✓
ผลการเรียนวิชา 202507 (a_5)	✓	✓
เกรดเฉลี่ยเทอม 1 (a_6)	✓	✓
ผลการเรียนวิชา 202506 (a_7)		✓
ผลการเรียนวิชา 105131 (a_8)		✓
ผลการเรียนวิชา 202101 (a_9)		✓
ผลการเรียนวิชา 901101 (a_{10})		✓
ผลการเรียนวิชา 202211 (a_{11})		✓
ผลการเรียนวิชา 202207 (a_{12})		✓
ผลการเรียนวิชา 101501 (a_{13})		✓
ผลการเรียนวิชา 903301 (a_{14})		✓
ค่าความถูกต้องการพยากรณ์	73.86	84.33

จากตารางที่ 4-2 รายการปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี พบว่า เทคนิคการพยากรณ์ผลการเรียนอัจฉริยะด้วยเทคนิค Naïve Bayes ที่ใช้ปัจจัยการพยากรณ์ 14 ปัจจัยให้ค่าความถูกต้องสูงกว่าเทคนิค Decision Tree ที่ใช้ปัจจัยการพยากรณ์ 6 ปัจจัย ดังนั้นผู้วิจัยจึงเลือกใช้เทคนิค Naïve Bayes ที่ 14 ปัจจัยในพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีในขั้นตอนถัดไป

4.2 ระยะที่ 2 ผลการพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

การพัฒนาตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree) และเทคนิค Naïve Bayes เครื่องมือที่ใช้ทำเหมืองข้อมูลด้วยโปรแกรม RapidMiner พบว่า เทคนิค Decision Tree จากปัจจัย 6 ปัจจัยซึ่งให้ค่าความถูกต้องของการพยากรณ์ร้อยละ 73.86 และเทคนิค Naïve Bayes จากปัจจัย 14 ปัจจัยซึ่งให้ค่าความถูกต้องของการพยากรณ์ร้อยละ 84.33 ดังนั้นผู้วิจัยนำโมเดล Naïve Bayes ที่ให้ค่าความถูกต้องสูงไปพัฒนาระบบเพื่อพยากรณ์ผลการเรียนนักศึกษาต่อไป

ผู้วิจัยจึงนำแบบพยากรณ์ผลการเรียนที่สร้างขึ้นมาประยุกต์ในส่วนของการพยากรณ์ผลการเรียนของนักศึกษาสาขาวิชาภาษาอังกฤษชั้นปีที่ 3 โดยการนำเทคนิคนาอีฟเบย์ (Naïve Bayes) มาในการพัฒนาระบบพยากรณ์ผลการเรียนของนักศึกษา ซึ่งเหมาะสมกับประเภทของข้อมูลและคุณลักษณะ

ข้อมูล โดยคุณลักษณะที่สำคัญสำหรับการพยากรณ์ผลการเรียนด้วยเทคนิคนาอีฟเบย์มีจำนวนทั้งหมด 14 แอททริบิวต์ ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา 202101 ผลการเรียนวิชา 901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301 โดยสมการพยากรณ์ผลการรเรียนด้วยเทคนิคนาอีฟเบย์ ดังนี้

$$P(C|A) = \frac{P(a_1|C) * P(a_2|C) * \ldots\ldots * P(a_{14}|C) * \mathrm{P(C)}}{P(A)} \qquad (4\text{-}1)$$

การทดสอบประสิทธิภาพของตัวแบบพยากรณ์ผลการเรียน สามารถทำการแบ่งส่วนข้อมูลออกเป็นหลายส่วนเท่า ๆ กันเพื่อทำการทดสอบ (Cross-Validation) ประกอบด้วยข้อมูลส่วน Training Set และ Testing Set ด้วยการใช้วิธี 10-fold Cross-validation ซึ่งแบ่งข้อมูลออกเป็น 10 ส่วนเท่า ๆ กัน แล้วนำข้อมูลที่แบ่งไว้ใช้ในการสร้างตัวแบบพยากรณ์และการทดสอบตัวแบบพยากรณ์ เวียนกันไปจนครบทุกส่วนของข้อมูล ซึ่งกระบวนการทดสอบจะทำการให้ข้อมูลชุดที่ 1 เป็นข้อมูลทดสอบ (Testing Set) และข้อมูลชุดที่ 2-10 ชุดเทรนนิ่ง (Training Set) วนไปเรื่อย ๆ จนกระทั่งชุดที่ 10 เป็นชุดทดสอบ (Testing Set) และชุดที่ 1-9 เป็นชุดเทรนนิ่ง (Training Set)

ตัวแบบพยากรณ์ผลการเรียนอัจฉริยะด้วยเทคนิค Naïve Bayes สามารถทำนายความน่าจะเป็นของผลการเรียนนักศึกษาในอนาคตได้โดยการนำข้อมูลที่ไม่ทราบคลาสหาความน่าจะเป็นด้วยการแทนสมการของ Bayes ได้เลย และสามารถสรุปตัวแบบพยากรณ์ผลการเรียนอัจฉริยะด้วยเทคนิค Naïve Bayes ที่ผ่านการปรับเรียบด้วยวิธี Laplace Smoothing ได้ดังตารางที่ 4-4

ตารางที่ 4-3 ความน่าจะเป็น P(C) ของคลาส GPAX

attribute	Class	P(C)
GPAX	High	0.266
	Good	0.254
	Middle	0.307
	Low	0.171

ตารางที่ 4-4 ตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes

attribute	Parameter	High	Good	Middle	Low
101501	value=0.0	0.021	0.022	0.018	0.033
101501	value=1.0	0.021	0.022	0.018	0.033
101501	value=1.5	0.021	0.022	0.018	0.033
101501	value=2.0	0.021	0.022	0.037	0.066
101501	value=2.5	0.021	0.022	0.056	0.233

ตารางที่ 4-4 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
101501	value=3.0	0.043	0.113	0.283	0.400
101501	value=3.5	0.130	0.295	0.490	0.133
101501	value=4.0	0.717	0.477	0.075	0.066
105131	value=0.0	0.021	0.022	0.018	0.033
105131	value=1.0	0.021	0.045	0.169	0.366
105131	value=1.5	0.021	0.09	0.320	0.300
105131	value=2.0	0.065	0.227	0.283	0.133
105131	value=2.5	0.152	0.340	0.094	0.066
105131	value=3.0	0.304	0.181	0.075	0.033
105131	value=3.5	0.217	0.068	0.018	0.033
105131	value=4.0	0.195	0.022	0.018	0.033
202101	value=0.0	0.021	0.022	0.018	0.033
202101	value=1.0	0.021	0.022	0.018	0.033
202101	value=1.5	0.021	0.022	0.018	0.133
202101	value=2.0	0.021	0.022	0.056	0.366
202101	value=2.5	0.021	0.045	0.226	0.200
202101	value=3.0	0.065	0.204	0.301	0.066
202101	value=3.5	0.173	0.318	0.301	0.133
202101	value=4.0	0.652	0.340	0.056	0.033
202207	value=0.0	0.021	0.022	0.018	0.033
202207	value=1.0	0.021	0.022	0.018	0.033
202207	value=1.5	0.021	0.022	0.018	0.033
202207	value=2.0	0.021	0.045	0.094	0.433
202207	value=2.5	0.043	0.181	0.433	0.300
202207	value=3.0	0.282	0.386	0.301	0.100
202207	value=3.5	0.195	0.272	0.094	0.033
202207	value=4.0	0.391	0.045	0.018	0.033
202211	value=0.0	0.021	0.022	0.018	0.033
202211	value=1.0	0.021	0.022	0.018	0.100
202211	value=1.5	0.021	0.022	0.018	0.066
202211	value=2.0	0.021	0.022	0.037	0.266
202211	value=2.5	0.021	0.022	0.150	0.200
202211	value=3.0	0.065	0.181	0.433	0.266
202211	value=3.5	0.304	0.363	0.226	0.033

ตารางที่ 4-4 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
202211	value=4.0	0.521	0.340	0.094	0.033
202506	value=0.0	0.021	0.022	0.018	0.166
202506	value=1.0	0.021	0.068	0.132	0.133
202506	value=1.5	0.021	0.09	0.283	0.300
202506	value=2.0	0.021	0.204	0.245	0.233
202506	value=2.5	0.086	0.204	0.207	0.066
202506	value=3.0	0.326	0.272	0.075	0.033
202506	value=3.5	0.239	0.090	0.018	0.033
202506	value=4.0	0.260	0.045	0.018	0.033
202507	value=0.0	0.021	0.045	0.075	0.200
202507	value=1.0	0.021	0.068	0.301	0.300
202507	value=1.5	0.021	0.090	0.207	0.066
202507	value=2.0	0.043	0.250	0.188	0.266
202507	value=2.5	0.086	0.204	0.169	0.066
202507	value=3.0	0.173	0.181	0.018	0.033
202507	value=3.5	0.260	0.045	0.018	0.033
202507	value=4.0	0.369	0.113	0.018	0.033
901101	value=0.0	0.021	0.022	0.018	0.033
901101	value=1.0	0.021	0.022	0.018	0.066
901101	value=1.5	0.021	0.022	0.018	0.200
901101	value=2.0	0.021	0.045	0.113	0.133
901101	value=2.5	0.021	0.113	0.339	0.400
901101	value=3.0	0.195	0.181	0.358	0.100
901101	value=3.5	0.369	0.454	0.113	0.033
901101	value=4.0	0.326	0.136	0.018	0.033
903301	value=0.0	0.021	0.022	0.018	0.033
903301	value=1.0	0.021	0.022	0.018	0.066
903301	value=1.5	0.021	0.022	0.018	0.033
903301	value=2.0	0.021	0.022	0.018	0.100
903301	value=2.5	0.021	0.045	0.056	0.166
903301	value=3.0	0.043	0.181	0.358	0.300
903301	value=3.5	0.195	0.272	0.471	0.266
903301	value=4.0	0.652	0.409	0.037	0.033
grade_T1	value=VL	0.023	0.024	0.020	0.185

ตารางที่ 4-4 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
grade_T1	value=L	0.023	0.048	0.100	0.370
grade_T1	value=M	0.023	0.073	0.540	0.296
grade_T1	value=H	0.116	0.560	0.260	0.111
grade_T1	value=VH	0.813	0.292	0.080	0.037
grade_T2	value=VL	0.023	0.024	0.020	0.037
grade_T2	value=L	0.023	0.024	0.040	0.111
grade_T2	value=M	0.023	0.024	0.160	0.629
grade_T2	value=H	0.023	0.439	0.760	0.185
grade_T2	value=VH	0.906	0.487	0.020	0.037
grade_T3	value=VL	0.023	0.024	0.020	0.074
grade_T3	value=L	0.023	0.024	0.080	0.481
grade_T3	value=M	0.023	0.097	0.620	0.333
grade_T3	value=H	0.209	0.731	0.260	0.074
grade_T3	value=VH	0.72	0.121	0.020	0.037
grade_T4	value=VL	0.023	0.024	0.020	0.037
grade_T4	value=L	0.023	0.024	0.020	0.296
grade_T4	value=M	0.023	0.073	0.480	0.555
grade_T4	value=H	0.116	0.560	0.440	0.074
grade_T4	value=VH	0.813	0.317	0.040	0.037
grade_T5	value=VL	0.023	0.024	0.020	0.037
grade_T5	value=L	0.023	0.024	0.020	0.185
grade_T5	value=M	0.023	0.024	0.300	0.629
grade_T5	value=H	0.139	0.463	0.580	0.111
grade_T5	value=VH	0.790	0.463	0.080	0.037

จากตารางที่ 4-4 แสดงถึงตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes สำหรับพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ซึ่งในการพยากรณ์ผลการเรียนของนักศึกษาในอนาคตเมื่อจบการศึกษาจะทำการแทนค่าของข้อมูลแต่ละพารามิเตอร์จำนวน 14 แอททริบิวต์ ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา 202101 ผลการเรียนวิชา 901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301 แสดงตัวอย่างการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาดังตารางที่ 4-5

ตารางที่ 4-5 การพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาของนักศึกษาปัจจุบัน

no	id_no	Predict (GPAX)	Confidence (High)	Confidence (Good)	Confidence (Middle)	Confidence (Low)
1	S61xxx101	Middle	0.000	0.090	0.910	0.000
2	S61xxx102	Good	0.000	1.000	0.000	0.000
3	S61xxx103	High	1.000	0.000	0.000	0.000
...	...	...	...	...	...	...
40	S61xxx225	Middle	0.000	0.004	0.996	0.000

จากตารางที่ 4-5 การพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาของนักศึกษาปัจจุบัน โดยการแทนค่าโมเดลพยากรณ์ผลการเรียนอัจฉริยะ ผลการวิจัยพบว่า นักศึกษาคนที่ 1 ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาจะมีผลการเรียนอยู่ระดับกลาง (Middle) ที่ค่าความเชื่อมั่นระดับสูง (Confidence=0.910) ซึ่งหมายถึงนักศึกษาจะมีผลการเรียนเฉลี่ยอยู่ระหว่าง 2.81-3.19 นักศึกษาคนที่ 2 ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาจะมีผลการเรียนอยู่ระดับดี (Good) ที่ค่าความเชื่อมั่นระดับสูง (Confidence=1.000) หมายถึงนักศึกษาจะมีผลการเรียนเฉลี่ยอยู่ระหว่าง 3.20-3.53 นักศึกษาคนที่ 3 ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาจะมีผลการเรียนอยู่ระดับสูง (High) ที่ค่าความเชื่อมั่นระดับสูง (Confidence=1.000) หมายถึง นักศึกษาจะมีผลการเรียนเฉลี่ยอยู่ระหว่าง 3.54-4.00 และนักศึกษาคนที่ 40 ผลการพยากรณ์ผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาจะมีผลการเรียนอยู่ระดับกลาง (Middle) ที่ค่าความเชื่อมั่นระดับสูง (Confidence=0.996) หมายถึง นักศึกษาจะมีผลการเรียนเฉลี่ยอยู่ระหว่าง 2.81-3.19

4.3 ระยะที่ 3 ผลการออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

การออกแบบแผนภาพทำงานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา ผู้วิจัยได้นำเสนอรูปแบบดังกล่าวมาพัฒนาเป็นแอพพลิเคชั่นพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี โดยหน้าจอการออกแบบระบบมีส่วนที่ติดต่อกับผู้ใช้งานแสดงรายละเอียดดังนี้

ภาพที่ 4-1 หน้าจอระบบพยากรณ์ผลการเรียน

4.3.1 เมนูหลักสำหรับการพยากรณ์ผลการเรียน ประกอบด้วยเมนูย่อยดังนี้

4.3.1.1 ข้อมูลนักศึกษาสำหรับการพยากรณ์ผลการเรียน ประกอบด้วยข้อมูลพื้นฐานส่วนตัวและข้อมูลผลการเรียนของนักศึกษาแต่ละรายวิชา

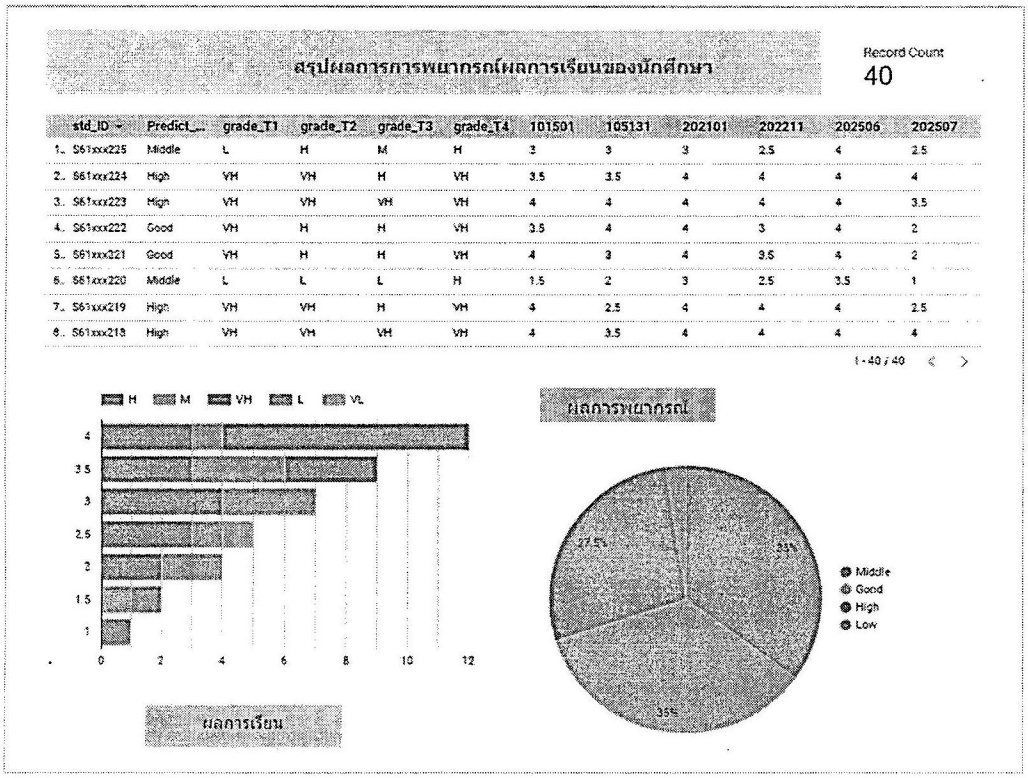
4.3.1.2 ระบบการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.3.1.3 การเสนอแนะแนวทางการพัฒนาผลการเรียนเพื่อลดการออกกลางค้นของนักศึกษาในกรณีที่นักศึกษามีผลการเรียนอยู่ในกลุ่มต่ำ (Low) ได้แก่ การพยากรณ์ผลการเรียนที่ลงทะเบียนในรายวิชาของเทอมถัดไป การพยากรณ์แผนการเรียนสำหรับนักศึกษาและแหล่งเรียบรู้เพิ่มเติมเพื่อเป็นการพัฒนาผลการเรียนของนักศึกษาให้สูงขึ้น

4.3.1.4 รายงานผลของระบบการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาโดยเสนอในรูปแบบข้อมูลของกราฟและสถิติข้อมูลผลการเรียนนักศึกษา

ภาพที่ 4-2 เมนูหลักการพยากรณ์ผลการเรียน

4.3.2 รายงานผลการพยากรณ์ผลการเรียนของนักศึกษาแต่ละคน ซึ่งได้จากวิเคราะห์ข้อมูลของนักศึกษา จำนวน 14 แอททริบิวต์



ภาพที่ 4-3 รายงานผลการพยากรณ์ผลการเรียน

4.4 ระยะที่ 4 ผลการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.4.1 ผลการประเมินประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ

การศึกษาประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ จำนนทั้งหมด 9 ท่าน ซึ่งสามารรถแบ่งรายการประเมินประสิทธิภาพไว้กว้าง ๆ เป็น 4 ด้าน ได้แก่ 1) ด้านความสามารถของระบบตรงความต้องการมากน้อยเพียงใด 2) ด้านการประเมินประสิทธิภาพการใช้งานของระบบ 3) ด้านผลลลัพธ์ที่ได้จากระบบและ 4) ด้านความปลอดภัย

ตารางที่ 4-6 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ

รายการประเมิน	ค่าเฉลี่ย	ส่วนเบี่ยงเบน	แปลผล
1. ด้านความสามารถของระบบตรงความต้องการมากน้อยเพียงใด	4.14	0.60	มาก
2. ด้านการประเมินประสิทธิภาพการใช้งานของระบบ	4.53	0.50	มากที่สุด
3. ด้านผลลลัพธ์ที่ได้จากระบบ	4.37	0.62	มาก
4. ด้านความปลอดภัย	5.00	0.00	มากที่สุด
รวมทั้งหมด	**4.48**	**0.57**	**มาก**

จากตารางที่ 4-6 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ จำนวน 9 ท่าน สามารถสรุปได้ว่าระบบพยากรณ์ผลการเรียนประสิทธิภาพอยู่ในระดับมาก ($\bar{X}$=4.48, S.D=0.57)

ตารางที่ 4-7 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ ด้านความสามารถของระบบตรงความต้องการ

รายการประเมิน	ค่าเฉลี่ย	ส่วนเบี่ยงเบน	แปลผล
ด้านความสามารถของระบบตรงความต้องการมากน้อยเพียงใด			
1. ตรงตามวัตถุประสงค์ในการพัฒนาระบบ	4.11	0.78	มาก
2. ตรงตามขอบเขตของโครงการพัฒนาระบบ	4.33	0.70	มาก
3. สามารถแก้ไขข้อผิดพลาดของระบบได้	4.00	0.00	มาก
รวมทั้งหมด	**4.14**	**0.6**	**มาก**

จากตารางที่ 4-7 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ด้านความสามารถขอระบบตรงความต้องการ จากผลการประเมินของผู้เชี่ยวชาญ จำนวน 9 ท่าน สามารถสรุปได้ว่าระบบพยากรณ์ผลการเรียนด้านความสามารถขอระบบตรงความต้องการมีประสิทธิภาพอยู่ในระดับมาก ($\bar{X}$=4.14, S.D=0.6)

ตารางที่ 4-8 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ ด้านการประเมินประสิทธิภาพการใช้งานของระบบ

รายการประเมิน	ค่าเฉลี่ย	ส่วนเบี่ยงเบน	แปลผล
ด้านการประเมินประสิทธิภาพการใช้งานของระบบ			
1. ความง่ายในการใช้งานของระบบ	4.55	0.52	มากที่สุด
2. ความสามารถของระบบในการจัดการข้อมูล	4.33	0.5	มาก
3. ความเหมาะสมของปฏิสัมพันธ์ของระบบกับผู้ใช้	4.33	0.5	มาก
4. ความรวดเร็วในการตอบสนองของระบบ	5	0	มากที่สุด
5. ประสิทธิภาพการพยากรณ์ผลการเรียน	4.55	0.52	มากที่สุด
6. ความสามารถของระบบในการรายงานผล	4.44	0.52	มาก
รวมทั้งหมด	**4.53**	**0.5**	**มากที่สุด**

จากตารางที่ 4-8 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ด้านการประเมินประสิทธิภาพการใช้งานของระบบจากผลการประเมินของผู้เชี่ยวชาญ จำนวน 9 ท่าน สามารถสรุปได้ว่าระบบพยากรณ์ผลการเรียนด้านการประเมินประสิทธิภาพการใช้งานของระบบมีประสิทธิภาพอยู่ในระดับมากที่สุด ($\bar{X}$=4.53, S.D=0.5)

ตารางที่ 4-9 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ ด้านผลลัพธ์ที่ได้จากระบบ

รายการประเมิน	ค่าเฉลี่ย	ส่วนเบี่ยงเบน	แปลผล
ด้านผลลัพธ์ที่ได้จากระบบ			
1. ความถูกต้องของผลลัพธ์ที่ได้จากระบบพยากรณ์ผลการเรียน	4.44	0.52	มาก
2. ความถูกต้องของระบบในการรายงานผล	4.55	0.52	มากที่สุด
3. การนำเสนอผลลัพธ์ของระบบแสดงข้อมูลได้ครบถ้วนและเหมาะสม	4.11	0.78	มาก
รวมทั้งหมด	**4.37**	**0.62**	**มาก**

จากตารางที่ 4-9 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ด้านผลลัพธ์ที่ได้จากระบบ จากผลการประเมินของผู้เชี่ยวชาญ จำนวน 9 ท่าน สามารถสรุปได้ว่าระบบพยากรณ์ผลการเรียนด้านผลลัพธ์ที่ได้จากระบบมีประสิทธิภาพอยู่ในระดับมาก ($\bar{X}$=4.37, S.D=0.62)

ตารางที่ 4-10 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ ด้านความปลอดภัย

รายการประเมิน	ค่าเฉลี่ย	ส่วนเบี่ยงเบน	แปลผล
ด้านความปลอดภัย			
1. การกำหนดรหัสผู้ใช้และรหัสผ่านในการตรวจสอบผู้ใช้งาน	5.00	0.00	มากที่สุด
2. การตรวจสอบสิทธิ์การใช้งานของระบบ	5.00	0.00	มากที่สุด
รวมทั้งหมด	**5.00**	**0.00**	**มากที่สุด**

จากตารางที่ 4-10 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ด้านความปลอดภัย จากผลการประเมินของผู้เชี่ยวชาญ จำนวน 9 ท่าน สามารถสรุปได้ว่าระบบพยากรณ์ผลการเรียนด้านความปลอดภัยมีประสิทธิภาพอยู่ในระดับมากที่สุด ($\bar{X}$=5.00, S.D=0.00)

4.5 ระยะที่ 5 ผลการศึกษาประสิทธิภาพการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

4.5.1 ผลการประเมินประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้ใช้งาน

การศึกษาประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้ใช้งานระบบ ได้ทำการประเมินประสิทธิภาพจากนักศึกษาสาขาวิชาภาษาอังกฤษ ที่ลงทะเบียนเรียนในประจำปีการศึกษา 2/2553 จำนวนทั้งหมด 40 คน ซึ่งสามารถแบ่งรายการประเมินประสิทธิภาพไว้กว้าง ๆ เป็น 9 รายการประเมินแสดงดังตารางที่ 4-11

ตารางที่ 4-11 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้ใช้งาน

รายการประเมิน	ค่าเฉลี่ย	ส่วนเบี่ยงเบน	แปลผล
1. ความง่ายในการใช้งานของระบบ	4.62	0.48	มากที่สุด
2. ความสามารถของระบบในการจัดการข้อมูล	4.22	0.82	มาก
3. ความเหมาะสมของปฏิสัมพันธ์ของระบบกับผู้ใช้	4.22	0.72	มาก

ตารางที่ 4-11 (ต่อ)

รายการประเมิน	ค่าเฉลี่ย	ส่วนเบี่ยงเบน	แปลผล
4. ความรวดเร็วในการตอบสนองของระบบ	4.50	0.50	มากที่สุด
5. ประสิทธิภาพการพยากรณ์ผลการเรียน	4.80	0.40	มากที่สุด
6. ความสามารถของระบบในการรายงานผล	4.87	0.33	มากที่สุด
7. ความถูกต้องของผลลัพธ์ที่ได้จากระบบพยากรณ์ผลการเรียน	4.70	0.55	มากที่สุด
8. ความถูกต้องของระบบในการรายงานผล	4.62	0.48	มากที่สุด
9. การนำเสนอผลลัพธ์ของระบบแสดงข้อมูลได้ครบถ้วนและเหมาะสม	4.85	0.35	มากที่สุด
รวมทั้งหมด	**4.60**	**0.58**	**มากที่สุด**

จากตารางที่ 4-11 ผลการประเมินประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้ใช้งาน จำนวน 40 คน สามารถสรุปได้ว่าระบบพยากรณ์ผลการเรียนประสิทธิภาพอยู่ในระดับมากที่สุด ($\bar{X}$=4.60, S.D=0.58)

4.4.2 ผลการประเมินความเชื่อมั่นของการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ผู้วิจัยได้ทำทดสอบการพยากรณ์ผลการเรียนนักศึกษาสาขาวิชาภาษาอังกฤษที่กำลังศึกษาประจำปีการศึกษา 2/2563 จำนวน 40 คน สามารถสรุปผลการพยากรณ์และค่าความเชื่อมั่นในการพยากรณ์ของนักศึกษา ได้ดังตารางที่ 4-12 – 4-13

ตารางที่ 4-12 การพยากรณ์ผลการเรียนของนักศึกษาที่มีระดับผลการเรียนอยู่ในกลุ่มต่ำ

no	id_no	Predict (GPAX)	Confidence (High)	Confidence (Good)	Confidence (Middle)	Confidence (Low)
...	.....	...	...	...	...	...
27	S61xxx207	Low	0.000	0.000	0.000	1.000
...	.....	...	...	...	...	...

จากตารางที่ 4-12 พบว่า นักศึกษารหัส S61xxx207 มีผลการพยากรณ์ผลการเรียนเฉลี่ยเมื่อจบการศึกษาอยู่ในกลุ่มต่ำ (Low) ที่ค่าความเชื่อมั่นของพยากรณ์ระดับสูง (Confidence=1.00) ดังนั้นนักศึกษาดังกล่าวจะต้องทำการพยากรณ์ผลการเรียนในรายวิชาที่ลงทะเบียนในเทอมถัดไปเพื่อเป็นการหาแนวทางในการพัฒนาผลการเรียน การวางแผนการเรียนสำหรับนักศึกษา และแหล่งเรียนรู้เพิ่มเติมเพื่อเป็นการพัฒนาผลการเรียนของนักศึกษาให้สูงขึ้น

ตารางที่ 4-13 ผลการพยากรณ์ผลการเรียนของนักศึกษากลุ่มตัวอย่าง

no	id_no	Predict (GPAX)	Confidence (High)	Confidence (Good)	Confidence (Middle)	Confidence (Low)
1	S61xxx101	Middle	0.000	0.090	0.910	0.000
2	S61xxx102	Good	0.000	1.000	0.000	0.000
3	S61xxx103	High	1.000	0.000	0.000	0.000
4	S61xxx104	Good	0.000	1.000	0.000	0.000
5	S61xxx105	Good	0.002	0.998	0.000	0.000
6	S61xxx109	Good	0.000	1.000	0.000	0.000
7	S61xxx110	Good	0.007	0.993	0.000	0.000
8	S61xxx111	High	1.000	0.000	0.000	0.000
9	S61xxx112	Middle	0.000	0.000	0.992	0.008
10	S61xxx114	Good	0.000	0.996	0.004	0.000
11	S61xxx115	Middle	0.000	0.000	1.000	0.000
12	S61xxx116	Good	0.008	0.992	0.000	0.000
13	S61xxx117	Middle	0.000	0.000	0.999	0.001
14	S61xxx119	Good	0.000	0.650	0.350	0.000
15	S61xxx120	Middle	0.000	0.352	0.648	0.000
16	S61xxx121	Middle	0.000	0.000	0.961	0.039
17	S61xxx122	Middle	0.000	0.000	1.000	0.000
18	S61xxx123	High	1.000	0.000	0.000	0.000
19	S61xxx124	High	0.578*	0.422	0.000	0.000
20	S61xxx125	High	0.740	0.260	0.000	0.000
21	S61xxx201	Middle	0.000	0.002	0.998	0.000
22	S61xxx202	Good	0.000	1.000	0.000	0.000
23	S61xxx203	Middle	0.000	0.015	0.985	0.000
24	S61xxx204	Good	0.000	0.570*	0.430	0.000
25	S61xxx205	Middle	0.000	0.294	0.706	0.000
26	S61xxx206	Good	0.007	0.993	0.000	0.000

ตารางที่ 4-13 (ต่อ)

no	id_no	Predict (GPAX)	Confidence (High)	Confidence (Good)	Confidence (Middle)	Confidence (Low)
27	S61xxx207	Low	0.000	0.000	0.000	1.000
28	S61xxx208	Middle	0.000	0.000	1.000	0.000
29	S61xxx213	Middle	0.000	0.432	0.567	0.001
30	S61xxx214	High	1.000	0.000	0.000	0.000
31	S61xxx216	High	1.000	0.000	0.000	0.000
32	S61xxx217	Good	0.001	0.986	0.012	0.000
33	S61xxx218	High	1.000	0.000	0.000	0.000
34	S61xxx219	High	0.998	0.002	0.000	0.000
35	S61xxx220	Middle	0.000	0.000	0.998	0.002
36	S61xxx221	Good	0.062	0.938	0.000	0.000
37	S61xxx222	Good	0.306	0.694	0.000	0.000
38	S61xxx223	High	1.000	0.000	0.000	0.000
39	S61xxx224	High	1.000	0.000	0.000	0.000
40	S61xxx225	Middle	0.000	0.004	0.996	0.000

จากตางที่ 4-13 ผลการพยากรณ์ผลการเรียนของนักศึกษา จำนวน 40 คน พบว่า ระบบการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีมีความเชื่อมั่นระดับสูง โดยการพยากรณ์ผลการเรียนมีนักศึกษา 2 คนที่ค่าความเชื่อมั่นในการพยากรณ์ผลการเรียนที่ความเชื่อมั่นระดับปานกลาง ได้แก่ นักศึกษารหัส S61xxx124 การพยากรณ์ผลการเรียนอยู่ในกลุ่มสูง (High) ที่ค่าความเชื่อมั่น 0.578 และ S61xxx204 การพยากรณ์ผลการเรียนอยู่ในกลุ่มดี (Good) ที่ค่าความเชื่อมั่น 0.570

บทที่ 5
ระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคัน ของนักศึกษาระดับปริญญาตรี

ระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี เป็นการวิจัยและพัฒนา (Research and Development) โดยดำเนินการวิจัยตามจุดประสงค์ของงานวิจัย ได้แก่ 1) เพื่อวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา 2) เพื่อพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา 3) เพื่อออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี 4) เพื่อพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี และ 5) เพื่อศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ซึ่งผลการวิจัยแสดงให้เห็นว่าระบบพยากรณ์ผลการเรียนที่พัฒนาขึ้นจากการศึกษาประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ จำนนทั้งหมด 9 ท่าน ซึ่งสามารรถแบ่งรายการประเมินประสิทธิภาพไว้กว้าง ๆ เป็น 4 ด้าน ได้แก่ 1) ด้านความสามารถของระบบตรงความต้องการมากน้อยเพียงใด 2) ด้านการประเมินประสิทธิภาพการใช้งานของระบบ 3) ด้านผลลัพธ์ที่ได้จากระบบและ 4) ด้านความปลอดภัย พบว่าระบบพยากรณ์ผลการเรียนประสิทธิภาพอยู่ในระดับมากและผู้ใช้ประเมินประสิทธิภาพอยู่ในระดับมากที่สุด ในบทนี้ผู้วิจัยขอนำเสนอรายละเอียดของระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีดังนี้

5.1 บทนำ ประกอบด้วย

5.1.1 ความเป็นมา เหตุผลและความจำเป็นของการพัฒนาระบบ

5.1.2 วัตถุประสงค์

5.2 ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

5.1 บทนำ

5.1.1 ความเป็นมา เหตุผลและความจำเป็นของการพัฒนาระบบ

การศึกษาไทยแบ่งการจัดการศึกษาออกเป็น 2 ระดับ คือ ระดับการศึกษาขั้นพื้นฐาน ซึ่งต้องจัดอย่าง 12 ปี และระดับการศึกษาอุดมศึกษา หรือหลังการศึกษาขั้นพื้นฐานโดยจะแบ่งออกเป็นระดับต่ำกว่าปริญญา และปริญญา ซึ่งการจัดการศึกษาระดับอุดมศึกษา มีจุดประสงค์เพื่อพัฒนาผู้เรียนให้มีลักษณะของความเป็นมนุษย์ที่สมบูรณ์ สามารถดำรงตนอยู่ในสังคมพหุวัฒนธรรมภายใต้กระแสโลกาภิวัฒน์ ที่มีการสื่อสารแบบไร้พรมแดน มีศักยภาพในการเรียนรู้ตลอดชีวิตมีความสามารถในการปฏิบัติงานได้ ตามกรอบมาตรฐานและจรรยาบรรณที่กำหนด สามารถสร้างสรรค์งานที่เกิดประโยชน์ต่อตนเองและสังคม ทั้งในระดับท้องถิ่นและสากล

การจัดการศึกษาในระดับอุดมศึกษา ตามพระราชบัญญัติการอุดมศึกษา พ.ศ. 2562 กำหนดวัตถุประสงค์ไว้ดังนี้ 1) พัฒนาทรัพยากรบุคคลให้มีความเชี่ยวชาญตามสาขาวิชาหรือวิชาชีพที่ตนถนัด สามารถตอบสนองต่อความต้องการของประเทศและสร้าขีดความสามารถในการแข่งขันในระดับโลกได้ 2) พัฒนาบุคคลให้เป็นมนุษย์ที่สมบูรณ์ทั้งร่างกาย จิตใจ สติปัญญา ความรู้ และทักษะที่จำเป็น เป็นคนดี มีวิจัย ภูมิใจในชาติ เข้าใจสังคมและวัฒนธรรมในการดำรงชีวิต สามารถปรับเปลี่ยนตนเอเพื่อรองรับสังคโลกที่จะเปลี่ยนแปลงในอนาคต มีความรับผิดชอบต่อครอบครัว ชุมชน สังคม และประเทศชาติ ร่วมกันแก้ปัญหาสังคม และสามารถอยู่ร่วมกับผู้อื่นได้อย่างมีความสุขและ 3) ตอบสนองยุทธศาสตร์ชาติ แผนแม่บท แผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ แผนการศึกษาแห่งชาติและแผนด้านการอุดมศึกษา และต้องเชื่อมโยงกับการศึกษาในระสดับที่ต่ำกว่า เพื่อเตรียมความพร้อมให้กับทรายากรบุคคลของประเทศในการเข้ารับการศึกษาในระดับอุดมศึกษา รวมทั้งส่งเสริมให้มีการศึกษาอบรมเพื่อเสริมสร้างทักษะในการประกอบอาชีพของบุคคลและการศึกษาตลอดชีวิต

สถาบันอุดมศึกษาเป็นสถาบันทางสังคมซึ่งมีบทบาทโดยตรงต่อการพัฒนาทรัพยากรมนุษย์ การจัดศึกษาในมหาวิทยาลัยมุ่งพัฒนานักศึกษาให้มีความเจริญก้าวหน้าทางสติปัญญาและความคิดอันนำไปสู่ความก้าวหน้าทางวิชาการ ที่จะสร้างสรรค์กำลังคนในระดับวิชาชีพชั้นสูงเพื่อพัฒนาสังคมและมุ่งเน้นพัฒนาคนให้มีความสมบูรณ์ด้านคุณธรรม จริยธรรม เพื่อให้สามารถดำเนินชีวิตอันมีคุณค่าต่อตนเอง สังคมและประเทศชาติ

ตามข้อกำหนดพระราชบัญญัติการศึกษาแห่งชาติกำหนดให้เด็กต้องรับการศึกษาขั้นพื้นฐานและเข้าศึกษาภาคบังคับ จำนวน 9 ปี และเมื่อเข้าศึกษาแล้วจะต้องพยายามให้เด็กศึกษาจนจบภาคการศึกษาภาคบังคับและได้ศึกษาในระดับที่สูงขึ้นต่อไป แต่ปัญหาหลักคือนักเรียนศึกษาเล่าเรียนไม่จบการศึกษาตามเกณฑ์ที่กำหนดและได้ทำการลาออกกลางคัน แต่สาเหตุของการลาออกกลางคันมาจากเหตุและปัจจัยหลายอย่างขึ้นอยู่แต่ละสถานที่แตกต่างกันไป เพื่อเป็นการแก้ปัญหาทำให้เกิดผลเสียต่อประเทศหลายประกาศทั้งความสูญเสียงบประมาณทางการศึกษาที่รัฐจัดให้มีการศึกษาทำให้ทรัพยากรบุคคลที่ควรจะมีคุณค่ากลายเป็นทรัพยากรที่มีปัญหา และต้องเสียงยประมาณในการแก้ปัญหาที่ถือเป็นการสูญเปล่าทางการศึกษาเมื่อเทียบกับการลงทุนรวมถึงด้านต่าง ๆ อีกมากมาย ดังนั้นการวิเคราะห์สาเหตุและปัจจัยที่ส่งผลให้นักศึกษาลาออกกลางคันจะสามารถทำให้โรงเรียนนหรือสถาบันระดับอุดมศึกษานำมาเป็นแนวทางในการแก้ปัญหาและลดจำนวนการออกกลางคันของนักศึกษาได้เป็นอย่างดี

มหาวิทยาลัยราชภัฏกาญจนบุรี จัดการเรียนการสอนในระดับอุดมศึกษา เปิดรับนักศึกษารุ่นแรกในปีการศึกษา 2519 ในการบริหารจัดการสิ่งที่เป็นความเสี่ยงของมหาวิทยาลัยด้านหนึ่งคือจำนวนนักศึกษา โดยมหาวิทยาลัยได้จัดทำข้อบังคับมหาวิทยาลัยราชภัฏกาญจนบุรีว่าด้วยการศึกษาระดับอนุปริญญา ปริญญาตรี และปริญญาตรี (ต่อเนื่อง) พ.ศ. 2551 แบ่งออกเป็น 8 หมวด โดยหมวดที่ 6 ได้กล่าวถึงสถานภาพนักศึกษาในข้อที่ 30 สถานภาพนักศึกษา ซึ่งมีข้อย่อยดังนี้ ข้อ 30.1 สถานภาพนักศึกษาสิ้นสุดลงด้วย 3.1.1 ตาย 30.1.2 ลาออก 30.1.3 ขาดคุณสมบัติการเข้าเป็นนักศึกษา 30.1.4 ไม่ลงทะเบียนเรียน ในภาคการศึกษาใดภาคการศึกษาหนึ่ง และไม่ลาพักการเรียน 30.1.5 นักศึกษาภาคปกติได้ค่าระดับคะแนนเฉลี่ยสะสมต่ำกว่า 1.60 เมื่อลงทะเบียนเรียนครบหลักสูตรหรือมีผลการเรียนแล้ว 2 ภาคการศึกษาปกติ หรือได้ค่าระดับคะแนนเฉลี่ยสะสมต่ำกว่า 1.80

เมื่อลงทะเบียนเรียนครบหลักสูตรหรือมีผลการเรียนแล้ว 4 ภาคการศึกษาปกตินับแต่วันเข้าเรียนและทุก ๆ สองภาคการศึกษาปกติถัดไป สำหรับนักศึกษาภาคพิเศษ จะพ้นสภาพการเป็นนักศึกษาเมื่อผลการประเมินได้คะแนนเฉลี่ยต่ำกว่า 1.80 เมื่อสิ้นภาคการศึกษาที่ 4 นับตั้งแต่เริ่มเข้าเรียนกรณีหลักสูตรระดับอนุปริญญาและระดับปริญญาตรี (ต่อเนื่อง) และเมื่อสิ้นภาคการศึกษาที่ 7 นับตั้งแต่เริ่มเข้าเรียนกรณีหลักสูตรปริญญาตรี 4 ปี และสิ้นภาคการศึกษาที่ 8 กรณีหลักสูตรปริญญาตรี 5 ปี หรือนักศึกษาลงทะเบียนครบตามที่หลักสูตรกำหนดแต่ยังได้ค่าระดับคะแนนเฉลี่ยสะสมต่ำกว่า 1.80 การนับภาคการศึกษา นับรวมทั้งภาคการศึกษาที่มีการลาพักการเรียนด้วย เมื่อนักศึกษาเรียนครบตามโครงสร้างของหลักสูตรแล้ว และได้ค่าระดับคะแนนเฉลี่ยสะสมตั้งแต่ 1.80 แต่ไม่ถึง 2.00 ให้เลือกเรียนรายวิชาเพิ่มเติมเพื่อทำค่าระดับคะแนนเฉลี่ยสะสมให้ถึง 2.00 ทั้งนี้ต้องอยู่ในระยะเวลาที่กำหนด หรือตามระยะเวลาการศึกษาที่กำหนดของการจัดการศึกษาภาคพิเศษนั้น ๆ 30.1.6 เมื่อลงทะเบียนเรียนครบกำหนดระยะเวลาการศึกษา 30.1.7 นักศึกษาไม่ผ่านรายวิชาเตรียมฝึกประสบการณ์วิชาชีพ หรือรายวิชาการฝึกประสบการณ์วิชาชีพเป็นครั้งที่ 2 30.1.8 ให้พ้นสถานภาพการเป็นนักศึกษาตามระเบียบว่าด้วยวินัยนักศึกษาและ 30.1.9 เรียนครบตามหลักสูตรและได้รับอนุมติให้สำเร็จการศึกษา

ระบบงานทะเบียนและประมวลผลของมหาวิทยาลัยราชภัฏกาญจนบุรี มีหน้าที่ติดตามและประเมินผลการเรียนของนักศึกษาที่เรียนทุกภาคการศึกษาตลอดจนครบตลอดหลักสูตร โดยนักศึกษาจะจบการศึกษาได้จะต้องมีผลการเรียนเฉลี่ย (GPAX) ไม่ต่ำกว่าเกณฑ์ที่กำหนด ซึ่งผลการเรียนเฉลี่ยเกิดจากการเรียนของแต่ละวิชาในหลักสูตร ระบบงานทะเบียนและประมวลผลจะบันทึกผลการเรียนนักศึกษาที่ฐานข้อมูลทะเบียนนักศึกษา ซึ่งระบบเหล่านี้มีการเก็บข้อมูลไว้เป็นจำนวนมากและมีการเก็บข้อมูลอย่างต่อเนื่องตั้งแต่นักศึกษาเริ่มเข้าศึกษาจนกระทั่งจบการศึกษาแต่ข้อมูลเหล่านี้ไม่ถูกนำมาใช้ประโยชน์เท่าที่ควร จากการวิเคราะห์ข้อมูลนักศึกษาที่มีการรายงานตัวเข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 4,372 คน ปรากฏว่ามีข้อมูลนักศึกษามารายงานตัวการเป็นนักศึกษาในสาขาวิชาที่สมัครแต่ไม่มีการเข้าศึกษาในสาขาดังกล่าว จำนวน 459 คน คิดเป็นร้อยละ 10.49 ซึ่งเป็นปัญหาของมหาวิทยาลัยหลาย ๆ มหาวิทยาลัยที่นักศึกษาสมัครเข้าศึกษาหลายสถาบันการศึกษาหรือหลายสาขาวิชาแต่ไม่มีการเข้าศึกษาจริงในวันเปิดเทอมโดยปัญหาดังกล่าวยากต่อการควบคุมดังนั้นการวิเคราะห์ข้อมูลนักศึกษาจะทำการวิเคราะห์ข้อมูลนักศึกษาเฉพาะที่มีการเข้ามาศึกษาในมหาวิทยาลัยเท่านั้น ได้แก่ นักศึกษา จำนวน 3,913 คน โดยผลการวิเคราะห์สถานะนักศึกษาที่มีการเข้าศึกษาในมหาวิทยาลัยฯสามารถสรุปได้ดังนี้ 1) นักศึกษาจบการศึกษา จำนวน 2,379 คน คิดเป็นร้อยละ 60.79 2) นักศึกษาตกค้าง จำนวน 243 คน คิดเป็นร้อยละ 6.21 และ 3) ออกกลางคัน จำนวน 1,291 คน คิดเป็นร้อยละ 32.99 ซึ่งจากข้อมูลดังกล่าวการออกกลางคันของนักศึกษาถือว่าเป็นสิ่งที่สำคัญในการบริหารจัดการศึกษา เพื่อหาสาเหตุและแนวทางในการป้องกันแก้ไขปัญหา ดังนั้นการสืบค้นความรู้จากระบบทะเบียนนักศึกษาจะเป็นประโยชน์ โดยการนำข้อมูลในอดีตมาวิเคราะห์เพื่อสร้างต้นแบบพยากรณ์หรือทำนายแนวโน้มสิ่งที่จะเกิดขึ้นในอนาคต เช่น พยากรณ์สถานภาพนักศึกษา พยากรณ์โอกาสการพ้นสภาพของนักศึกษาพยากรณ์โอกาสสำเร็จการศึกษาหรือการพยากรณ์ผลการเรียนของนักศึกษา โดยการนำเทคนิคเหมืองข้อมูล (Data Mining) มาสืบค้นความรู้ในอดีตเพื่อมาวิเคราะห์และสร้างต้นแบบการพยากรณ์ด้วยเทคนิคที่หลากหลาย อาทิเช่น การใช้เทคนิคกฎความสัมพันธ์

(Association Rule) หรือการจำแนกประเภท (Classification) ซึ่งเป็นเทคนิคหนึ่งของการทำเหมืองข้อมูล (Data Mining) เพื่อหาความสัมพันธ์ของผลการเรียนแต่ละวิชาและใช้เทคนิคฟัซซีกฎความสัมพันธ์ (Fuzzy Association Rule Mining) ในการคาดการณ์ผลการเรียนเฉลี่ยว่าอยู่ระดับใด เพื่อนำผลไปวิเคราะห์ผลการเรียนก่อนลงทะเบียน ถ้าผลการคาดการณ์ได้ผลการเรียนอยู่ในระดับต่ำ ผู้เรียนจะได้ปรับแผนการเรียนก่อนการลงทะเบียนและให้ผลการเรียนเป็นไปตามเกณฑ์วัดผล เทคนิคการจำแนกประเภทข้อมูล (Classification) และเทคนิคการสร้างโมเดลเพื่อใช้ในการทำนายหรือพยากรณ์สิ่งที่จะเกิดขึ้น

การออกกลางคันของนักศึกษาเป็นปัญหาสำคัญที่สถาบันการศึกษาจะต้องทำการวิเคราะห์หาปัจจัย สาเหตุ หรือแนวทางการป้องกันปัญหาทำให้นักศึกษาไม่สามารถสำเร็จการศึกษาได้ตามแผนการบริหารการศึกษาของหลักสูตร ซึ่งการออกกลางคันของนักศึกษาสะท้อนและมีผลกระทบต่อประสิทธิภาพในการจัดการศึกษาและการจัดการงบประมาณขององค์กรด้วย สำหรับการออกกลางคันของนักศึกษามีทั้งนักศึกษาที่มีผลการเรียนไม่เป็นไปตามเกณฑ์ที่กำหนด หรือออกกลางคันด้วยเหตุผลอื่น ๆ ซึ่งไม่ว่าจะเป็นเพราะปัจจัยใดนักศึกษาดังกล่าวถือว่าไม่สำเร็จการศึกษาตามที่หลักสูตรกำหนด ถือว่าเป็นการสูญเสียโอกาสในการผลิตบัณฑิต เสียทรัพยากรในการลงทุน ซึ่งการออกกลางคันของนักศึกษาถือว่าเป็นสิ่งที่สำคัญอย่างยิ่งในการบริหารจัดการ เพราะเป็นการสูญเปล่าและเป็นความล้มเหลวในการจัดการศึกษา ดังนั้นการนำฐานข้อมูลขนาดใหญ่ของระบบงานทะเบียนและประมวลผลของนักศึกษาตั้งแต่ปีการศึกษา พ.ศ. 2555-2557 วิเคราะห์ข้อมูลด้วยเทคนิคเหมืองข้อมูล (Data Mining) มาใช้วิเคราะห์ข้อมูลนักศึกษาในอดีตที่จัดเก็บไว้เพื่อใช้ในการพยากรณ์ผลการเรียนของแต่ละรายวิชาที่มีความสัมพันธ์นั้นจะเป็นแนวทางแก้ปัญหาหรือป้องกันการลดจำนวนการออกกลางคันน้อยลง ทั้งนี้เพื่อเป็นแนวทางให้อาจารย์ที่ปรึกษา นักศึกษา และผู้บริหารสามารถให้ความช่วยเหลือในการวางแผนการเรียน ดูแลการลงทะเบียนของนักศึกษาได้ จากปัญหาดังกล่าวผู้วิจัยจึงเล็งเห็นถึงความสำคัญในพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

5.1.2 วัตถุประสงค์

5.1.2.1 เพื่อวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

5.1.2.2 เพื่อพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

5.1.2.3 เพื่อออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

5.1.2.4 เพื่อพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

5.1.2.5 เพื่อศึกษาผลการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

5.1.3 ประชากรและกลุ่มตัวอย่าง

5.1.3.1 ประชากร

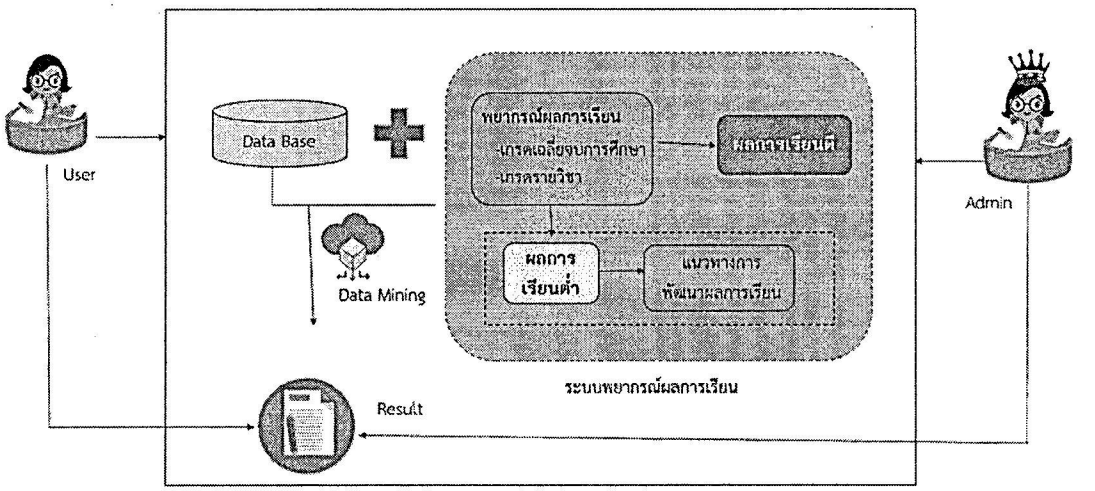
นักศึกษาภาคปกติ มหาวิทยาลัยราชภัฏกาญจนบุรี เข้าศึกษาในปีการศึกษา 2555-2557 จำนวนทั้งหมด 3,913 คน ที่สำเร็จการศึกษา จำนวน 2,379 คน และนักศึกษาที่ลงทะเบียนเรียนในภาคการศึกษา 2/2563 จำนวน 3,579 คน

5.1.3.2 กลุ่มตัวอย่าง

นักศึกษาสาขาภาษาอังกฤษ มหาวิทยาลัยราชภัฏกาญจนบุรี ที่จบการศึกษาในปีการศึกษา 2555-2557 จำนวน 141 คน และนักศึกษาปัจจุบันที่ลงทะเบียนเรียนภาคการศึกษา 2/2563 จำนวน 40 คน ได้จากวิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling)

5.2 ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

5.2.1 แผนภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ สามารถอธิบายขั้นตอนการพัฒนาระบบได้ดังนี้



ภาพที่ 5-1 แผนภาพการทำงานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา

จากแผนภาพทำงานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา แสดงถึงกระบวนการทำงานของระบบประกอบด้วย

5.2.1.1 ผู้ใช้ (User) หมายถึง นักศึกษาที่เข้าใช้ระบบพยากรณ์ผลการเรียน

5.2.1.2 ฐานข้อมูล (Database) หมายถึง ข้อมูลที่เก็บรวบรวมจากฐานข้อมูลนักศึกษาสำหรับพยากรณ์ผลการเรียน

5.2.1.3 การทำเหมืองข้อมูล หมายถึง กระบวนการที่กระทำกับข้อมูลเพื่อหาความสัมพันธ์ของข้อมูลด้วยเทคนิคต่าง ๆ

5.2.1.4 การพยากรณ์ผลการเรียน หมายถึง การพยากรณ์ผลการเรียนนักศึกษาเมื่อจบการศึกษาและการพยากรณ์ผลการเรียนรายวิชา เพื่อลดการออกกลางคันของนักศึกษาและทำการเสนอแนวทางการพัฒนาผลการเรียน

5.2.1.5 ผู้ดูแลระบบ (Admin) หมายถึง ผู้ที่จัดการและดูแลระบบด้านการใช้งานฐานข้อมูลและการรายงานประเมินผลย้อนกลับ

5.2.1.6 รายงานผล (Result) หมายถึง รายงานผลการพยากรณ์ผลการเรียน

5.2.2 การพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

จากการออกแบบแผนภาพทำงานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา ผู้วิจัยได้นำเสนอรูปแบบดังกล่าวมาพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยเพื่อลดการออกกลางคันของนักศึกษา โดยหน้าจอการออกแบบระบบมีส่วนที่ติดต่อกับผู้ใช้งานแสดงรายละเอียดดังนี้

ภาพที่ 5-2 หน้าจอระบบพยากรณ์ผลการเรียน

5.2.2.1 เมนูหลักสำหรับการพยากรณ์ผลการเรียน ประกอบด้วยเมนูย่อยดังนี้

5.2.2.1.1 ข้อมูลนักศึกษาที่ต้องการพยากรณ์ผลการเรียน

5.2.2.1.2 การพยากรณ์ผลการเรียน

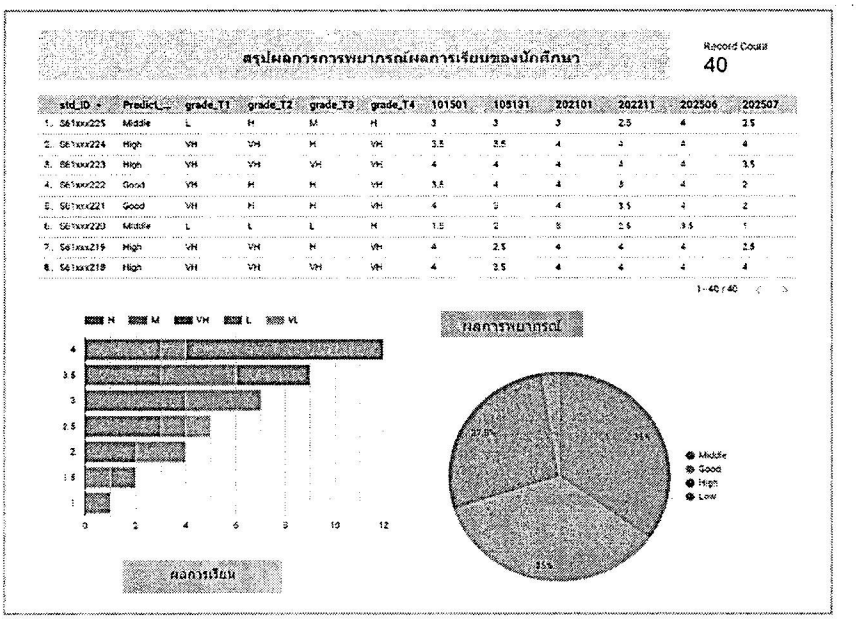
5.2.2.1.3 การเสนอแนะแนวทางการพัฒนาผลการเรียนเพื่อลดการออกกลางค้นของนักศึกษาในกรณีที่นักศึกษามีผลการเรียนอยู่ในกลุ่มต่ำ (Low) ได้แก่ การพยากรณ์ผลการเรียนที่

ลงทะเบียนในรายวิชาของเทอมถัดไปการพยากรณ์แผนการเรียนสำหรับนักศึกษาและแหล่งเรียบรู้เพิ่มเติมเพื่อเป็นการพัฒนาผลการเรียนของนักศึกษาให้สูงขึ้น

5.2.2.1.4 รายงานผลของระบบการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา โดยเสนอในรูปแบบข้อมูลของกราฟและสถิติข้อมูลผลการเรียนนักศึกษา

ภาพที่ 5-3 เมนูหลักการพยากรณ์ผลการเรียน

5.2.2.2 รายงานสรุปการพยากรณ์ผลการเรียนของนักศึกษา



ภาพที่ 5-4 สรุปรายงานผลการพยากรณ์ผลการเรียน

บทที่ 6
สรุป อภิปรายผล ข้อเสนอแนะ

การวิจัยเรื่อง ระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี มีผลการสรุปการวิจัย อภิปรายผล และข้อเสนอแนะตามลำดับดังต่อไปนี้

6.1 สรุปผลการวิจัย

6.2 อภิปรายผล

6.3 ข้อเสนอแนะ

6.1 สรุปผลการวิจัย

การสรุปผลการวิจัยเรื่อง ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีสามารถสรุปตามวัตถุประสงค์ แสดงรายละเอียดดังนี้

6.1.1 สรุปผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียน

ผลการวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษาด้วยเทคนิค Information Gain จากปัจจัยทั้งหมด 13 ปัจจัย ได้แก่ ข้อมูลภูมิหลังของนักศึกษา ข้อมูลผู้ปกครอง ข้อมูลเกรดเฉลี่ยและสถานะของนักศึกษาที่เข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 3,913 คน ประกอบด้วยข้อมูลนักศึกษาที่ออกกลางคัน จำนวน 1,291 คน และไม่ออกกลางคันจำนวน 2,622 คน พบว่า ปัจจัยที่ส่งผลต่อการออกลางคันของนักศึกษา 5 ปัจจัยแรก ได้แก่ เกรดเฉลี่ย (GPAX) สาขาวิชา (Id_program) คณะที่สังกัด (Id_faculty) และเพศ (Sex) ซึ่งข้อมูลภูมิหลังของนักศึกษาข้อมูลผู้ปกครองไม่ส่งผลต่อการออกกลางคันของนักศึกษาแต่อย่างใดจากการสังเกตพบว่าปัจจัยสำคัญที่สุดได้แก่ เกรดเฉลี่ย (GPAX) ซึ่งส่งผลต่อการออกลางคันถึงร้อยละ 78.8 ดังนั้นการวิเคราะห์ปัจจัยที่ส่งผลต่อผลการเรียนของนักศึกษาจึงเป็นปัจจัยที่ต้องทำการแก้ไขอย่างเร่งด่วน ดังนั้นผู้วิจัยจึงทำการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนของนักศึกษา เพื่อเป็นการหาแนวทางการป้องกัน แก้ปัญหาและส่งเสริมผลการเรียนของนักศึกษาให้นักศึกษามีผลการเรียนสูงขึ้นและจะสามารถลดอัตราการออกกลางคันของนักศึกษาต่อไปได้

ปัจจัยการพยากรณ์ผลการเรียนด้วยเทคนิคการคัดเลือกคุณลักษณะ (Feature Select) สำคัญกับการพยากรณ์เหมืองข้อมูลด้วยเทคนิค Information Gain กับเทคนิค Decision Tree ผลปรากฏว่าปัจจัยการพยากรณ์ผลการเรียนของนักศึกษา ประกอบด้วย 6 ปัจจัย ได้แก่ ประกอบด้วย 6 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 และเกรดเฉลี่ยเทอม 1 เป็นต้น

ปัจจัยการพยากรณ์ผลการเรียนด้วยเทคนิคการคัดเลือกคุณลักษณะ (Feature Select) สำคัญกับการพยากรณ์เหมืองข้อมูลด้วยเทคนิค Information Gain และเทคนิค Naïve Bayes ประกอบด้วย 14 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการ

เรียนวิชา 202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา 202101 ผลการเรียนวิชา 901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301 เป็นต้น

6.1.2 สรุปผลการพัฒนาตัวแบบพยากรณ์ผลการเรียนอัจฉริยะ

การพัฒนาตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree) และเทคนิคNaïve Bayes พบว่า ตัวแบบพยากรณ์ผลการเรียนเทคนิค Decision Tree จากแอททริบิวต์ 6 แอททริบิวต์ให้ค่าความถูกต้องของการพยากรณ์ร้อยละ 73.86 และตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes จากแอททริบิวต์ 14 แอททริบิวต์ให้ค่าความถูกต้องของการพยากรณ์ร้อยละ 84.33 ซึ่งผู้วิจัยนำต้นแบบการพยากรณ์ด้วยเทคนิค Naïve Bayes ที่ให้ค่าความถูกต้องสูงไปพัฒนาระบบเพื่อพยากรณ์ผลการเรียนนักศึกษาต่อไป

6.1.3 สรุปการออกแบบระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

การออกแบบแผนภาพทำงานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา ผู้วิจัยได้นำเสนอการพัฒนาเป็นแอพพลิเคชั่นพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา ประกอบด้วยเมนูหลัก 4 เมนู ได้แก่ 1) ข้อมูลนักศึกษาที่ต้องการพยากรณ์ผลการเรียน 2) การพยากรณ์ผลการเรียน 3) แนวทางการเสนอแนะการพัฒนาผลการเรียน ในกรณีที่นักศึกษามีผลการเรียนอยู่ในกลุ่มต่ำ (Low) ระบบจะเสนอแนวทางการพัฒนาผลการเรียนด้วยการเสนอแผนการเรียนและและเสนอแหล่งที่ผู้เรียนสามารถศึกษาเพิ่มเติมด้วยตนเองและ 4) รายงานผลของระบบ โดยเสนอในรูปแบบข้อมูลของกราฟและสถิติข้อมูลผลการเรียนนักศึกษา

6.1.4 สรุปผลการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

6.1.4.1 ผลการประเมินประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ

ผลการประเมินประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ จำนวนทั้งหมด 9 ท่าน ซึ่งสามารรถแบ่งรายการประเมินประสิทธิภาพไว้กว้าง ๆ เป็น 4 ด้าน ได้แก่ 1. ด้านความสามารถของระบบตรงความต้องการมากน้อยเพียงใด 2. ด้านการประเมินประสิทธิภาพการใช้งานของระบบ 3. ด้านผลลลัพธ์ที่ได้จากระบบ และ 4. ด้านความปลอดภัย พบว่า ระบบพยากรณ์ผลการเรียนประสิทธิภาพอยู่ในระดับมาก ($\bar{X}$=4.48, S.D=0.57)

6.1.5 สรุปผลการศึกษาประสิทธิภาพการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

6.1.5.1 ผลการประเมินประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้ใช้งาน

การศึกษาประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้ใช้งานระบบ ผู้วิจัยได้ทำการศึกษาและประเมินประสิทธิภาพจากนักศึกษาสาขาวิชาภาษาอังกฤษ ที่ลงทะเบียนเรียนในประจำปีการศึกษา 2/2553 จำนวน 40 คน

ซึ่งสามารถแบ่งรายการประเมินประสิทธิภาพไว้กว้าง ๆ เป็น 9 รายการประเมิน พบว่า ระบบพยากรณ์ผลการเรียนประสิทธิภาพอยู่ในระดับมากที่สุด ($\bar{X}$=4.60, S.D=0.58)

6.1.5.2 ผลการประเมินค่าความเชื่อมั่นของการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาได้ทำการศึกษาทดลองกับกลุ่มตัวอย่างของนักศึกษาสาขาวิชาภาษาอังกฤษที่ลงทะเบียนเรียนในภาคการศึกษาที่ 2/2563 จำนวน 40 คน ผลการวิจัย พบว่า ผลการพยากรณ์ผลการเรียนของนักศึกษา จำนวน 38 คน มีความเชื่อมั่นในการพยากรณ์ระดับสูงและผลการพยากรณ์ผลการเรียนของนักศึกษา จำนวน 2 คน มีค่าความเชื่อมั่นในการพยากรณ์ระดับกลาง ได้แก่ นักศึกษา รหัส S61xxx124 ผลพยากรณ์ผลการเรียนอยู่ในกลุ่มสูง (High) ที่ค่าความเชื่อมั่น 0.578 และ S61xxx204 พยากรณ์ผลการเรียนอยู่ในกลุ่มดี (Good) ที่ค่าความเชื่อมั่น 0.570

6.2 อภิปรายผล

จากการวิเคราะห์และสังเคราะห์เอกสารและงานวิจัยที่เกี่ยวข้องกับระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี สามารถอภิปรายผลตามวัตถุประสงค์ดังนี้

6.2.1 การวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียน

ก่อนการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนผู้วิจัยได้ทำการวิเคราะห์ปัจจัยการออกกลางคันของนักศึกษาด้วยเทคนิค Information Gain จากปัจจัยทั้งหมด 14 ปัจจัย ได้แก่ ข้อมูลภูมิหลังของนักศึกษา ข้อมูลผู้ปกครอง ข้อมูลเกรดเฉลี่ย และสถานะการออกกลางคันของนักศึกษาที่เข้าศึกษาในปีการศึกษา 2555-2557 จำนวน 3,913 คน ปรากฏว่าปัจจัยที่ส่งผลต่อการออกกลางคันของนักศึกษา 5 ปัจจัยแรก ได้แก่ เกรดเฉลี่ย (GPAX) สาขาวิชา (Id_program) คณะที่สังกัด (Id_faculty) และเพศ (Sex) ซึ่งข้อมูลภูมิหลังของนักศึกษาและข้อมูลผู้ปกครองไม่ส่งผลต่อการออกกลางคันของนักศึกษาแต่อย่างใด แต่ปัจจัยสำคัญที่สุดได้แก่ เกรดเฉลี่ย (GPAX) ซึ่งส่งผลต่อการออกลางคันถึงร้อยละ 78.8 ดังนั้นการวิเคราะห์ปัจจัยที่ส่งผลต่อผลการเรียนของนักศึกษาจึงเป็นปัจจัยที่ต้องทำการแก้ไขอย่างเร่งด่วน เพื่อเป็นการหาแนวทางการป้องกัน แก้ปัญหา และส่งเสริมผลการเรียนของนักศึกษาให้นักศึกษามีผลการเรียนสูงขึ้นและจะสามารถลดอัตราการออกกลางคันของนักศึกษาได้ และผลการวิเคราะห์ปัจจัยการพยากรณ์ผลการเรียนด้วยเทคนิคการคัดเลือกคุณลักษณะ (Feature Select) สำคัญกับการพยากรณ์เหมืองข้อมูลด้วยเทคนิค Information Gain กับเทคนิค Decision Tree ผลปรากฏว่าปัจจัยการพยากรณ์ผลการเรียนของนักศึกษา ประกอบด้วย 6 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 และเกรดเฉลี่ยเทอม 1 เป็นต้น และเทคนิค Naïve Bayes ประกอบด้วย 14 ปัจจัย ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา202101 ผลการเรียนวิชา 901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301 เป็นต้น และการพัฒนาตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree) และเทคนิค Naïve Bayes พบว่า ตัวแบบพยากรณ์ผลการเรียนเทคนิค Decision Tree จากคุณลักษณะ 6 คุณลักษณะซึ่งให้ค่าความถูกต้องของการพยากรณ์ร้อยละ 73.86 และตัว

รูปแบบข้อมูลของกราฟและสถิติข้อมูลผลการเรียนนักศึกษา ซึ่งสอดคล้องกับชุติมา และคณะ (2553) ได้ศึกษาวิจัยเรื่องการพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา ซึ่งเป็นการวิจัยที่มีความมุ่งหมายเพื่อสร้างตัวแบบสำหรับหาปัจจัยที่ส่งผลต่อการตัดสินใจเลือกสาขาวิชาเรียนของนักศึกษา และพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจอย่างอัตโนมัติแบบออนไลน์สำหรับแนะแนวทางการเลือกสาขาวิชาเรียนของนักศึกษาในระดับอุดมศึกษา ตัวแบบเชิงความน่าจะเป็นได้สร้างขึ้นภายใจเทคนิคการจัดทำเหมืองข้อมูล ซึ่งในการทดสอบแบบจำลองที่ได้จะทำการทดสอบผลบนพื้นฐานของวิธี K-ford cross validation โดยใช้โปรแกรม Weka ในการสร้างแบบจำลอง หลังจากนั้นได้นำผลลัพธ์ที่ดีที่สุดจากเทคนิคข่ายงานเบย์ มาสร้างตัวแบบในขั้นสุดท้ายและได้เรียบเทียบความถูกต้องของตัวแบบกับผลการวิเคราะห์ความถดถอยเชิงพหุคูณ ในการวิจัยครั้งนี้ใช้ข้อมูลด้านการเรียนของกลุ่มตัวอย่างจากนักศึกษาในระดับอุดมศึกษาจากมหาวิทยาลัยทั้งภาครัฐบาลและเอกชน จำนวน 9 มหาวิทยาลัย ผลการวิจัยที่ได้แสดงให้เห็นว่า ตัวแบบความน่าจะเป็นโดยใช้เทคนิคการจัดทำเหมืองข้อมูลตามวิธีของข่ายงานเบย์ สามารถบ่งบอกตัวแปรสำคัญที่มีผลต่อการตัดสินใจในการเลือกสาขาวิชาเรียนของนักศึกษาในระดับอุดมศึกษาได้ และให้ค่าความแม่นยำในการทำนายสูงถึงร้อยละ 91.35 จากตัวแบบดังกล่าวทำให้ทราบว่าตัวแปรสำคัญที่มีต่อการตัดสินใจในการเลือกสาขาวิชาเรียนของนักศึกษาในระดับอุดมศึกษา ได้แก่ เกรดเฉลี่ยวิชาคณิตศาสตร์ก่อนศึกษา เกรดเฉลี่ยวิชาคณิตศาสตร์ขณะกำลังศึกษา เกรดเฉลี่ยวิชาเขียนโปรแกรม ความรู้ในการพัฒนาระดับคอมพิวเตอร์และซอฟต์แวร์ ความรู้ด้านวิชาเรียน 1 ความรู้ด้านวิชากรเรียน 2 และความถนัด นอกจากนั้นตัวแปรรายวิชาที่ค้นพบที่มีผลต่อการเลือกสาขาวิชานั้น มีความสอดคล้องกับผลที่ได้จากการวิเคราะห์ความถดถอยเชิงหุคูณ จึงน่าเชื่อว่าตัวแปรที่ได้จากการพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจ ด้วยข่ายงานเบย์ดังกล่าวมีความน่าเชื่อถืออยู่ในระดับที่ยอมรับได้

6.2.4 การพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

ผลการประเมินประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้เชี่ยวชาญ จำนนทั้งหมด 9 ท่าน ซึ่งสามารรถแบ่งรายการประเมินประสิทธิภาพไว้กว้าง ๆ เป็น 4 ด้าน ได้แก่ 1) ด้านความสามารถของระบบตรงความต้องการมากน้อยเพียงใด 2) ด้านการประเมินประสิทธิภาพการใช้งานของระบบ 3) ด้านผลลัพธ์ที่ได้จากระบบ และ 4) ด้านความปลอดภัย พบว่า ระบบพยากรณ์ผลการเรียนประสิทธิภาพอยู่ในระดับมาก ($\bar{X}$=4.48, S.D=0.57) ซึ่งสอดคล้องกับภรัณยา และคณะ (2563) ได้พัฒนาระบบวิเคราะห์ข้อมูลเศรษฐกิจชุมชนเพื่อสนับสนุนการตัดสินใจด้วยเทคนิคเหมืองข้อมูล วัตถุประสงค์เพื่อ 1) ออกแบบและพัฒนาระบบวิเคราะห์ข้อมูลเศรฐกิจชุมชนด้วยเทคนิคเหมืองข้อมูลและ 2) ประเมินประสิทธิภาพการใช้งานและการยอมระบบระบบ ผลการวิจัยพบว่า 1) ระบบวิเคราะห์ข้อมูลเศรฐกิจชุมชนด้วยเทคนิคเหมืองข้อมูลแบ่งออกเป็น 4 ด้าน และด้านการวิเคราะห์ข้อมูลที่นำตัวแบบที่ได้จากอัคลอริทึม Random Tree ซึ่งประสิทธิภาพสูงสุดให้ค่าความถูกต้องร้อยละ 98.13 ค่าความแม่นยำเท่ากับร้อยละ 98.10 และ 2) ผลการประเมินประสิทธิภาพการใช้งานและการยอมรับระบบอยู่ในระดับมาก ($\bar{X}$=4.38, S.D=0.69)

6.2.5 สรุปผลการศึกษาประสิทธิภาพการใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

การศึกษาประสิทธิภาพของระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีจากผู้ใช้งานระบบ ผู้วิจัยได้ทำการศึกษาและประเมินประสิทธิภาพจากนักศึกษาสาขาวิชาภาษาอังกฤษ ที่ลงทะเบียนเรียนในประจำปีการศึกษา 2/2553 จำนวน 40 คน พบว่า ระบบพยากรณ์ผลการเรียนประสิทธิภาพอยู่ในระดับมากที่สุด ($\bar{X}$=4.60, S.D=0.58) และผลการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี พบว่า ผลการพยากรณ์ผลการเรียนของนักศึกษา จำนวน 38 คน มีความเชื่อมั่นในการพยากรณ์ระดับสูง และผลการพยากรณ์ผลการเรียนของนักศึกษาจำนวน 2 คน มีค่าความเชื่อมั่นในการพยากรณ์ระดับกลาง ได้แก่ นักศึกษารหัส S61xxx124 ผลพยากรณ์ผลการเรียนอยู่ในกลุ่มสูง (High) ที่ค่าความเชื่อมั่น 0.578 และ S61xxx204 พยากรณ์ผลการเรียนอยู่ในกลุ่มดี (Good) ที่ค่าความเชื่อมั่น 0.570 ซึ่งสอดคล้องกับ Sen (2012) ได้ศึกษาเรื่องการคาดการณ์และวิเคราะห์การทดสอบความรู้ในระดับมัธยมศึกษาด้วยวิธีการทำเหมืองข้อมูล โดยทำการศึกษาปัจจัยที่นำไปสู่ความสำเร็จ (หรือความล้มเหลว) ของนักเรียนในการทดสอบเป็นปัญหาที่น่าสนใจและท้าทาย ตั้งแต่การทดสอบส่วนกลางและความก้าวหน้าทางวิชาการจะถือว่าเป็นแนวความคิดที่เกี่ยวข้องกับการวิเคราะห์ปัจจัยพื้นฐานของความสำเร็จ การทดสอบอาจช่วยให้เข้าใจและอาจปรับปรุงผลสัมฤทธิ์ทางการเรียน ในการศึกษานี้ใช้ข้อมูลจากการศึกษาในระดับมัธยม ในตุรกีพัฒนาแบบจำลองในการทำนายผลการทดสอบในระดับมัธยมศึกษาและการใช้การวิเคราะห์การคาดการณ์ที่สำคัญที่สุด ผลการศึกษาพบว่าวิธีการตัดสินใจ C5 ต้นไม้การตัดสินใจมีความแม่นยำ 95% และเครือข่ายประสาทเทียม (กับความถูกต้องของ 89%) ใน รูปแบบโลจิสติกแบบถดถอยความถูกต้องโดยรวมของ 82% การวิเคราะห์ก่อนหน้านี้ไม่ว่าจะเป็นนักเรียนที่มีทุนการศึกษา จำนวนนักเรียนปีก่อนหน้า เกรดเฉลี่ยอยู่ในกลุ่มพยากรณ์ที่สำคัญที่สุดคือคะแนนการทดสอบ

6.3 ข้อเสนอแนะ

จากผลการพัฒนาระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีผู้วิจัยมีข้อเสนอแนะสำหรับผู้ที่ต้องพัฒนาระบบดังนี้

6.3.1 ข้อเสนอแนะในการวิจัยครั้งนี้

6.3.1.1 ตัวแบบการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีสามารถนำไปประยุกต์ใช้กับข้อมูลนักศึกษาในสาขาอื่น ๆ ได้ และยังสามารถทำนายผลการเรียนเฉลี่ยในอนาคตเมื่อจบการศึกษาได้

6.3.1.2 ตัวแบบการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีสามารถนำไปประยุกต์ใช้กับการทำนายผลการเรียนรายวิชาของนักศึกษาที่ลงทะเบียนในปีการศึกษาถัดไปได้และยังเป็นแนวทางในการจัดแผนการเรียนในหลักสูตรให้เหมาะสมกับผู้เรียนต่อไปได้

6.3.1.3 ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีที่พัฒนาขึ้นสามารถทำงานผ่านระบบเครือข่าย โดยการผู้ใช้ทำการล็อคอินเข้าสู่ระบบเพื่อเป็นการป้องการความปลอดภัยในการเข้าถึงข้อมูลของผู้ใช้งานแต่ละบุคคลได้อย่างสะดวกรวดเร็ว

6.3.2 ข้อเสนอแนะในการวิจัยครั้งต่อไป

6.3.2.1 ควรมีการพัฒนาต่อยอดงานวิจัยด้วยการเพิ่มโมดูลในการบริหารจัดการระบบสำหรับอาจารย์ที่ปรึกษา นักวิชาการศึกษา หรือสำหรับผู้บริหารสถานศึกษา เพื่อเป็นการหาสาเหตุปัจจัยหรือแนวทางการป้องกันปัญหาที่หลายหลายด้านที่ทำให้นักศึกษาไม่สามารถสำเร็จการศึกษาได้ตามแผนการบริหารการศึกษาของหลักสูตรมากยิ่งขึ้น

6.3.2.2 การพัฒนาระบบควรพัฒนาและกำหนดรูปแบบการใช้งานและขับเคลื่อนเชิงนโยบายของผู้บริหารสำหรับการบริหารจัดการนักศึกษาในสถาบันอุดมศึกษา เพื่อเป็นการป้องการปัญหาการออกกลางคันของนักศึกษาในสถาบันอุดมศึกษาให้มีอัตราลดลงด้วยการใช้ระบบพยากรณ์ผลการเรียนอัจฉริยะของนักศึกษาก่อนที่นักศึกษาจะทำการลงทะเบียนในแต่ละรายวิชาต่าง ๆ ซึ่งสามารถเป็นแนวทางในการจัดแผนการเรียนในหลักสูตรให้เหมาะสมกับผู้เรียนต่อไปได้

บรรณานุกรม

ภาษาไทย

กิตติ ภักดีวัฒนะกุล. (2550). **คัมภีร์การวิเคราะห์และออกแบบระบบ.** กรุงเทพฯ : เคทีพี คอมพ์ แอนด์ คอนซัลท์.

คำนาย อภิปรัชญาสกุล. (2557). **คู่มือซอฟต์แวร์การวางแผนทรัพยากรองค์กร (Enterprise Resource Planning Software Handbook).** พิมพ์ครั้งที่ 2. กรุงเทพฯ : โฟกัสมีเดีย แอนด์ พับลิชชิ่ง.

จตุพล นุกูลวุฒิโอภาส. (2552). **การพัฒนาตัวแบบด้วยเทคนิคการทำเหมืองข้อมูลสำหรับการวางแผนและ ควบคุมการผลิต.** กรุงเทพฯ : มหาวิทยาลัยธรรมศาสตร์.

จามรกุล เหล่าเกียรติกุล และจิรารัตน์ สิทธิวรชาติ. (2549). "ระบบอัจฉริยะเพื่อสนับสนุนการให้คำปรึกษาทางวิชาการระดับอุดมศึกษาในสาขาวิชาเทคโนโลยีสารนเทศโดยใช้เทคนิคแบบผสมผสาน." ใน **The proceedings of the first national conference of information technology (NCIT 2006).** กรุงเทพฯ : คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี : 61-68.

_____. (2551). "ตัวแบบพยากรณ์ลักษณะความเหมาะสมของนักศึกษาใหม่สาขาวิชาเทคโนโลยีสารสนเทศด้วยกฎการจำแนกประเภทเชิงสัมพันธ์." **วารสารเทคโนโลยีสารสนเทศ.** ปีที่ 4 ฉบับที่ 8 : 7-13.

จามรกุล เหล่าเกียรติกุล. (2552). **รูปแบบการให้คำปรึกษาด้านวิชาการเพื่อจัดแผนการเรียนในระดับอุดมศึกษา ด้วยเทคนิคผนมผสานภายใต้วิธีการให้เหตุผลโดยใช้กรณีเป็นฐาน.** วิทยานิพนธ์ปรัชญาดุษฎีบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

จีระนันต์ เจริญรัตน์. (2559). "การวิเคราะห์กฎความสัมพันธ์ของการสั่งอาหารด้วยเทคนิคเหมืองข้อมูล กรณีศึกษา ร้านอาหารบ้านฟ้าโปร่ง จ.สกลนคร." ใน **การประชุมวิชาการระดับชาติราชมงคลสุรินทร์วิชาการครั้งที่ 8.** สุรินทร์ : มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์ : 90-98.

_____. (2559). "การวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาที่มีผลการเรียนปกติโดยใช้ต้นไม้ตัดสินใจ." SNRU Journal of Science and Technology. ปีที่ 8 ฉบับที่ 2 : 256-267.

ชณิดาภา บุญประสม และจรัญ แสนราช. (2561). "การวิเคราะห์การทำนายการลาออกกลางคันของนักศึกษาระดับปริญญาตรี โดยใช้เทคนิควิธีการทำเหมืองข้อมูล." **วารสารวิชาการครุศาสตร์อุตสาหกรรม พระจอมเกล้าพระนครเหนือ.** ปีที่ 9 ฉบับที่ 1 : 142-151.

ชลนิศา สาระ. (2550). **การจำแนกกลุ่มสถานภาพการสำเร็จการศึกษาโดยแบบจำลองต้นไม้ตัดสินใจ.** สารนิพนธ์วิทยาศาสตมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเพตอร์ ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

ชุติมา อุตมะมุณีย์ และประสงค์ ปราณีตพลกรัง. (2553). "การพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา." **Journal of information science and technology.** ปีที่ 1 ฉบับที่ 2 : 39-48.

ซอและ เกปัน พิมลพรรณ ลีลาภัทรพันธุ์ และอัจฉราพร ยกขุน. (2561). "การวิเคราะห์ปัจจัยที่มีผลต่อการพ้นสภาพของนักศึกษาโดยใช้เทคนิคเหมืองข้อมูล กรณีศึกษา หลักสูตรวิทยาการคอมพิวเตอร์และหลักสูตรเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏยะลา." **วารสาร Veridian E-Journal สาขาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยศิลปากร.** ปีที่ 5 ฉบับที่ 4 : 96-110.

ณรงค์ศักดิ์ คงทิม และจิรัฏฐา ภูบุญอบ. (2554). "การประยุกต์ใช้เอฟพี-กโรธกับงานแนะแนวการศึกษาต่อในระดับอุดมศึกษา." ใน **National Conference on Computer Information Technologies 2011 (CIT2011) and UniNet Network Operation and Management Workshop (UniNOMS 2011).** กรุงเทพฯ : มหาวิทยาลัยมหิดล : 13-17.

ณัฏฐพร พิมพายน. (2543). **ระบบสนับสนุนการตัดสินใจทางธุรกิจ.** นนทบุรี : สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมธิราช.

ณัฐธิดา สุวรรณโณ และอันธิกา สิงห์เอี่ยม. (2554). "การหาปัจจัยที่ส่งผลต่อความเสี่ยงของนักศึกษาเรียนอ่อนด้วยเทคนิคกฎความสัมพันธ์ กรณีศึกษา : มหาวิทยาลัยสงขลานครินทร์." **วารสารวิทยาการจัดการ.** ปีที่ 28 ฉบับที่ 1 : 65-79.

ทิพย์หทัย ทองธรรมชาติ. (2560). "การคัดเลือกคุณลักษณะเพื่อสร้างโมเดลสำหรับการพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วยเทคนิคเหมืองข้อมูล." **รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ ครั้งที่ 4.** สถาบันวิจัย มหาวิทยาลัยราชภัฏกำแพงเพชร : 338-348.

ธนพัฒน์ ทองมา. (2563). "การพัฒนาโมเดลทำนายแผนการเรียนในการศึกษาต่อระดับมัธยมศึกษาตอนปลายของโรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม : การประยุกต์ใช้เทคนิคเอ็นเซ็มเบิ้ลโหวตร่วมกันระหว่างเครือข่ายใยประสาท ซัพพอร์ตเวกเตอร์แมชชีน และต้นไม้ตัดสินใจ." **วารสารครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.** ปีที่ 48 ฉบับที่ 3 : 125-143.

นงลักษณ์ วิรัชชัย. (2542). **โมเดลลิสเรล สถิติวิเคราะห์สำหรับการวิจัย.** กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

นนทวัฒน์ ทวีชาติ และคณะ. (2562). "การวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาระดับปริญญาตรีด้วยเทคนิคการทำเหมืองข้อมูล." ใน **การประชุมวิชาการระดับชาติการจัดการเทคโนโลยีและนวัตกรรม ครั้งที่ 5.** มหาวิทยาลัยราชภัฏมหาสารคาม จังหวัดมหาสารคาม : 1-8.

พงษ์เทพ รักผกาวงศ์ และอุไรวรรณ์ รักผกาวงศ์. (2558). "การใช้กฎความสัมพันธ์ร่วมกับฟัซซี่กฎความสัมพันธ์เพื่อคาดการณ์ผลการเรียนของนักศึกษา." **วารสารวิทยาศาสตร์มหาวิทยาลัยขอนแก่น**. ปีที่ 43 ฉบับที่ 3 : 542-551.

พงษ์เทพ รักผกาวงศ์. (2558). "การหาความสัมพันธ์ผลการเรียนจากโรงเรียนเดิมกับผลการเรียนในขณะศึกษาในมหาวิทยาลัยโดยใช้เทคนิคกฎความสัมพันธ์ฟัซซี่." ใน **รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ พิบูลสงครามวิจัย 2558**. สถาบันวิจัยและพัฒนา มหาวิทยาลัยพิบูลสงคราม : 47-52.

พีระศักดิ์ เสรีกุล. (2548). "ระบบสนับสนุนการตัดสินใจในการวางแผนการเรียน." **วารสารเทคโนโลยีสารสนเทศ**. ปีที่ 1 ฉบับที่ 1 : 92-97.

ภรัณยา ปาลวิสุทธิ์ และคณะ. (2563). "การพัฒนาระบบวิเคราะห์ข้อมูลเศรษกิจชุมชนเพื่อสนับสนุนการตัดสินใจด้วยเทคนิคเหมืองข้อมูล." **วารสารวิชาการจัดการเทคโนโลยีมหาวิทยาลัยมหาสารคาม**. ปีที่ 7 ฉบับที่ 2 : 72-83.

ภรัณยา อำมฤครัตน์ และพยุง มีสัจ. (2553). "การเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มข้อมูลของโรคลมร้อนด้วยวิธีการทางเครือข่ายประสาทเทียม." ใน **The 6TH National Conference on Computing and Information Technology**. กรุงเทพฯ : มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ : 116-121.

ภัทรพงศ์ พงศ์ภัทรากานต์. (2553). "การวิเคราะห์ปัจจัยที่ส่งผลต่อการพ้นสภาพของนักศึกษาปริญญาตรี โดยใช้คอมมิตตีแมชชีน." ใน **The 6th National Conference on Computer and Information Technology**. กรุงเทพฯ : มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ : 491-496.

ราชกิจจานุเบกษา. (1 พฤษภาคม 2562). **พระราชบัญญัติการอุดมศึกษา พ.ศ. 2562**. เล่มที่ 136 ตอนที่ 57 ก : 54-78.

_____. (25 พฤษภาคม 2548). ประกาศกระทรวงศึกษาธิการ. **เกณฑ์มาตรฐานหลักสูตรระดับปริญญาตรี พ.ศ. 2548**. เล่มที่ 122 ตอนพิเศษ 36 ง : 7-13.

สายชล สินสมบูรณ์ทอง. (2558). **การทำเหมืองข้อมูล**. กรุงเทพฯ : จามจุรีโปรดักส์.

สุวรรณี ธูปจีน มณเฑียร รัตนศิริวงศ์วุฒิ และมนต์ชัย เทียนทอง. (2559). "การพัฒนาขั้นตอนวิธีแบบผสมผสานสำหรับพยากรณ์ผลสัมฤทธิ์ทางการเรียน." **Science and Technology RMUTT Journal**. ปีที่ 6 ฉบับที่ 1 : 139-152.

อสมา กุลวานิชไชยนันท์. (2561). **Big Data Series II : Think Like a Data Scientist คิดแบบนักวิทยาศาสตร์ข้อมูล**. พิมพ์ครั้งที่ 1. กรุงเทพฯ : พราว เพรส.

อุกฤษ ปัจฉิม. (2546). **การประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลในการทำนายระดับน้ำสูงสุด**. วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมทรัพยากรน้ำ คณะวิศวกรรมศาสตร์ บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี.

เอกสิทธิ์ พัชรวงศ์ศักดา. (2557). **การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้า ไมน์นิง เบื้องต้น**. พิมพ์ครั้งที่ 2. กรุงเทพฯ : เอเชีย ดิจิตอลการพิมพ์.

_____. (2560). **Practical Data Mining with RapidMiner Studio8.** กรุงเทพฯ : เอเชียดิจิตอลการพิมพ์.

ภาษาอังกฤษ

Bose, I. and Mahapatra, R. K. (2001). "Business Data Mining-a Machine Learning Perspective." **Information and Management.** Vol.39 : 211-225.

Dash, Raj., Paramguru, R. L. and Dash, Ras. (2011). "Comparative Analysis of Supervised and Unsupervised Discretization Techniques." **International Journal of Advances in Science and Technology.** Vol. No.3 : 29-37.

Delavari, N., Phon-amnuaisuk, S. and BEikzadeh M.R. (2008). "Data Mining Application in Higher Learning Institutions." **Informatics in Education.** Vol.7 No.1 : 31-54.

Joreskog, K.G. and Sorbom, D. (1996). **LISREL 8 user's reference guide.** Chicago : Scientific Software International.

Kardan, A., et al. (2013). "Prediction of student course selection in online higher education institutes using neural network." **Computers & Education.** Vol.65 : 1–11.

Magdin, M. and Turčáni, M. (2015). "Personalization of student in course management system on the basis using method of Data Mining." **TOJET : The Turkish Online Journal of Educational Technology.** Vol.14 No.1 : 58-64.

Mardikyan, S. and Badur, B. (2011). "Analyzing Teaching Performance of Instructors Using Data Mining Techniques." **Informatics in Education.** Vol.10 No.2 : 245–257.

Mashael, A,. Barrak, A. and Muna, A. R. (2016). "Prediction Student Final GPA Using Decision Trees : A Case Study." **International Journal of Information and Education Technology.** Vol.6 No.7 : 528-533.

Mulika, H. (2019). [online]. **Machine Learning with Clustering Model.** [cited Oct 15, 2020]. Available form : การทำ Machine Learning ด้วย Clustering Model | TNI University (medium.com).

Oancea, B., Dragoescu, R. and Ciucu, S. (2013). "Predicting students' results in higher education using neural networks." **International Conference on Applied Information and Communication Technologies.** : 190-193.

Sen, B., Uçar, E. and Delen, D. (2012). "Predicting and analyzing secondary education placement-test scores: A data mining approach." **Expert Systems with Applications.** Vol.39 No.10 : 9468–9476.

Shrivastava, P. and Rajput, A. (2013). "Predicting Course and Branch Interest among Higher Education Students from Rural and Semi-Urban Area using Data Mining Techniques." **International Journal of Advanced Research in Computer Science.** Vol.4 No.10 : 179-188.

Tekin, A. (2014). "Early Prediction of Students' Grade Point Averages at Graduation : A Data Mining Approach." **Eurasian Journal of Educational Research.** No.54 : 207-226.

Turban, E., Aronson, J.E. and Liang, T. P. (2001). **Decision Support Systems and Intelligent Systems.** Prentice Hall.

Yangyang, S., Martha L. and Catholijn, M. J. (2015) **"Recurrent neural network language model adaptation with curriculum learning." Computer Speech and Language.** Vol.33 No.1 : 136–154.

ภาคผนวก ก

รายนามผู้เชี่ยวชาญ

รายนามผู้เชี่ยวชาญ

1. ผู้ช่วยศาสตราจารย์ ดร.อัครวิทย์ โพชะเรือง
   ตำแหน่ง อาจารย์สาขาวิชาเทคโนโลยีการศึกษาและคอมพิวเตอร์
   สังกัด คณะครุศาสตร์ มหาวิทยาลัยราชภัฏกาญจนบุรี
2. ผู้ช่วยศาสตราจารย์ธีรเดช เทวาภินันท์
   ตำแหน่ง อาจารย์สาขาวิชาวิทยาการคอมพิวเตอร์
   สังกัด คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี
3. ผู้ช่วยศาสตราจารย์จรายุทธ ประทีปวรกาญจน์
   ตำแหน่ง อาจารย์สาขาวิชาเทคโนโลยีอุตสาหกรรม
   สังกัด คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏกาญจนบุรี
4. ผู้ช่วยศาสตราจารย์ลัดดาวัลย์ จำปา
   ตำแหน่ง อาจารย์สาขาวิชาเทคโนโลยีอุตสาหกรรม
   สังกัด คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏกาญจนบุรี
5. อาจารย์ ดร.ณรงค์ พันธุ์คง
   ตำแหน่ง อาจารย์สาขาวิชาคอมพิวเตอร์ธุรกิจ
   สังกัด คณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏกาญจนบุรี
6. อาจารย์สะใบแพร อาจศรี
   ตำแหน่ง อาจารย์สาขาวิชาวิทยาการคอมพิวเตอร์
   สังกัด คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี
7. ผู้ช่วยศาสตราจารย์วิลัยพร ไชยสิทธิ์
   ตำแหน่ง อาจารย์สาขาวิชาเทคโนโลยีและวัตกรรมคอมพิวเตอร์
   สังกัด คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏหมู่บ้านจอมบึง
8. ดร.บุรินทร์ นรินทร์
   ตำแหน่ง คณบดี คณะวิทยาศาสตร์และเทคโนโลยี อาจารย์สาขาวิชาคอมพิวเตอร์ศึกษา
   สังกัด คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏหมู่บ้านจอมบึง
9. อาจารย์ ดร.นาวิน คงรักษา
   ตำแหน่ง อาจารย์สาขาวิชาเทคโนโลยีดิจิทัลมีเดีย
   สังกัด คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏหมู่บ้านจอมบึง

ภาคผนวก ข

เครื่องมือที่ใช้ในการวิจัย

- แบบประเมินความคิดเห็นจากผู้เชี่ยวชาญที่มีต่อประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี
- แบบประเมินความคิดเห็นจากผู้ใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

แบบประเมินความคิดเห็นจากผู้เชี่ยวชาญที่มีต่อประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

ชื่องานวิจัย เรื่องระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี
อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร.พัลลภ พิริยะสุรวงศ์
ผู้วิจัย นางสาวบุษบา สังข์วรรณะ

คำชี้แจง

แบบประเมินชุดนี้เป็นแบบประเมินความคิดเห็นเกี่ยวกับการใช้งานระบบพยากรณ์ผลการเรียนของนักศึกษาที่พัฒนาขึ้นประกอบด้วยข้อคำถามและเกณฑ์การประเมินแบบมาตราส่วนประมาณค่า 5 ระดับ โดยให้ผู้เชี่ยวชาญทำเครื่องหมาย ✓ ลงในช่องที่ตรงกับความคิดเห็นของท่าน โดยกำหนดค่าระดับคะแนนดังนี้

5	หมายถึง	มีประสิทธิภาพระดับมากที่สุด
4	หมายถึง	มีประสิทธิภาพระดับมาก
3	หมายถึง	มีประสิทธิภาพระดับปานกลาง
2	หมายถึง	มีประสิทธิภาพระดับน้อย
1	หมายถึง	มีประสิทธิภาพระดับน้อยที่สุด

แบบประเมินความคิดเห็นจากผู้เชี่ยวชาญที่มีต่อประสิทธิภาพระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ผู้วิจัยแบ่งการประเมินออกเป็น 3 ตอน ดังนี้

ตอนที่ 1 ข้อมูลส่วนตัวของผู้ประเมิน
ชื่อผู้ประเมิน .....
ตำแหน่ง .....
ประสบการณ์ด้านการทำงาน .....

ตอนที่ 2 ประเมินความคิดเห็นของผู้เชี่ยวชาญต่อประสิทธิภาพของระบบพยากรณ์ผลการเรียนของนักศึกษา

รายการประเมิน	ระดับความคิดเห็น					หมายเหตุ
	5	4	3	2	1	
1. ด้านความสามารถของระบบตรงความต้องการมากน้อยเพียงใด						
1.1 ตรงตามวัตถุประสงค์ในการพัฒนาระบบ						
1.2 ตรงตามขอบเขตของโครงการพัฒนาระบบ						
1.3 สามารถแก้ไขข้อผิดพลาดของระบบได้						
2. ด้านการประเมินประสิทธิภาพการใช้งานของระบบ						
2.1 ความง่ายในการใช้งานของระบบ						
2.2 ความสามารถของระบบในการจัดการข้อมูล						
2.3 ความเหมาะสมของปฏิสัมพันธ์ของระบบกับผู้ใช้						
2.4 ความรวดเร็วในการตอบสนองของระบบ						
2.5 ประสิทธิภาพการพยากรณ์ผลการเรียน						
2.6 ความสามารถของระบบในการรายงานผล						
3. ด้านผลลัพธ์ที่ได้จากระบบ						
3.1 ความถูกต้องของผลลัพธ์ที่ได้จากระบบพยากรณ์ผลการเรียน						
3.2 ความถูกต้องของระบบในการรายงานผล						
3.3 การนำเสนอผลลัพธ์ของระบบแสดงข้อมูลได้ครบถ้วนและเหมาะสม						
4. ด้านความปลอดภัย						
4.1 การกำหนดรหัสผู้ใช้และรหัสผ่านในการตรวจสอบผู้ใช้งาน						
4.2 การตรวจสอบสิทธิ์การใช้งานของระบบ						

ตอนที่ 3 ข้อเสนอแนะอื่น ๆ เพิ่มเติม

.....

.....

ลงชื่อ.....ผู้เชี่ยวชาญ

(.....)

วันที่ประเมิน ...../...../.....

แบบประเมินความคิดเห็นจากผู้ใช้งานระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

ชื่องานวิจัย	เรื่องระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร.พัลลภ พิริยะสุรวงศ์
ผู้วิจัย	นางสาวบุษบา สังข์วรรณะ

คำชี้แจง

แบบประเมินชุดนี้เป็นแบบประเมินความคิดเห็นเกี่ยวกับการใช้งานระบบพยากรณ์ผลการเรียนของนักศึกษาที่พัฒนาขึ้นประกอบด้วยข้อคำถามและเกณฑ์การประเมินแบบมาตราส่วนประมาณค่า 5 ระดับ โดยให้ผู้ใช้ทำเครื่องหมาย ✓ ลงในช่องที่ตรงกับความคิดเห็นของท่านโดยกำหนดค่าระดับคะแนนดังนี้

5	หมายถึง	เห็นด้วยระดับมากที่สุด
4	หมายถึง	เห็นด้วยระดับมาก
3	หมายถึง	เห็นด้วยระดับปานกลาง
2	หมายถึง	เห็นด้วยระดับน้อย
1	หมายถึง	เห็นด้วยระดับน้อยที่สุด

แบบประเมินความคิดเห็นจากผู้ใช้งานที่มีต่อระบบพยากรณ์ผลการเรียนอัจฉริยะ เพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี ผู้วิจัยแบ่งการประเมินออกเป็น 3 ตอน ดังนี้

ตอนที่ 1 ข้อมูลส่วนตัวของผู้ประเมิน

ชื่อผู้ประเมิน .....

สาขาวิชา .....ชั้นปีที่ .....

ตอนที่ 2 ประเมินความคิดเห็นของผู้เชี่ยวชาญต่อประสิทธิภาพของระบบพยากรณ์ผลการเรียนของนักศึกษา

รายการประเมิน	ระดับความคิดเห็น					หมายเหตุ
	5	4	3	2	1	
1. ความง่ายในการใช้งานของระบบ						
2. ความสามารถของระบบในการจัดการข้อมูล						
3. ความเหมาะสมของปฏิสัมพันธ์ของระบบกับผู้ใช้						
4. ความรวดเร็วในการตอบสนองของระบบ						
5. ประสิทธิภาพการพยากรณ์ผลการเรียน						
6. ความสามารถของระบบในการรายงานผล						
7. ความถูกต้องของผลลัพธ์ที่ได้จากระบบพยากรณ์ผลการเรียน						
8. ความถูกต้องของระบบในการรายงานผล						
9. การนำเสนอผลลัพธ์ของระบบแสดงข้อมูลได้ครบถ้วนและเหมาะสม						

ตอนที่ 3 ข้อเสนอแนะอื่น ๆ เพิ่มเติม

.....

.....

.....

.....

.....

.....

.....

.....

ลงชื่อ.....ผู้ประเมิน

(.....)

วันที่ประเมิน ...../...../.....

ภาคผนวก ค

การวิเคราะห์ข้อมูลการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคัน ของนักศึกษาระดับปริญญาตรี

1.

การวิเคราะห์ข้อมูลการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษา

1.1 ข้อมูลผลการเรียนของนักศึกษากลุ่มตัวอย่างสำหรับการสร้างตัวแบบพยากรณ์ผลการเรียนอัจฉริยะ จำนวน 141 คน แสดงดังตารางที่ ค-1

ตารางที่ ค-1 ผลการเรียนของนักศึกษาสำหรับการสร้างตัวแบบพยากรณ์

ผลการเรียนแต่ละรายวิชา	A	B+	B	C+	C	D+	D	F	รวม
การสอนวิชาเอกบังคับ	**49**	**22**	**18**	**25**	**18**	**6**	**3**		**141**
105142	49	22	18	25	18	6	3		141
ครูบังคับ	**377**	**395**	**272**	**53**	**25**	**3**		**1**	**1126**
101001	58	67	13	1	1			1	141
101101	60	38	32	5	3	3			141
101201	48	55	30	6	2				141
101203	36	55	44	6					141
101301	55	45	32	6	3				141
101401	36	45	45	6	7				139
101501	56	45	30	8	2				141
101502	28	45	46	15	7				141
ครูเลือก	**72**	**57**	**9**	**2**	**1**				**141**
108201	72	57	9	2	1				141
พื้นฐานภาษาและการสื่อสาร	**211**	**81**	**53**	**48**	**21**	**5**	**2**		**421**
901101	19	40	35	32	8	5	1		140
901201	84	23	12	10	10		1		140
901202	108	18	6	6	3				141
พื้นฐานมนุษย์	**125**	**87**	**52**	**10**	**2**		**2**	**1**	**279**
902101	61	52	23	4	1				141
902301	64	35	29	6	1		2	1	138

ตารางที่ ค-1 (ต่อ)

ผลการเรียนแต่ละรายวิชา	A	B+	B	C+	C	D+	D	F	รวม
พื้นฐานวิทย์-คณิต	**206**	**113**	**64**	**18**	**10**	**2**	**1**		**414**
904101	104	30	6	1					141
904201	77	29	22	4	7		1		140
904301	25	54	36	13	3	2			133
พื้นฐานสังคม	**61**	**74**	**98**	**33**	**14**	**1**	**1**		**282**
903101	14	24	64	26	12	1			141
903301	47	50	34	7	2		1		141
เอกบังคับ	**395**	**309**	**425**	**306**	**250**	**141**	**112**	**46**	**1984**
105131	8	11	23	25	28	27	19		141
105133	43	17	23	25	19	8	6		141
202101	45	38	26	17	11	3			140
202102	41	26	49	11	13		1		141
202103	42	20	33	23	16	6	1		141
202204	12	16	33	23	30	15	9	14	152
202207	18	23	45	38	17				141
202211	41	39	38	12	8	1	2		141
202303	36	31	30	26	12	2	3	1	141
202307	20	31	45	23	15	2	5		141
202505	19	16	17	27	24	22	15	10	150
202506	12	13	28	22	22	25	11	8	141
202507	20	12	14	20	20	14	25	9	134
202508	38	16	21	14	15	16	15	4	139
เอกเลือก	**55**	**38**	**25**	**20**	**2**	**1**			**141**
202702	55	38	25	20	2	1			141
ผลรวมทั้งหมด	**1551**	**1176**	**1016**	**515**	**343**	**159**	**121**	**48**	**4929**

1.2 ข้อมูลนักศึกษาสำหรับเทรนนิ่งระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

ผู้วิจัยดำเนินการสร้างตัวแบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรีด้วยเทคนิค Naïve Bayes โดยคัดเลือกกลุ่มตัวอย่างสำหรับการสร้างตัวแบบพยากรณ์ผลการเรียนอัจฉริยะ ดังนี้ กลุ่มตัวอย่างสำหรับการสร้างตัวแบบพยากรณ์ได้แก่ นักศึกษาสาขาภาษาอังกฤษที่จบการศึกษาแล้ว จำนวน 141 คน และกลุ่มตัวอย่างสำหรับการทดสอบตัวแบบพยากรณ์ผลการเรียนอัจฉริยะ ได้แก่นักศึกษาสาขาวิชาภาษาอังกฤษชั้นปีที่ 3 จำนวน 40 คน โดยแอททริบิวต์ที่สำคัญสำหรับการพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษามีจำนวนทั้งหมด 14 แอททริบิวต์ ได้แก่ เกรดเฉลี่ยเทอม 2 เกรดเฉลี่ยเทอม 3 เกรดเฉลี่ยเทอม 4 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202507 เกรดเฉลี่ยเทอม 1 ผลการเรียนวิชา 202506 ผลการเรียนวิชา 105131 ผลการเรียนวิชา 202101 ผลการเรียนวิชา 901101 ผลการเรียนวิชา 202211 ผลการเรียนวิชา 202207 ผลการเรียนวิชา 101501 และผลการเรียนวิชา 903301

ตารางที่ ค-2 ข้อมูลผลการเรียนเฉลี่ย (GPAX) ของนักศึกษาแต่ละคลาส

คลาส	จำนวน
High	38
Good	36
Middle	45
Low	22
ทั้งหมด	**141**

ตารางที่ ค-3 ข้อมูลผลการเรียนรายวิชาและผลการเรียนเฉลี่ยแต่ละเทอม

Attribute	Parameter	High	Good	Middle	Low
101501	0.0	0	0	0	0
101501	1.0	0	0	0	0
101501	1.5	0	0	0	0
101501	2.0	0	0	1	1
101501	2.5	0	0	2	6
101501	3.0	1	4	14	11
101501	3.5	5	12	25	3
101501	4.0	32	20	3	1
105131	0.0	0	0	0	0
105131	1.0	0	1	8	10
105131	1.5	0	3	16	8
105131	2.0	2	9	14	3

ตารางที่ ค-3 (ต่อ)

Attribute	Parameter	High	Good	Middle	Low
105131	2.5	6	14	4	1
105131	3.0	13	7	3	0
105131	3.5	9	2	0	0
105131	4.0	8	0	0	0
202101	0.0	0	0	0	0
202101	1.0	0	0	0	0
202101	1.5	0	0	0	3
202101	2.0	0	0	2	10
202101	2.5	0	1	11	5
202101	3.0	2	8	15	1
202101	3.5	7	13	15	3
202101	4.0	29	14	2	0
202207	0.0	0	0	0	0
202207	1.0	0	0	0	0
202207	1.5	0	0	0	0
202207	2.0	0	1	4	12
202207	2.5	1	7	22	8
202207	3.0	12	16	15	2
202207	3.5	8	11	4	0
202207	4.0	17	1	0	0
202211	0.0	0	0	0	0
202211	1.0	0	0	0	2
202211	1.5	0	0	0	1
202211	2.0	0	0	1	7
202211	2.5	0	0	7	5
202211	3.0	2	7	22	7
202211	3.5	13	15	11	0
202211	4.0	23	14	4	0
202506	0.0	0	0	0	4
202506	1.0	0	2	6	3
202506	1.5	0	3	14	8
202506	2.0	0	8	12	6
202506	2.5	3	8	10	1
202506	3.0	14	11	3	0

ตารางที่ ค-3 (ต่อ)

Attribute	Parameter	High	Good	Middle	Low
202506	3.5	10	3	0	0
202506	4.0	11	1	0	0
202507	0.0	0	1	3	5
202507	1.0	0	2	15	8
202507	1.5	0	3	10	1
202507	2.0	1	10	9	7
202507	2.5	3	8	8	1
202507	3.0	7	7	0	0
202507	3.5	11	1	0	0
202507	4.0	16	4	0	0
901101	0.0	0	0	0	0
901101	1.0	0	0	0	1
901101	1.5	0	0	0	5
901101	2.0	0	1	5	3
901101	2.5	0	4	17	11
901101	3.0	8	7	18	2
901101	3.5	16	19	5	0
901101	4.0	14	5	0	0
903301	0.0	0	0	0	0
903301	1.0	0	0	0	1
903301	1.5	0	0	0	0
903301	2.0	0	0	0	2
903301	2.5	0	1	2	4
903301	3.0	1	7	18	8
903301	3.5	8	11	24	7
903301	4.0	29	17	1	0
grade_T1	VL	0	0	0	4
grade_T1	L	0	1	4	9
grade_T1	M	0	2	26	7
grade_T1	H	4	22	12	2
grade_T1	VH	34	11	3	0
grade_T2	VL	0	0	0	0
grade_T2	L	0	0	1	2

ตารางที่ ค-3 (ต่อ)

Attribute	Parameter	High	Good	Middle	Low
grade_T2	M	0	0	7	16
grade_T2	H	0	17	37	4
grade_T2	VH	38	19	0	0
grade_T3	VL	0	0	0	1
grade_T3	L	0	0	3	12
grade_T3	M	0	3	30	8
grade_T3	H	8	29	12	1
grade_T3	VH	30	4	0	0
grade_T4	VL	0	0	0	0
grade_T4	L	0	0	0	7
grade_T4	M	0	2	23	14
grade_T4	H	4	22	21	1
grade_T4	VH	34	12	1	0
grade_T5	VL	0	0	0	0
grade_T5	L	0	0	0	4
grade_T5	M	0	0	14	16
grade_T5	H	5	18	28	2
grade_T5	VH	33	18	3	0

2. ขั้นตอนการคำนวณความน่าจะเป็นของข้อมูลแต่ละคลาส

2.1 คำนวณความน่าจะเป็น P(C) ของคลาส GPAX

P(GPAX=High) = 38/141 =0.269

P(GPAX=Good) = 36/141 =0.255

P(GPAX=Middle) = 45/141 =0.319

P(GPAX=Low) = 22/141 =0.156

2.2 คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2)

2.2.1 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2) เมื่อ GPAX = High

2.2.1.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = VH เมื่อ GPAX = High ผลลัพธ์ 38/38 = 1.000

2.2.1.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = H เมื่อ GPAX = High ผลลัพธ์ 0/38 = 0.000

2.2.1.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = M เมื่อ GPAX = High ผลลัพธ์ 0/38 = 0.000

2.2.1.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = L เมื่อ GPAX = High ผลลัพธ์ 0/38 = 0.000

2.2.1.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = VL เมื่อ GPAX = High ผลลัพธ์ 0/38 = 0.000

2.2.2 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2) เมื่อ GPAX = Good

2.2.2.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = VH เมื่อ GPAX = Good ผลลัพธ์ 19/36 = 0.527

2.2.2.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = H เมื่อ GPAX = Good ผลลัพธ์ 17/36 = 0.472

2.2.2.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = M เมื่อ GPAX = Good ผลลัพธ์ 0/36 = 0.000

2.2.2.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = L เมื่อ GPAX = Good ผลลัพธ์ 0/36 = 0.000

2.2.2.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = VL เมื่อ GPAX = Good ผลลัพธ์ 0/36 = 0.000

2.2.3 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2) เมื่อ GPAX = Middle

2.2.3.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = VH เมื่อ GPAX = Middle ผลลัพธ์ 0/45 = 0.000

2.2.3.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = H เมื่อ GPAX = Middle ผลลัพธ์ 37/45 = 0.822

2.2.3.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = M เมื่อ GPAX = Middle ผลลัพธ์ 7/45 = 0.156

2.2.3.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = L เมื่อ GPAX = Middle ผลลัพธ์ 1/45 = 0.022

2.2.3.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = VL เมื่อ GPAX = Middle ผลลัพธ์ 0/45 = 0.000

2.2.4 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 1 เทอม 2 (Grade_T2) เมื่อ GPAX = Low

2.2.4.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = VH เมื่อ GPAX = Low ผลลัพธ์ 0/22 = 0.000

2.2.4.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = H เมื่อ GPAX = Low ผลลัพธ์ 4/22 = 0.182

2.2.4.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = M เมื่อ GPAX = Low ผลลัพธ์ 16/22 = 0.727

2.2.4.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = L เมื่อ GPAX = Low ผลลัพธ์ 2/22 = 0.091

2.2.4.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T2 = VL เมื่อ GPAX = Low ผลลัพธ์ 0/22 = 0.000

2.3 คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 1 (Grade_T3)

2.3.1 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 1 (Grade_T3) เมื่อ GPAX = High

2.3.1.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = VH เมื่อ GPAX = High

2.3.1.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = H เมื่อ GPAX = High

2.3.1.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = M เมื่อ GPAX = High

2.3.1.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = L เมื่อ GPAX = High

2.3.1.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = VL เมื่อ GPAX = High

2.3.2 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี2เทอม1 (Grade_T3) เมื่อ GPAX = Good

2.3.2.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = VH เมื่อ GPAX = Good

2.3.2.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = H เมื่อ GPAX = Good

2.3.2.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = M เมื่อ GPAX = Good

2.3.2.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = L เมื่อ GPAX = Good

2.3.2.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = VL เมื่อ GPAX = Good

2.3.3 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี2เทอม1 (Grade_T3) เมื่อ GPAX = Middle

2.3.3.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = VH เมื่อ GPAX = Middle

2.3.3.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = H เมื่อ GPAX = Middle

2.3.3.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = M เมื่อ GPAX = Middle

2.3.3.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = L เมื่อ GPAX = Middle

2.3.3.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = VL เมื่อ GPAX = Middle

2.3.4 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี2เทอม1 (Grade_T3) เมื่อ GPAX = Low

2.3.4.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = VH เมื่อ GPAX =Low

2.3.4.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = H เมื่อ GPAX = Low

2.3.4.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = M เมื่อ GPAX = Low

2.3.4.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = L เมื่อ GPAX = Low

2.3.4.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T3 = VL เมื่อ GPAX = Low

2.4 คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 2 เทอม 2 (Grade_T4)

2.4.1 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี2เทอม2 (Grade_T4) เมื่อ GPAX = High

2.4.1.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = VH เมื่อ GPAX = High

2.4.1.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = H เมื่อ GPAX = High

2.4.1.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = M เมื่อ GPAX = High

2.4.1.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = L เมื่อ GPAX = High

2.4.1.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = VL เมื่อ GPAX = High

2.4.2 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี2เทอม2 (Grade_T4) เมื่อ GPAX = Good
2.4.2.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = VH เมื่อ GPAX = Good
2.4.2.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = H เมื่อ GPAX = Good
2.4.2.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = M เมื่อ GPAX = Good
2.4.2.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = L เมื่อ GPAX = Good
2.4.2.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = VL เมื่อ GPAX = Good
2.4.3 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี2เทอม2 (Grade_T4) เมื่อ GPAX = Middle
2.4.3.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = VH เมื่อ GPAX = Middle
2.4.3.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = H เมื่อ GPAX = Middle
2.4.3.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = M เมื่อ GPAX = Middle
2.4.3.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = L เมื่อ GPAX = Middle
2.4.3.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = VL เมื่อ GPAX = Middle
2.4.4 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี2เทอม2 (Grade_T4) เมื่อ GPAX = Low
2.4.4.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = VH เมื่อ GPAX = Low
2.4.4.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = H เมื่อ GPAX = Low
2.4.4.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = M เมื่อ GPAX = Low
2.4.4.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = L เมื่อ GPAX = Low
2.4.4.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T4 = VL เมื่อ GPAX = Low
2.5 คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี1เทอม1(Grade_T1)
2.5.1 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี1เทอม1 (Grade_T1) เมื่อ GPAX = High
2.5.1.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = VH เมื่อ GPAX = High
2.5.1.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = H เมื่อ GPAX = High
2.5.1.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = M เมื่อ GPAX = High
2.5.1.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = L เมื่อ GPAX = High
2.5.1.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = VL เมื่อ GPAX = High
2.5.2 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี1เทอม1 (Grade_T1) เมื่อ GPAX = Good
2.5.2.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = VH เมื่อ GPAX = Good
2.5.2.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = H เมื่อ GPAX = Good
2.5.2.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = M เมื่อ GPAX = Good
2.5.2.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = L เมื่อ GPAX = Good
2.5.2.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = VL เมื่อ GPAX = Good
2.5.3 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี1เทอม1 (Grade_T1) เมื่อ GPAX = Middle
2.5.3.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = VH เมื่อ GPAX = Middle
2.5.3.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = H เมื่อ GPAX = Middle
2.5.3.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = M เมื่อ GPAX = Middle
2.5.3.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = L เมื่อ GPAX = Middle
2.5.3.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = VL เมื่อ GPAX = Middle

2.5.4 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี1เทอม1 (Grade_T1) เมื่อ GPAX = Low

2.5.4.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = VH เมื่อ GPAX = Low

2.5.4.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = H เมื่อ GPAX = Low

2.5.4.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = M เมื่อ GPAX = Low

2.5.4.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = L เมื่อ GPAX = Low

2.5.4.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T1 = VL เมื่อ GPAX = Low

2.6 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202507

2.6.1 ความน่าจะเป็นของผลการเรียนรายวิชา 202507 เมื่อ GPAX = High

2.6.1.1 ความน่าจะเป็นของแอททริบิวต์ 202507 = 4.0 เมื่อ GPAX = High

2.6.1.2 ความน่าจะเป็นของแอททริบิวต์ 202507 = 3.5 เมื่อ GPAX = High

2.6.1.3 ความน่าจะเป็นของแอททริบิวต์ 202507 = 3.0 เมื่อ GPAX = High

2.6.1.4 ความน่าจะเป็นของแอททริบิวต์ 202507 = 2.5 เมื่อ GPAX = High

2.6.1.5 ความน่าจะเป็นของแอททริบิวต์ 202507 = 2.0 เมื่อ GPAX = High

2.6.1.6 ความน่าจะเป็นของแอททริบิวต์ 202507 = 1.5 เมื่อ GPAX = High

2.6.1.7 ความน่าจะเป็นของแอททริบิวต์ 202507 = 1.0 เมื่อ GPAX = High

2.6.1.8 ความน่าจะเป็นของแอททริบิวต์ 202507 = 0.0 เมื่อ GPAX = High

2.6.2 ความน่าจะเป็นของผลการเรียนรายวิชา 202507 เมื่อ GPAX = Good

2.6.2.1 ความน่าจะเป็นของแอททริบิวต์ 202507 = 4.0 เมื่อ GPAX = Good

2.6.2.2 ความน่าจะเป็นของแอททริบิวต์ 202507 = 3.5 เมื่อ GPAX = Good

2.6.2.3 ความน่าจะเป็นของแอททริบิวต์ 202507 = 3.0 เมื่อ GPAX = Good

2.6.2.4 ความน่าจะเป็นของแอททริบิวต์ 202507 = 2.5 เมื่อ GPAX = Good

2.6.2.5 ความน่าจะเป็นของแอททริบิวต์ 202507 = 2.0 เมื่อ GPAX = Good

2.6.2.6 ความน่าจะเป็นของแอททริบิวต์ 202507 = 1.5 เมื่อ GPAX = Good

2.6.2.7 ความน่าจะเป็นของแอททริบิวต์ 202507 = 1.0 เมื่อ GPAX = Good

2.6.2.8 ความน่าจะเป็นของแอททริบิวต์ 202507 = 0.0 เมื่อ GPAX = Good

2.6.3 ความน่าจะเป็นของผลการเรียนรายวิชา 202507 เมื่อ GPAX = Middle

2.6.3.1 ความน่าจะเป็นของแอททริบิวต์ 202507 = 4.0 เมื่อ GPAX = Middle

2.6.3.2 ความน่าจะเป็นของแอททริบิวต์ 202507 = 3.5 เมื่อ GPAX = Middle

2.6.3.3 ความน่าจะเป็นของแอททริบิวต์ 202507 = 3.0 เมื่อ GPAX = Middle

2.6.3.4 ความน่าจะเป็นของแอททริบิวต์ 202507 = 2.5 เมื่อ GPAX = Middle

2.6.3.5 ความน่าจะเป็นของแอททริบิวต์ 202507 = 2.0 เมื่อ GPAX = Middle

2.6.3.6 ความน่าจะเป็นของแอททริบิวต์ 202507 = 1.5 เมื่อ GPAX = Middle

2.6.3.7 ความน่าจะเป็นของแอททริบิวต์ 202507 = 1.0 เมื่อ GPAX = Middle

2.6.3.8 ความน่าจะเป็นของแอททริบิวต์ 202507 = 0.0 เมื่อ GPAX = Middle

2.6.4 ความน่าจะเป็นของผลการเรียนรายวิชา 202507 เมื่อ GPAX = Low

2.6.4.1 ความน่าจะเป็นของแอททริบิวต์ 202507 = 4.0 เมื่อ GPAX =Low

2.6.4.2 ความน่าจะเป็นของแอททริบิวต์ 202507 = 3.5 เมื่อ GPAX = Low

2.6.4.3 ความน่าจะเป็นของแอททริบิวต์ 202507 = 3.0 เมื่อ GPAX = Low

2.6.4.4 ความน่าจะเป็นของแอททริบิวต์ 202507 = 2.5 เมื่อ GPAX = Low

2.6.4.5 ความน่าจะเป็นของแอททริบิวต์ 202507 = 2.0 เมื่อ GPAX = Low

2.6.4.6 ความน่าจะเป็นของแอททริบิวต์ 202507 = 1.5 เมื่อ GPAX = Low

2.6.4.7 ความน่าจะเป็นของแอททริบิวต์ 202507 = 1.0 เมื่อ GPAX = Low

2.6.4.8 ความน่าจะเป็นของแอททริบิวต์ 202507 = 0.0 เมื่อ GPAX = Low

2.7 คำนวณความน่าจะเป็นของผลการเรียนเฉลี่ยปี 3 เทอม 1 (Grade_T5)

2.7.1 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี 3 เทอม 1 (Grade_T5) เมื่อ GPAX = High

2.7.1.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = VH เมื่อ GPAX = High

2.7.1.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = H เมื่อ GPAX = High

2.7.1.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = M เมื่อ GPAX = High

2.7.1.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = L เมื่อ GPAX = High

2.7.1.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = VL เมื่อ GPAX = High

2.7.2 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี3เทอม1 (Grade_T5) เมื่อ GPAX = Good

2.7.2.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = VH เมื่อ GPAX = Good

2.7.2.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = H เมื่อ GPAX = Good

2.7.2.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = M เมื่อ GPAX = Good

2.7.2.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = L เมื่อ GPAX = Good

2.7.2.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = VL เมื่อ GPAX = Good

2.7.3 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี3เทอม1 (Grade_T5) เมื่อ GPAX = Middle

2.7.3.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = VH เมื่อ GPAX = Middle

2.7.3.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = H เมื่อ GPAX = Middle

2.7.3.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = M เมื่อ GPAX = Middle

2.7.3.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = L เมื่อ GPAX = Middle

2.7.3.6 ความน่าจะเป็นของแอททริบิวต์ Grade_T5 = VL เมื่อ GPAX = Middle

2.7.4 ความน่าจะเป็นของผลการเรียนเฉลี่ยปี3เทอม1(Grade_T5) เมื่อ GPAX=Low

2.7.4.1 ความน่าจะเป็นของแอททริบิวต์ Grade_T5=VH เมื่อ GPAX=Low

2.7.4.2 ความน่าจะเป็นของแอททริบิวต์ Grade_T5=H เมื่อ GPAX= Low

2.7.4.3 ความน่าจะเป็นของแอททริบิวต์ Grade_T5=M เมื่อ GPAX= Low

2.7.4.4 ความน่าจะเป็นของแอททริบิวต์ Grade_T5=L เมื่อ GPAX= Low

2.7.4.5 ความน่าจะเป็นของแอททริบิวต์ Grade_T5=VL เมื่อ GPAX= Low

2.8 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202506

2.8.1 ความน่าจะเป็นของผลการเรียนรายวิชา 202506 เมื่อ GPAX = High

2.8.1.1 ความน่าจะเป็นของแอททริบิวต์ 202506 = 4.0 เมื่อ GPAX = High

2.8.1.2 ความน่าจะเป็นของแอททริบิวต์ 202506 = 3.5 เมื่อ GPAX = High

2.8.1.3 ความน่าจะเป็นของแอททริบิวต์ 202506 = 3.0 เมื่อ GPAX = High

2.8.1.4 ความน่าจะเป็นของแอททริบิวต์ 202506 = 2.5 เมื่อ GPAX = High

2.8.1.5 ความน่าจะเป็นของแอททริบิวต์ 202506 = 2.0 เมื่อ GPAX = High

2.8.1.6 ความน่าจะเป็นของแอททริบิวต์ 202506 = 1.5 เมื่อ GPAX = High

2.8.1.7 ความน่าจะเป็นของแอททริบิวต์ 202506 = 1.0 เมื่อ GPAX = High

2.8.1.8 ความน่าจะเป็นของแอททริบิวต์ 202506 = 0.0 เมื่อ GPAX = High

2.8.2 ความน่าจะเป็นของผลการเรียนรายวิชา 202506 เมื่อ GPAX=Good

2.8.2.1 ความน่าจะเป็นของแอททริบิวต์ 202506=4.0 เมื่อ GPAX=Good

2.8.2.2 ความน่าจะเป็นของแอททริบิวต์ 202506=3.5 เมื่อ GPAX=Good

2.8.2.3 ความน่าจะเป็นของแอททริบิวต์ 202506=3.0 เมื่อ GPAX=Good

2.8.2.4 ความน่าจะเป็นของแอททริบิวต์ 202506=2.5 เมื่อ GPAX=Good

2.8.2.5 ความน่าจะเป็นของแอททริบิวต์ 202506=2.0 เมื่อ GPAX=Good

2.8.2.6 ความน่าจะเป็นของแอททริบิวต์ 202506=1.5 เมื่อ GPAX=Good

2.8.2.7 ความน่าจะเป็นของแอททริบิวต์ 202506=1.0 เมื่อ GPAX=Good

2.8.2.8 ความน่าจะเป็นของแอททริบิวต์ 202506 = 0.0 เมื่อ GPAX = Good

2.8.3 ความน่าจะเป็นของผลการเรียนรายวิชา 202506 เมื่อ GPAX = Middle

2.8.3.1 ความน่าจะเป็นของแอททริบิวต์ 202506 = 4.0 เมื่อ GPAX=Middle

2.8.3.2 ความน่าจะเป็นของแอททริบิวต์ 202506 = 3.5 เมื่อ GPAX = Middle

2.8.3.3 ความน่าจะเป็นของแอททริบิวต์ 202506 = 3.0 เมื่อ GPAX = Middle

2.8.3.4 ความน่าจะเป็นของแอททริบิวต์ 202506 = 2.5 เมื่อ GPAX = Middle

2.8.3.5 ความน่าจะเป็นของแอททริบิวต์ 202506 = 2.0 เมื่อ GPAX = Middle

2.8.3.6 ความน่าจะเป็นของแอททริบิวต์ 202506 = 1.5 เมื่อ GPAX = Middle

2.8.3.7 ความน่าจะเป็นของแอททริบิวต์ 202506 = 1.0 เมื่อ GPAX = Middle

2.8.3.8 ความน่าจะเป็นของแอททริบิวต์ 202506 = 0.0 เมื่อ GPAX = Middle

2.8.4 ความน่าจะเป็นของผลการเรียนรายวิชา 202506 เมื่อ GPAX = Low

2.8.4.1 ความน่าจะเป็นของแอททริบิวต์ 202506 = 4.0 เมื่อ GPAX =Low

2.8.4.2 ความน่าจะเป็นของแอททริบิวต์ 202506 = 3.5 เมื่อ GPAX = Low

2.8.4.3 ความน่าจะเป็นของแอททริบิวต์ 202506 = 3.0 เมื่อ GPAX = Low

2.8.4.4 ความน่าจะเป็นของแอททริบิวต์ 202506 = 2.5 เมื่อ GPAX = Low

2.8.4.5 ความน่าจะเป็นของแอททริบิวต์ 202506 = 2.0 เมื่อ GPAX = Low

2.8.4.6 ความน่าจะเป็นของแอททริบิวต์ 202506 = 1.5 เมื่อ GPAX = Low

2.8.4.7 ความน่าจะเป็นของแอททริบิวต์ 202506 = 1.0 เมื่อ GPAX = Low

2.8.4.8 ความน่าจะเป็นของแอททริบิวต์ 202506 = 0.0 เมื่อ GPAX = Low

2.9 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 105131

2.9.1 ความน่าจะเป็นของผลการเรียนรายวิชา 105131 เมื่อ GPAX = High

2.9.1.1 ความน่าจะเป็นของแอททริบิวต์ 105131 = 4.0 เมื่อ GPAX = High

2.9.1.2 ความน่าจะเป็นของแอททริบิวต์ 105131 = 3.5 เมื่อ GPAX = High

2.9.1.3 ความน่าจะเป็นของแอททริบิวต์ 105131 = 3.0 เมื่อ GPAX = High

2.9.1.4 ความน่าจะเป็นของแอททริบิวต์ 105131 = 2.5 เมื่อ GPAX = High

2.9.1.5 ความน่าจะเป็นของแอททริบิวต์ 105131 = 2.0 เมื่อ GPAX = High

2.9.1.6 ความน่าจะเป็นของแอททริบิวต์ 105131 = 1.5 เมื่อ GPAX = High

2.9.1.7 ความน่าจะเป็นของแอททริบิวต์ 105131 = 1.0 เมื่อ GPAX = High

2.9.1.8 ความน่าจะเป็นของแอททริบิวต์ 105131 = 0.0 เมื่อ GPAX = High

2.9.2 ความน่าจะเป็นของผลการเรียนรายวิชา 105131 เมื่อ GPAX = Good

2.9.2.1 ความน่าจะเป็นของแอททริบิวต์ 105131 = 4.0 เมื่อ GPAX = Good

2.9.2.2 ความน่าจะเป็นของแอททริบิวต์ 105131 = 3.5 เมื่อ GPAX = Good

2.9.2.3 ความน่าจะเป็นของแอททริบิวต์ 105131 = 3.0 เมื่อ GPAX = Good

2.9.2.4 ความน่าจะเป็นของแอททริบิวต์ 105131 = 2.5 เมื่อ GPAX = Good

2.9.2.5 ความน่าจะเป็นของแอททริบิวต์ 105131 = 2.0 เมื่อ GPAX = Good

2.9.2.6 ความน่าจะเป็นของแอททริบิวต์ 105131 = 1.5 เมื่อ GPAX = Good

2.9.2.7 ความน่าจะเป็นของแอททริบิวต์ 105131 = 1.0 เมื่อ GPAX = Good

2.9.2.8 ความน่าจะเป็นของแอททริบิวต์ 105131 = 0.0 เมื่อ GPAX = Good

2.9.3 ความน่าจะเป็นของผลการเรียนรายวิชา 105131 เมื่อ GPAX = Middle

2.9.3.1 ความน่าจะเป็นของแอททริบิวต์ 105131 = 4.0 เมื่อ GPAX = Middle

2.9.3.2 ความน่าจะเป็นของแอททริบิวต์ 105131 = 3.5 เมื่อ GPAX = Middle

2.9.3.3 ความน่าจะเป็นของแอททริบิวต์ 105131 = 3.0 เมื่อ GPAX = Middle

2.9.3.4 ความน่าจะเป็นของแอททริบิวต์ 105131 = 2.5 เมื่อ GPAX = Middle

2.9.3.5 ความน่าจะเป็นของแอททริบิวต์ 105131 = 2.0 เมื่อ GPAX = Middle

2.9.3.6 ความน่าจะเป็นของแอททริบิวต์ 105131 = 1.5 เมื่อ GPAX = Middle

2.9.3.7 ความน่าจะเป็นของแอททริบิวต์ 105131 =1 .0 เมื่อ GPAX = Middle

2.9.3.8 ความน่าจะเป็นของแอททริบิวต์ 105131 = 0.0 เมื่อ GPAX = Middle

2.9.4 ความน่าจะเป็นของผลการเรียนรายวิชา 105131 เมื่อ GPAX = Low

2.9.4.1 ความน่าจะเป็นของแอททริบิวต์ 105131 = 4.0 เมื่อ GPAX = Low

2.9.4.2 ความน่าจะเป็นของแอททริบิวต์ 105131 = 3.5 เมื่อ GPAX = Low

2.9.4.3 ความน่าจะเป็นของแอททริบิวต์ 105131 = 3.0 เมื่อ GPAX = Low

2.9.4.4 ความน่าจะเป็นของแอททริบิวต์ 105131 = 2.5 เมื่อ GPAX = Low

2.9.4.5 ความน่าจะเป็นของแอททริบิวต์ 105131 = 2.0 เมื่อ GPAX = Low

2.9.4.6 ความน่าจะเป็นของแอททริบิวต์ 105131 = 1.5 เมื่อ GPAX = Low

2.9.4.7 ความน่าจะเป็นของแอททริบิวต์ 105131 = 1.0 เมื่อ GPAX = Low

2.9.4.8 ความน่าจะเป็นของแอททริบิวต์ 105131 = 0.0 เมื่อ GPAX = Low

2.10 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202101

2.10.1 ความน่าจะเป็นของผลการเรียนรายวิชา 202101 เมื่อ GPAX = High

2.10.1.1 ความน่าจะเป็นของแอททริบิวต์ 202101 = 4.0 เมื่อ GPAX = High

2.10.1.2 ความน่าจะเป็นของแอททริบิวต์ 202101 = 3.5 เมื่อ GPAX = High

2.10.1.3 ความน่าจะเป็นของแอททริบิวต์ 202101 = 3.0 เมื่อ GPAX = High

2.10.1.4 ความน่าจะเป็นของแอททริบิวต์ 202101 = 2.5 เมื่อ GPAX = High

2.10.1.5 ความน่าจะเป็นของแอททริบิวต์ 202101 = 2.0 เมื่อ GPAX = High

2.10.1.6 ความน่าจะเป็นของแอททริบิวต์ 202101 = 1.5 เมื่อ GPAX = High

2.10.1.7 ความน่าจะเป็นของแอททริบิวต์ 202101 = 1.0 เมื่อ GPAX = High

2.10.1.8 ความน่าจะเป็นของแอททริบิวต์ 202101 = 0.0 เมื่อ GPAX = High

2.10.2 ความน่าจะเป็นของผลการเรียนรายวิชา 202101 เมื่อ GPAX = Good

2.10.2.1 ความน่าจะเป็นของแอททริบิวต์ 202101 = 4.0 เมื่อ GPAX = Good

2.10.2.2 ความน่าจะเป็นของแอททริบิวต์ 202101 = 3.5 เมื่อ GPAX = Good

2.10.2.3 ความน่าจะเป็นของแอททริบิวต์ 202101 = 3.0 เมื่อ GPAX = Good

2.10.2.4 ความน่าจะเป็นของแอททริบิวต์ 202101 = 2.5 เมื่อ GPAX = Good

2.10.2.5 ความน่าจะเป็นของแอททริบิวต์ 202101 = 2.0 เมื่อ GPAX = Good

2.10.2.6 ความน่าจะเป็นของแอททริบิวต์ 202101 = 1.5 เมื่อ GPAX = Good

2.10.2.7 ความน่าจะเป็นของแอททริบิวต์ 202101 = 1.0 เมื่อ GPAX = Good

2.10.2.8 ความน่าจะเป็นของแอททริบิวต์ 202101 = 0.0 เมื่อ GPAX = Good

2.10.3 ความน่าจะเป็นของผลการเรียนรายวิชา 202101 เมื่อ GPAX = Middle

2.10.3.1 ความน่าจะเป็นของแอททริบิวต์ 202101 = 4.0 เมื่อ GPAX = Middle

2.10.3.2 ความน่าจะเป็นของแอททริบิวต์ 202101 = 3.5 เมื่อ GPAX = Middle

2.10.3.3 ความน่าจะเป็นของแอททริบิวต์ 202101 = 3.0 เมื่อ GPAX = Middle

2.10.3.4 ความน่าจะเป็นของแอททริบิวต์ 202101 = 2.5 เมื่อ GPAX = Middle

2.10.3.5 ความน่าจะเป็นของแอททริบิวต์ 202101 = 2.0 เมื่อ GPAX = Middle

2.10.3.6 ความน่าจะเป็นของแอททริบิวต์ 202101 = 1.5 เมื่อ GPAX = Middle

2.10.3.7 ความน่าจะเป็นของแอททริบิวต์ 202101 = 1.0 เมื่อ GPAX = Middle

2.10.3.8 ความน่าจะเป็นของแอททริบิวต์ 202101 = 0.0 เมื่อ GPAX = Middle

2.10.4 ความน่าจะเป็นของผลการเรียนรายวิชา 202101 เมื่อ GPAX = Low

2.10.4.1 ความน่าจะเป็นของแอททริบิวต์ 202101 = 4.0 เมื่อ GPAX = Low

2.10.4.2 ความน่าจะเป็นของแอททริบิวต์ 202101 = 3.5 เมื่อ GPAX = Low

2.10.4.3 ความน่าจะเป็นของแอททริบิวต์ 202101 = 3.0 เมื่อ GPAX = Low

2.10.4.4 ความน่าจะเป็นของแอททริบิวต์ 202101 = 2.5 เมื่อ GPAX = Low

2.10.4.5 ความน่าจะเป็นของแอททริบิวต์ 202101 = 2.0 เมื่อ GPAX = Low

2.10.4.6 ความน่าจะเป็นของแอททริบิวต์ 202101 = 1.5 เมื่อ GPAX = Low

2.10.4.7 ความน่าจะเป็นของแอททริบิวต์ 202101 = 1.0 เมื่อ GPAX = Low

2.10.4.8 ความน่าจะเป็นของแอททริบิวต์ 202101 = 0.0 เมื่อ GPAX = Low

2.11 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 901101

2.11.1 ความน่าจะเป็นของผลการเรียนรายวิชา 901101 เมื่อ GPAX = High

2.11.1.1 ความน่าจะเป็นของแอททริบิวต์ 901101 = 4.0 เมื่อ GPAX = High

2.11.1.2 ความน่าจะเป็นของแอททริบิวต์ 901101 = 3.5 เมื่อ GPAX = High

2.11.1.3 ความน่าจะเป็นของแอททริบิวต์ 901101 = 3.0 เมื่อ GPAX = High

2.11.1.4 ความน่าจะเป็นของแอททริบิวต์ 901101 = 2.5 เมื่อ GPAX = High

2.11.1.5 ความน่าจะเป็นของแอททริบิวต์ 901101 = 2.0 เมื่อ GPAX = High

2.11.1.6 ความน่าจะเป็นของแอททริบิวต์ 901101 = 1.5 เมื่อ GPAX = High

2.11.1.7 ความน่าจะเป็นของแอททริบิวต์ 901101 = 1.0 เมื่อ GPAX = High

2.11.1.8 ความน่าจะเป็นของแอททริบิวต์ 901101 = 0.0 เมื่อ GPAX = High

2.11.2 ความน่าจะเป็นของผลการเรียนรายวิชา 901101 เมื่อ GPAX = Good

2.11.2.1 ความน่าจะเป็นของแอททริบิวต์ 901101 = 4.0 เมื่อ GPAX = Good

2.11.2.2 ความน่าจะเป็นของแอททริบิวต์ 901101 = 3.5 เมื่อ GPAX = Good

2.11.2.3 ความน่าจะเป็นของแอททริบิวต์ 901101 = 3.0 เมื่อ GPAX = Good

2.11.2.4 ความน่าจะเป็นของแอททริบิวต์ 901101 = 2.5 เมื่อ GPAX = Good

2.11.2.5 ความน่าจะเป็นของแอททริบิวต์ 901101 = 2.0 เมื่อ GPAX = Good

2.11.2.6 ความน่าจะเป็นของแอททริบิวต์ 901101 = 1.5 เมื่อ GPAX = Good

2.11.2.7 ความน่าจะเป็นของแอททริบิวต์ 901101 = 1.0 เมื่อ GPAX = Good

2.11.2.8 ความน่าจะเป็นของแอททริบิวต์ 901101 = 0.0 เมื่อ GPAX = Good

2.11.3 ความน่าจะเป็นของผลการเรียนรายวิชา 901101 เมื่อ GPAX = Middle

2.11.3.1 ความน่าจะเป็นของแอททริบิวต์ 901101 = 4.0 เมื่อ GPAX = Middle

2.11.3.2 ความน่าจะเป็นของแอททริบิวต์ 901101 = 3.5 เมื่อ GPAX = Middle

2.11.3.3 ความน่าจะเป็นของแอททริบิวต์ 901101 = 3.0 เมื่อ GPAX = Middle

2.11.3.4 ความน่าจะเป็นของแอททริบิวต์ 901101 = 2.5 เมื่อ GPAX = Middle

2.11.3.5 ความน่าจะเป็นของแอททริบิวต์ 901101 = 2.0 เมื่อ GPAX = Middle

2.11.3.6 ความน่าจะเป็นของแอททริบิวต์ 901101 = 1.5 เมื่อ GPAX = Middle

2.11.3.7 ความน่าจะเป็นของแอททริบิวต์ 901101 = 1.0 เมื่อ GPAX = Middle

2.11.3.8 ความน่าจะเป็นของแอททริบิวต์ 901101 = 0.0 เมื่อ GPAX = Middle

2.11.4 ความน่าจะเป็นของผลการเรียนรายวิชา 901101 เมื่อ GPAX = Low

2.11.4.1 ความน่าจะเป็นของแอททริบิวต์ 901101 = 4.0 เมื่อ GPAX = Low

2.11.4.2 ความน่าจะเป็นของแอททริบิวต์ 901101 = 3.5 เมื่อ GPAX = Low

2.11.4.3 ความน่าจะเป็นของแอททริบิวต์ 901101 = 3.0 เมื่อ GPAX = Low

2.11.4.4 ความน่าจะเป็นของแอททริบิวต์ 901101 = 2.5 เมื่อ GPAX = Low

2.11.4.5 ความน่าจะเป็นของแอททริบิวต์ 901101 = 2.0 เมื่อ GPAX = Low

2.11.4.6 ความน่าจะเป็นของแอททริบิวต์ 901101 = 1.5 เมื่อ GPAX = Low
2.11.4.7 ความน่าจะเป็นของแอททริบิวต์ 901101 = 1.0 เมื่อ GPAX = Low
2.11.4.8 ความน่าจะเป็นของแอททริบิวต์ 901101 = 0.0 เมื่อ GPAX = Low

2.12 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202211

2.12.1 ความน่าจะเป็นของแอททริบิวต์ 202211 เมื่อ GPAX=High
2.12.1.1 ความน่าจะเป็นของแอททริบิวต์ 202211 = 4.0 เมื่อ GPAX = High
2.12.1.2 ความน่าจะเป็นของแอททริบิวต์ 202211 = 3.5 เมื่อ GPAX = High
2.12.1.3 ความน่าจะเป็นของแอททริบิวต์ 202211 = 3.0 เมื่อ GPAX = High
2.12.1.4 ความน่าจะเป็นของแอททริบิวต์ 202211 = 2.5 เมื่อ GPAX = High
2.12.1.5 ความน่าจะเป็นของแอททริบิวต์ 202211 = 2.0 เมื่อ GPAX = High
2.12.1.6 ความน่าจะเป็นของแอททริบิวต์ 202211 = 1.5 เมื่อ GPAX = High
2.12.1.7 ความน่าจะเป็นของแอททริบิวต์ 202211 = 1.0 เมื่อ GPAX = High
2.12.1.8 ความน่าจะเป็นของแอททริบิวต์ 202211 = 0.0 เมื่อ GPAX = High

2.12.2 ความน่าจะเป็นของผลการเรียนรายวิชา 202211 เมื่อ GPAX = Good
2.12.2.1 ความน่าจะเป็นของแอททริบิวต์ 202211 = 4.0 เมื่อ GPAX = Good
2.12.2.2 ความน่าจะเป็นของแอททริบิวต์ 202211 = 3.5 เมื่อ GPAX = Good
2.12.2.3 ความน่าจะเป็นของแอททริบิวต์ 202211 = 3.0 เมื่อ GPAX = Good
2.12.2.4 ความน่าจะเป็นของแอททริบิวต์ 202211 = 2.5 เมื่อ GPAX = Good
2.12.2.5 ความน่าจะเป็นของแอททริบิวต์ 202211 = 2.0 เมื่อ GPAX = Good
2.12.2.6 ความน่าจะเป็นของแอททริบิวต์ 202211 = 1.5 เมื่อ GPAX = Good
2.12.2.7 ความน่าจะเป็นของแอททริบิวต์ 202211 = 1.0 เมื่อ GPAX = Good
2.12.2.8 ความน่าจะเป็นของแอททริบิวต์ 202211 = 0.0 เมื่อ GPAX = Good

2.12.3 ความน่าจะเป็นของผลการเรียนรายวิชา 202211 เมื่อ GPAX = Middle
2.12.3.1 ความน่าจะเป็นของแอททริบิวต์ 202211 = 4.0 เมื่อ GPAX = Middle
2.12.3.2 ความน่าจะเป็นของแอททริบิวต์ 202211 = 3.5 เมื่อ GPAX = Middle
2.12.3.3 ความน่าจะเป็นของแอททริบิวต์ 202211 = 3.0 เมื่อ GPAX = Middle
2.12.3.4 ความน่าจะเป็นของแอททริบิวต์ 202211 = 2.5 เมื่อ GPAX = Middle
2.12.3.5 ความน่าจะเป็นของแอททริบิวต์ 202211 = 2.0 เมื่อ GPAX = Middle
2.12.3.6 ความน่าจะเป็นของแอททริบิวต์ 202211 = 1.5 เมื่อ GPAX = Middle
2.12.3.7 ความน่าจะเป็นของแอททริบิวต์ 202211 = 1.0 เมื่อ GPAX = Middle
2.12.3.8 ความน่าจะเป็นของแอททริบิวต์ 202211 = 0.0 เมื่อ GPAX = Middle

2.12.4 ความน่าจะเป็นของผลการเรียนรายวิชา 202211 เมื่อ GPAX = Low
2.12.4.1 ความน่าจะเป็นของแอททริบิวต์ 202211 = 4.0 เมื่อ GPAX = Low
2.12.4.2 ความน่าจะเป็นของแอททริบิวต์ 202211 = 3.5 เมื่อ GPAX = Low
2.12.4.3 ความน่าจะเป็นของแอททริบิวต์ 202211 = 3.0 เมื่อ GPAX = Low
2.12.4.4 ความน่าจะเป็นของแอททริบิวต์ 202211 = 2.5 เมื่อ GPAX = Low

2.12.4.5 ความน่าจะเป็นของแอททริบิวต์ 202211 = 2.0 เมื่อ GPAX = Low
2.12.4.6 ความน่าจะเป็นของแอททริบิวต์ 202211 = 1.5 เมื่อ GPAX = Low
2.12.4.7 ความน่าจะเป็นของแอททริบิวต์ 202211 = 1.0 เมื่อ GPAX = Low
2.12.4.8 ความน่าจะเป็นของแอททริบิวต์ 202211 = 0.0 เมื่อ GPAX = Low

2.13 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 202207

2.13.1 ความน่าจะเป็นของแอททริบิวต์ 202207 เมื่อ GPAX = High
2.13.1.1 ความน่าจะเป็นของแอททริบิวต์ 202207 = 4.0 เมื่อ GPAX = High
2.13.1.2 ความน่าจะเป็นของแอททริบิวต์ 202207 = 3.5 เมื่อ GPAX = High
2.13.1.3 ความน่าจะเป็นของแอททริบิวต์ 202207 = 3.0 เมื่อ GPAX = High
2.13.1.4 ความน่าจะเป็นของแอททริบิวต์ 202207 = 2.5 เมื่อ GPAX = High
2.13.1.5 ความน่าจะเป็นของแอททริบิวต์ 202207 = 2.0 เมื่อ GPAX = High
2.13.1.6 ความน่าจะเป็นของแอททริบิวต์ 202207 = 1.5 เมื่อ GPAX = High
2.13.1.7 ความน่าจะเป็นของแอททริบิวต์ 202207 = 1.0 เมื่อ GPAX = High
2.13.1.8 ความน่าจะเป็นของแอททริบิวต์ 202207 = 0.0 เมื่อ GPAX = High

2.13.2 ความน่าจะเป็นของผลการเรียนรายวิชา 202207 เมื่อ GPAX = Good
2.13.2.1 ความน่าจะเป็นของแอททริบิวต์ 202207 = 4.0 เมื่อ GPAX = Good
2.13.2.2 ความน่าจะเป็นของแอททริบิวต์ 202207 = 3.5 เมื่อ GPAX = Good
2.13.2.3 ความน่าจะเป็นของแอททริบิวต์ 202207 = 3.0 เมื่อ GPAX = Good
2.13.2.4 ความน่าจะเป็นของแอททริบิวต์ 202207 = 2.5 เมื่อ GPAX = Good
2.13.2.5 ความน่าจะเป็นของแอททริบิวต์ 202207 = 2.0 เมื่อ GPAX = Good
2.13.2.6 ความน่าจะเป็นของแอททริบิวต์ 202207 = 1.5 เมื่อ GPAX = Good
2.13.2.7 ความน่าจะเป็นของแอททริบิวต์ 202207 = 1.0 เมื่อ GPAX = Good
2.13.2.8 ความน่าจะเป็นของแอททริบิวต์ 202207 = 0.0 เมื่อ GPAX = Good

2.13.3 ความน่าจะเป็นของผลการเรียนรายวิชา 202207 เมื่อ GPAX = Middle
2.13.3.1 ความน่าจะเป็นของแอททริบิวต์ 202207 = 4.0 เมื่อ GPAX = Middle
2.13.3.2 ความน่าจะเป็นของแอททริบิวต์ 202207 = 3.5 เมื่อ GPAX = Middle
2.13.3.3 ความน่าจะเป็นของแอททริบิวต์ 202207 = 3.0 เมื่อ GPAX = Middle
2.13.3.4 ความน่าจะเป็นของแอททริบิวต์ 202207 = 2.5 เมื่อ GPAX = Middle
2.13.3.5 ความน่าจะเป็นของแอททริบิวต์ 202207 = 2.0 เมื่อ GPAX = Middle
2.13.3.6 ความน่าจะเป็นของแอททริบิวต์ 202207 = 1.5 เมื่อ GPAX = Middle
2.13.3.7 ความน่าจะเป็นของแอททริบิวต์ 202207 = 1.0 เมื่อ GPAX = Middle
2.13.3.8 ความน่าจะเป็นของแอททริบิวต์ 202207 = 0.0 เมื่อ GPAX = Middle

2.13.4 ความน่าจะเป็นของผลการเรียนรายวิชา 202207 เมื่อ GPAX = Low
2.13.4.1 ความน่าจะเป็นของแอททริบิวต์ 202207 = 4.0 เมื่อ GPAX=Low
2.13.4.2 ความน่าจะเป็นของแอททริบิวต์ 202207 = 3.5 เมื่อ GPAX= Low
2.13.4.3 ความน่าจะเป็นของแอททริบิวต์ 202207 = 3.0 เมื่อ GPAX= Low

2.13.4.4 ความน่าจะเป็นของแอททริบิวต์ 202207 = 2.5 เมื่อ GPAX= Low
2.13.4.5 ความน่าจะเป็นของแอททริบิวต์ 202207 = 2.0 เมื่อ GPAX= Low
2.13.4.6 ความน่าจะเป็นของแอททริบิวต์ 202207 = 1.5 เมื่อ GPAX= Low
2.13.4.7 ความน่าจะเป็นของแอททริบิวต์ 202207 = 1.0 เมื่อ GPAX= Low
2.13.4.8 ความน่าจะเป็นของแอททริบิวต์ 202207 = 0.0 เมื่อ GPAX= Low

2.14 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 101501
2.14.1 ความน่าจะเป็นของผลการเรียนรายวิชา 101501 เมื่อ GPAX = High
2.14.1.1 ความน่าจะเป็นของแอททริบิวต์ 101501=4.0 เมื่อ GPAX = High
2.14.1.2 ความน่าจะเป็นของแอททริบิวต์ 101501=3.5 เมื่อ GPAX = High
2.14.1.3 ความน่าจะเป็นของแอททริบิวต์ 101501=3.0 เมื่อ GPAX = High
2.14.1.4 ความน่าจะเป็นของแอททริบิวต์ 101501=2.5 เมื่อ GPAX = High
2.14.1.5 ความน่าจะเป็นของแอททริบิวต์ 101501=2.0 เมื่อ GPAX = High
2.14.1.6 ความน่าจะเป็นของแอททริบิวต์ 101501=1.5 เมื่อ GPAX = High
2.14.1.7 ความน่าจะเป็นของแอททริบิวต์ 101501=1.0 เมื่อ GPAX = High
2.14.1.8 ความน่าจะเป็นของแอททริบิวต์ 101501=0.0 เมื่อ GPAX = High
2.14.2 ความน่าจะเป็นของผลการเรียนรายวิชา 101501 เมื่อ GPAX = Good
2.14.2.1 ความน่าจะเป็นของแอททริบิวต์ 101501=4.0 เมื่อ GPAX = Good
2.14.2.2 ความน่าจะเป็นของแอททริบิวต์ 101501 = 3.5 เมื่อ GPAX = Good
2.14.2.3 ความน่าจะเป็นของแอททริบิวต์ 101501 = 3.0 เมื่อ GPAX = Good
2.14.2.4 ความน่าจะเป็นของแอททริบิวต์ 101501 = 2.5 เมื่อ GPAX = Good
2.14.2.5 ความน่าจะเป็นของแอททริบิวต์ 101501 = 2.0 เมื่อ GPAX = Good
2.14.2.6 ความน่าจะเป็นของแอททริบิวต์ 101501 = 1.5 เมื่อ GPAX = Good
2.14.2.7 ความน่าจะเป็นของแอททริบิวต์ 101501 = 1.0 เมื่อ GPAX = Good
2.14.2.8 ความน่าจะเป็นของแอททริบิวต์ 101501 = 0.0 เมื่อ GPAX = Good
2.14.3 ความน่าจะเป็นของผลการเรียนรายวิชา 101501 เมื่อ GPAX = Middle
2.14.3.1 ความน่าจะเป็นของแอททริบิวต์ 101501 = 4.0 เมื่อ GPAX = Middle
2.14.3.2 ความน่าจะเป็นของแอททริบิวต์ 101501 = 3.5 เมื่อ GPAX = Middle
2.14.3.3 ความน่าจะเป็นของแอททริบิวต์ 101501 = 3.0 เมื่อ GPAX = Middle
2.14.3.4 ความน่าจะเป็นของแอททริบิวต์ 101501 = 2.5 เมื่อ GPAX = Middle
2.14.3.5 ความน่าจะเป็นของแอททริบิวต์ 101501 = 2.0 เมื่อ GPAX = Middle
2.14.3.6 ความน่าจะเป็นของแอททริบิวต์ 101501 = 1.5 เมื่อ GPAX = Middle
2.14.3.7 ความน่าจะเป็นของแอททริบิวต์ 101501 = 1.0 เมื่อ GPAX = Middle
2.14.3.8 ความน่าจะเป็นของแอททริบิวต์ 101501 = 0.0 เมื่อ GPAX = Middle
2.14.4 ความน่าจะเป็นของผลการเรียนรายวิชา 101501 เมื่อ GPAX = Low
2.14.4.1 ความน่าจะเป็นของแอททริบิวต์ 101501 = 4.0 เมื่อ GPAX = Low
2.14.4.2 ความน่าจะเป็นของแอททริบิวต์ 101501 = 3.5 เมื่อ GPAX = Low

2.14.4.3 ความน่าจะเป็นของแอททริบิวต์ 101501 = 3.0 เมื่อ GPAX = Low
2.14.4.4 ความน่าจะเป็นของแอททริบิวต์ 101501 = 2.5 เมื่อ GPAX = Low
2.14.4.5 ความน่าจะเป็นของแอททริบิวต์ 101501 = 2.0 เมื่อ GPAX = Low
2.14.4.6 ความน่าจะเป็นของแอททริบิวต์ 101501 = 1.5 เมื่อ GPAX = Low
2.14.4.7 ความน่าจะเป็นของแอททริบิวต์ 101501 = 1.0 เมื่อ GPAX = Low
2.14.4.8 ความน่าจะเป็นของแอททริบิวต์ 101501 = 0.0 เมื่อ GPAX = Low

2.15 คำนวณความน่าจะเป็นของผลการเรียนรายวิชา 903301

2.15.1 ความน่าจะเป็นของผลการเรียนรายวิชา 903301 เมื่อ GPAX = High
2.15.1.1 ความน่าจะเป็นของแอททริบิวต์ 903301 = 4.0 เมื่อ GPAX = High
2.15.1.2 ความน่าจะเป็นของแอททริบิวต์ 903301 = 3.5 เมื่อ GPAX = High
2.15.1.3 ความน่าจะเป็นของแอททริบิวต์ 903301 = 3.0 เมื่อ GPAX = High
2.15.1.4 ความน่าจะเป็นของแอททริบิวต์ 903301 = 2.5 เมื่อ GPAX = High
2.15.1.5 ความน่าจะเป็นของแอททริบิวต์ 903301 = 2.0 เมื่อ GPAX = High
2.15.1.6 ความน่าจะเป็นของแอททริบิวต์ 903301 = 1.5 เมื่อ GPAX = High
2.15.1.7 ความน่าจะเป็นของแอททริบิวต์ 903301 = 1.0 เมื่อ GPAX = High
2.15.1.8 ความน่าจะเป็นของแอททริบิวต์ 903301 = 0.0 เมื่อ GPAX = High

2.15.2 ความน่าจะเป็นของผลการเรียนรายวิชา 903301เมื่อ GPAX = Good
2.15.2.1 ความน่าจะเป็นของแอททริบิวต์ 903301 = 4.0 เมื่อ GPAX = Good
2.15.2.2 ความน่าจะเป็นของแอททริบิวต์ 903301 = 3.5 เมื่อ GPAX = Good
2.15.2.3 ความน่าจะเป็นของแอททริบิวต์ 903301 = 3.0 เมื่อ GPAX = Good
2.15.2.4 ความน่าจะเป็นของแอททริบิวต์ 903301 = 2.5 เมื่อ GPAX = Good
2.15.2.5 ความน่าจะเป็นของแอททริบิวต์ 903301 = 2.0 เมื่อ GPAX = Good
2.15.2.6 ความน่าจะเป็นของแอททริบิวต์ 903301 = 1.5 เมื่อ GPAX = Good
2.15.2.7 ความน่าจะเป็นของแอททริบิวต์ 903301 = 1.0 เมื่อ GPAX = Good
2.15.2.8 ความน่าจะเป็นของแอททริบิวต์ 903301 = 0.0 เมื่อ GPAX = Good

2.15.3 ความน่าจะเป็นของผลการเรียนรายวิชา 903301เมื่อ GPAX = Middle
2.15.3.1 ความน่าจะเป็นของแอททริบิวต์ 903301 = 4.0 เมื่อ GPAX = Middle
2.15.3.2 ความน่าจะเป็นของแอททริบิวต์ 903301 = 3.5 เมื่อ GPAX = Middle
2.15.3.3 ความน่าจะเป็นของแอททริบิวต์ 903301 = 3.0 เมื่อ GPAX = Middle
2.15.3.4 ความน่าจะเป็นของแอททริบิวต์ 903301 = 2.5 เมื่อ GPAX = Middle
2.15.3.5 ความน่าจะเป็นของแอททริบิวต์ 903301 = 2.0 เมื่อ GPAX = Middle
2.15.3.6 ความน่าจะเป็นของแอททริบิวต์ 903301 = 1.5 เมื่อ GPAX = Middle
2.15.3.7 ความน่าจะเป็นของแอททริบิวต์ 903301 = 1.0 เมื่อ GPAX = Middle
2.15.3.8 ความน่าจะเป็นของแอททริบิวต์ 903301 = 0.0 เมื่อ GPAX = Middle

2.15.4 ความน่าจะเป็นของผลการเรียนรายวิชา 903301เมื่อ GPAX = Low

2.15.4.1 ความน่าจะเป็นของแอททริบิวต์ 903301 = 4.0 เมื่อ GPAX = Low

2.15.4.2 ความน่าจะเป็นของแอททริบิวต์ 903301 = 3.5 เมื่อ GPAX = Low

2.15.4.3 ความน่าจะเป็นของแอททริบิวต์ 903301 = 3.0 เมื่อ GPAX = Low

2.15.4.4 ความน่าจะเป็นของแอททริบิวต์ 903301 = 2.5 เมื่อ GPAX = Low

2.15.4.5 ความน่าจะเป็นของแอททริบิวต์ 903301 = 2.0 เมื่อ GPAX = Low

2.15.4.6 ความน่าจะเป็นของแอททริบิวต์ 903301 = 1.5 เมื่อ GPAX = Low

2.15.4.7 ความน่าจะเป็นของแอททริบิวต์ 903301 = 1.0 เมื่อ GPAX = Low

2.15.4.8 ความน่าจะเป็นของแอททริบิวต์ 903301 = 0.0 เมื่อ GPAX = Low

จากความน่าจะเป็นที่คำนวณได้สามารถสรุปออกมาได้ดังตารางที่ ค-4 และตารางที่ ค-5 และเรียกตารางดังกล่าวว่าตัวแบบการพยากรณ์ผลการเรียนอัจฉริยะด้วยเทคนิค Naïve Bayes ซึ่งสามารถนำตัวแบบการพยากรณ์ผลการเรียนดังกล่าวไปทำนายความน่าจะเป็นต่าง ๆ ไปแทนสมการของ Bayes ได้เลย

ตัวแบบพยากรณ์ผลการเรียนอัจฉริยะด้วยเทคนิค Naïve Bayes สามารถทำนายความน่าจะเป็นของผลการเรียนนักศึกษาในอนาคตได้โดยการนำข้อมูลที่ไม่ทราบคลาสหาความน่าจะเป็นด้วยการแทนสมการของ Bayes ได้เลย และสามารถสรุปตัวแบบพยากรณ์ผลการเรียนอัจฉริยะด้วยเทคนิค Naïve Bayes ที่ผ่านการปรับเรียบด้วยวิธี Laplace Smoothing ได้ดังตารางที่ ค-5

ตารางที่ ค-4 ความน่าจะเป็น P(C) ของคลาส GPAX

attribute	Class	P(C)
GPAX	High	0.266
	Good	0.254
	Middle	0.307
	Low	0.171

ตารางที่ ค-5 ตัวแบบพยากรณ์ผลการเรียนด้วยเทคนิค Naïve Bayes

attribute	Parameter	High	Good	Middle	Low
101501	value=0.0	0.021	0.022	0.018	0.033
101501	value=1.0	0.021	0.022	0.018	0.033
101501	value=1.5	0.021	0.022	0.018	0.033
101501	value=2.0	0.021	0.022	0.037	0.066
101501	value=2.5	0.021	0.022	0.056	0.233
101501	value=3.0	0.043	0.113	0.283	0.4
101501	value=3.5	0.13	0.295	0.49	0.133
101501	value=4.0	0.717	0.477	0.075	0.066

ตารางที่ ค-5 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
105131	value=0.0	0.021	0.022	0.018	0.033
105131	value=1.0	0.021	0.045	0.169	0.366
105131	value=1.5	0.021	0.09	0.32	0.3
105131	value=2.0	0.065	0.227	0.283	0.133
105131	value=2.5	0.152	0.34	0.094	0.066
105131	value=3.0	0.304	0.181	0.075	0.033
105131	value=3.5	0.217	0.068	0.018	0.033
105131	value=4.0	0.195	0.022	0.018	0.033
202101	value=0.0	0.021	0.022	0.018	0.033
202101	value=1.0	0.021	0.022	0.018	0.033
202101	value=1.5	0.021	0.022	0.018	0.133
202101	value=2.0	0.021	0.022	0.056	0.366
202101	value=2.5	0.021	0.045	0.226	0.2
202101	value=3.0	0.065	0.204	0.301	0.066
202101	value=3.5	0.173	0.318	0.301	0.133
202101	value=4.0	0.652	0.34	0.056	0.033
202207	value=0.0	0.021	0.022	0.018	0.033
202207	value=1.0	0.021	0.022	0.018	0.033
202207	value=1.5	0.021	0.022	0.018	0.033
202207	value=2.0	0.021	0.045	0.094	0.433
202207	value=2.5	0.043	0.181	0.433	0.3
202207	value=3.0	0.282	0.386	0.301	0.1
202207	value=3.5	0.195	0.272	0.094	0.033
202207	value=4.0	0.391	0.045	0.018	0.033
202211	value=0.0	0.021	0.022	0.018	0.033
202211	value=1.0	0.021	0.022	0.018	0.1
202211	value=1.5	0.021	0.022	0.018	0.066
202211	value=2.0	0.021	0.022	0.037	0.266
202211	value=2.5	0.021	0.022	0.15	0.2
202211	value=3.0	0.065	0.181	0.433	0.266
202211	value=3.5	0.304	0.363	0.226	0.033
202211	value=4.0	0.521	0.34	0.094	0.033

ตารางที่ ค-5 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
202506	value=0.0	0.021	0.022	0.018	0.166
202506	value=1.0	0.021	0.068	0.132	0.133
202506	value=1.5	0.021	0.09	0.283	0.3
202506	value=2.0	0.021	0.204	0.245	0.233
202506	value=2.5	0.086	0.204	0.207	0.066
202506	value=3.0	0.326	0.272	0.075	0.033
202506	value=3.5	0.239	0.09	0.018	0.033
202506	value=4.0	0.26	0.045	0.018	0.033
202507	value=0.0	0.021	0.045	0.075	0.2
202507	value=1.0	0.021	0.068	0.301	0.3
202507	value=1.5	0.021	0.09	0.207	0.066
202507	value=2.0	0.043	0.25	0.188	0.266
202507	value=2.5	0.086	0.204	0.169	0.066
202507	value=3.0	0.173	0.181	0.018	0.033
202507	value=3.5	0.26	0.045	0.018	0.033
202507	value=4.0	0.369	0.113	0.018	0.033
901101	value=0.0	0.021	0.022	0.018	0.033
901101	value=1.0	0.021	0.022	0.018	0.066
901101	value=1.5	0.021	0.022	0.018	0.2
901101	value=2.0	0.021	0.045	0.113	0.133
901101	value=2.5	0.021	0.113	0.339	0.4
901101	value=3.0	0.195	0.181	0.358	0.1
901101	value=3.5	0.369	0.454	0.113	0.033
901101	value=4.0	0.326	0.136	0.018	0.033
903301	value=0.0	0.021	0.022	0.018	0.033
903301	value=1.0	0.021	0.022	0.018	0.066
903301	value=1.5	0.021	0.022	0.018	0.033
903301	value=2.0	0.021	0.022	0.018	0.1
903301	value=2.5	0.021	0.045	0.056	0.166
903301	value=3.0	0.043	0.181	0.358	0.3
903301	value=3.5	0.195	0.272	0.471	0.266
903301	value=4.0	0.652	0.409	0.037	0.033
grade_T1	value=VL	0.023	0.024	0.02	0.185
grade_T1	value=L	0.023	0.048	0.1	0.37

ตารางที่ ค-5 (ต่อ)

attribute	Parameter	High	Good	Middle	Low
grade_T1	value=M	0.023	0.073	0.54	0.296
grade_T1	value=H	0.116	0.56	0.26	0.111
grade_T1	value=VH	0.813	0.292	0.08	0.037
grade_T2	value=VL	0.023	0.024	0.02	0.037
grade_T2	value=L	0.023	0.024	0.04	0.111
grade_T2	value=M	0.023	0.024	0.16	0.629
grade_T2	value=H	0.023	0.439	0.76	0.185
grade_T2	value=VH	0.906	0.487	0.02	0.037
grade_T3	value=VL	0.023	0.024	0.02	0.074
grade_T3	value=L	0.023	0.024	0.08	0.481
grade_T3	value=M	0.023	0.097	0.62	0.333
grade_T3	value=H	0.209	0.731	0.26	0.074
grade_T3	value=VH	0.72	0.121	0.02	0.037
grade_T4	value=VL	0.023	0.024	0.02	0.037
grade_T4	value=L	0.023	0.024	0.02	0.296
grade_T4	value=M	0.023	0.073	0.48	0.555
grade_T4	value=H	0.116	0.56	0.44	0.074
grade_T4	value=VH	0.813	0.317	0.04	0.037
grade_T5	value=VL	0.023	0.024	0.02	0.037
grade_T5	value=L	0.023	0.024	0.02	0.185
grade_T5	value=M	0.023	0.024	0.3	0.629
grade_T5	value=H	0.139	0.463	0.58	0.111
grade_T5	value=VH	0.79	0.463	0.08	0.037

ภาคผนวก ง

บทความวิจัยที่ได้รับการตอบรับตีพิมพ์ในวารสารวิชาการระดับนานาชาติ

ISSN 2332-3205

Universal Journal of Education Research

http://www.hrpub.org

Horizon Research Publishing,USA

Horizon Research Publishing Corporation, USA

Date: 09/15/2021

Universal Journal of Educational Research

ISSN: 2332-3205 (Print) ISSN: 2332-3213 (Online)

Acceptance Letter

Dear Budsaba Sungwanna,

Congratulations! As a result of the reviews and revisions, we are pleased to inform you that your following paper has been accepted for publication.

Paper Title: A Modeling of an Intelligent System for Learning Result Prediction to Reduce Drop-out of Undergraduate Students

Paper ID: 19524568

Contributor (s): Budsaba Sungwanna, Pallop Piriyasurawong

It is scheduled for publication on Universal Journal of Educational Research, Vol 9 , No 10 .

The publication fee $ 380 should be paid within 2 weeks.

Should you have any questions, please feel free to let us know by quoting your **Paper ID** in any future inquiries.

Best wishes,

editorialboard@hrpub.org
Journal Manager
Horizon Research Publishing, USA
http://www.hrpub.org

ประวัติผู้วิจัย

ชื่อ : นางสาวบุษบา สังข์วรรณะ

ชื่อวิทยานิพนธ์ : ระบบพยากรณ์ผลการเรียนอัจฉริยะเพื่อลดการออกกลางคันของนักศึกษาระดับปริญญาตรี

สาขาวิชา : เทคโนโลยีสารสนเทศและการสื่อสารเพื่อการศึกษา

ประวัติการศึกษา

พ.ศ. 2546 สำเร็จการศึกษาหลักสูตรครุศาสตร์บัณฑิต สาขาวิชาคอมพิวเตอร์ศึกษา สถาบันราชภัฏหมู่บ้านจอมบึง

พ.ศ. 2549 สำเร็จการศึกษาหลักสูตรครุศาสตร์อุตสาหกรรมมหาบัณฑิต สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

ประวัติการทำงาน

พ.ศ. 2546-2549 อาจารย์สาขาวิชาคอมพิวเตอร์ มหาวิทยาลัยราชภัฏกาญจนบุรี

พ.ศ. 2549-2552 บริษัท โชคนำชัยกรุ๊ป จังหวัดสุพรรณบุรี

พ.ศ. 2552-2553 โรงเรียนบ้านมะนาวหวาน จังหวัดลพบุรี

พ.ศ. 2553-ปัจจุบัน อาจารย์สาขาวิชาเทคโนโลยีการศึกษาและคอมพิวเตอร์ มหาวิทยาลัยราชภัฏกาญจนบุรี

ผลงานวิชาการ

บุษบา สังข์วรรณะ และกาญจนา เทพสร. (2563). “การพัฒนาระบบบันทึกการเข้าร่วมกิจกรรมด้วย AR Code.” **วารสารวิทยาศาสตร์และเทคโนโลยี มทร.สุวรรณภูมิ.** ปีที่ 4 ฉบับที่ 1 : 33-42.

บุษบา สังข์วรรณะ เพียงจันทร์ โมฟเฟ็ทท์ และพาทีเกศธนากร. (2562). “การพัฒนาบรรจุภัณฑ์ขนมทองโยะเพื่อการถนอมอาหาร.” ใน **การประชุมวิชาการระดับชาติ “ราชภัฏกรุงเก่า” ประจำปี พ.ศ. 2562.** มหาวิทยาลัยราชภัฏพระนครศรีอยุธยา, (569-576).

บุษบา สังข์วรรณะ. (2559). “การยกระดับข้อมูลด้วยระบบการจัดการเนื้อหา.” **วารสารการอาชีวะและเทคนิคศึกษา.** ปีที่ 6 ฉบับที่ 11 : 60-63.

ผลงานวิจัย

บุษบา สังข์วรรณะ และกาญจนา เทพสร. การพัฒนาระบบบันทึกการเข้าร่วมกิจกรรม AR Code (หัวหน้าโครงการ) พ.ศ. 2561.

พาที เกศธนากร, บุษบา สังข์วรรณะ และเพียงจันทร์ โมฟเฟ็ทท์. **รูปแบบการพัฒนาเศรษฐกิจชุมชนที่ส่งผลต่อการพึ่งตนเองอย่างยั่งยืนของชุมชนอำเภอทองผาภูมิ จังหวัดกาญจนบุรี** (ผู้วิจัย) พ.ศ. 2561.

บุษบา สังข์วรรณะ, เพียงจันทร์ โมฟเฟ็ทท์ และพาที เกศธนากร. **การพัฒนาบรรจุภัณฑ์ขนมทองโยะเพื่อการถนอมอาหาร ตำบลท่าขนุน จังหวัดกาญจนบุรี** (หัวหน้าโครงการ) พ.ศ. 2561.