



มหาวิทยาลัยศรีปทุม

รายงานวิจัย

เรื่อง

การสร้างโมเดลการจัดการระบบนักศึกษาสัมพันธ์โดยใช้เทคนิคเหมืองข้อมูล

THE DEVELOPMENT OF STUDENT RELATIONSHIP MANAGEMENT MODEL  
USING DATA MINING TECHNIQUE

มหาวิทยาลัยศรีปทุม  
SRIPATUM UNIVERSITY

ปราณี มณีรัตน์

งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยศรีปทุม

ปีการศึกษา 2554

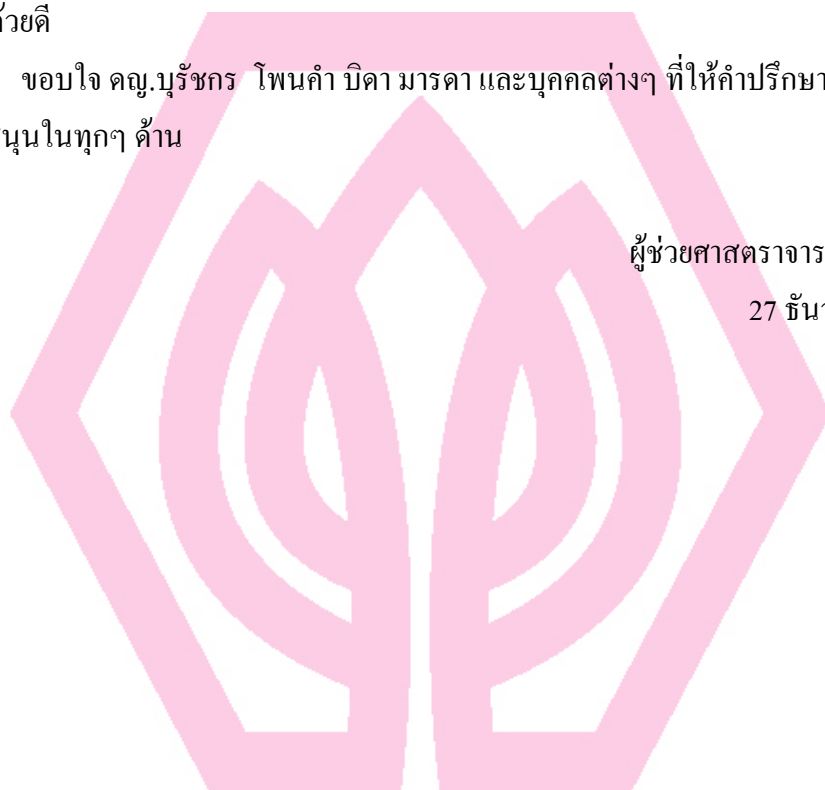
## กิตติกรรมประกาศ

งานวิจัยฉบับนี้ได้รับความอนุเคราะห์จาก ผศ.ดร.กฤษณะ ชินสาร และรศ.ดร.ประสงค์ ประณีตพลกรัง ที่ได้กรุณาสละเวลาอันมีค่ายิ่งในการคำแนะนำ และปรึกษา งานวิจัยนี้จนสำเร็จ  
ลุล่วงด้วยดี

ขอขอบใจ คณ.บุรุษกร โพนคำ บิดา มารดา และบุคคลต่างๆ ที่ให้คำปรึกษาและกำลังใจ  
สนับสนุนในทุกๆ ด้าน

ผู้ช่วยศาสตราจารย์ปราณี มณีรัตน์

27 ธันวาคม 2554



มหาวิทยาลัยศรีปทุม  
SRIPATUM UNIVERSITY

หัวข้อวิจัย : การสร้างโมเดลการจัดการระบบนักศึกษาสัมพันธ์โดยใช้เทคนิคเหมืองข้อมูล  
ผู้วิจัย : นางสาวปราณี มณีรัตน์  
หน่วยงาน : คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม  
ปีที่พิมพ์ : พ.ศ.2554

---

## บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและวิเคราะห์หาปัจจัยที่ผลต่อความพึงพอใจของนักศึกษา ในการสร้างระบบบริหารความสัมพันธ์นักศึกษา (Student Relationship Management) โดยใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) ซึ่งใช้การแบ่งกลุ่มข้อมูล (Clustering) โดยอัลกอริทึม K Mean และเทคนิค JRip การหาความสัมพันธ์ของระหว่างปัจจัยร่วมกัน ในงานวิจัยครั้งนี้ได้เก็บแบบสอบถามจำนวน 2,731 ชุด จากนักศึกษามหาวิทยาลัยรัฐและเอกชน จำนวน 7 แห่ง โดยการแจกแบบสอบถามแบบบังเอิญ (Accidental Sampling) ซึ่งลักษณะแบบสอบถาม จะสอบถามความพึงพอใจของนักศึกษาในด้านต่างๆ ดังนี้คือ ภาพรวมเนื้อหาหลักสูตร คุณสมบัติอาจารย์ และสิ่งอำนวยความสะดวกของแต่ละมหาวิทยาลัย

ในงานวิจัยครั้งนี้ผลลัพธ์ที่ได้จากการใช้การแบ่งกลุ่มข้อมูล (Clustering) โดยใช้เทคนิค K mean แบ่งกลุ่มนักศึกษาตามชั้นที่ 1,2,3 และ 4 ออกเป็นกลุ่มต่างๆ เป็นวิธีที่ช่วยให้สามารถวิเคราะห์คุณลักษณะและความพึงพอใจในด้านต่างๆ ของนักศึกษาแต่ละกลุ่ม ได้อย่างถูกต้องและชัดเจน โดยที่ชั้นปีที่ 1 มีค่าเฉลี่ย( $\bar{X}$ ) ความพึงใจรวมในด้านต่างๆ เท่ากับ 3.77 ชั้นปีที่ 2 มีค่าเฉลี่ย( $\bar{X}$ ) ความพึงใจรวมในด้านต่างๆ เท่ากับ 3.64 ชั้นปีที่ 3 มีค่าเฉลี่ย( $\bar{X}$ ) ความพึงใจรวมในด้านต่างๆ เท่ากับ 3.67 และชั้นปีที่ 4 มีค่าเฉลี่ย( $\bar{X}$ ) ความพึงใจรวมในด้านต่างๆ เท่ากับ 3.77 และใช้เทคนิค JRip ช่วยในการค้นหาความสัมพันธ์ในแต่ละกลุ่มที่สนใจศึกษา เพื่อวิเคราะห์ในรายละเอียดว่ากลุ่มนั้นมีความสัมพันธ์ระหว่างปัจจัยใดร่วมกัน เพื่อจะนำมาปรับปรุงการสร้าง ความพึงพอใจให้กับนักศึกษาให้ตรงตามความต้องการของกลุ่มนักศึกษาแต่ละกลุ่มให้มากที่สุด

คำสำคัญ : ระบบบริหารความสัมพันธ์นักศึกษา,การทำเหมืองข้อมูล, แบ่งกลุ่มข้อมูล ,  
K Mean, JRip

Research Title : The Development of Student Relationship Management Model using  
Data Mining Technique

Name of Researcher : Miss Paralee Maneerat

Name of Institution : Faculty of Information Technology, Sripatum University

Year of Publication : B. E. 2554

---

### ABSTRACT

This research is to study and analyze the factors that affecting the students satisfaction in Student Relationship Management (SRM) model using data mining technique .Groups of data were separated by clustering by K mean algorithm. Consequently, the individual data can be analyzed by JRip technique, which to find out the relationship to others factor. This research is based on quantitative study of a sample of undergraduate student of seven universities both public and private university in Thailand. Data were collected through individuals questionnaires from an accidental sampling of 2,731 to examine the satisfactions of students. The questionnaire composed of the overview of curriculum, the proprieties lecturer and the facilities in university.

The results of segmentation were separated by the levels of study year .Thereby ,the analysis provides more correctly and effective evaluation of students satisfaction factor .For example the average of the satisfaction ( $\bar{x}$ ) in the first year is 3.77, the average of the satisfaction ( $\bar{x}$ ) in the second year is 3.65, the average of the satisfaction ( $\bar{x}$ ) in the third year is 3.67 ,and the average of the satisfaction ( $\bar{x}$ ) in the fourth year is 3.77. Moreover, this research regard the important relationship of the others factors in lower satisfaction student segment . The relationships help the university improve the service to the best for theirs satisfaction of students.

Keywords : SRM, Data mining, Clustering , K Mean ,JRip

## สารบัญ

บทที่		หน้า
1	บทนำ.....	1
	1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
	1.2 วัตถุประสงค์ของการวิจัย.....	2
	1.3 คำถามการวิจัย.....	3
	1.4 สมมติฐานการวิจัย.....	3
	1.5 ขอบเขตของการวิจัย.....	3
	1.6 นิยามศัพท์เฉพาะ.....	3
	1.7 ประโยชน์ที่คาดว่าจะได้รับ.....	5
2	วรรณกรรมที่เกี่ยวข้อง.....	6
	2.1 การทำเหมืองข้อมูล.....	6
	2.2 การสุ่มตัวอย่าง.....	24
	2.3 ประเภทของสถิติ.....	27
	2.4 ข้อมูล.....	27
	2.5 ค่าเฉลี่ย.....	29
	2.6 ระบบบริหารความสัมพันธ์ลูกค้า.....	29
	2.7 งานวิจัยที่เกี่ยวข้อง.....	33
3	ระเบียบวิธีการวิจัย.....	36
	3.1 วิธีการและขั้นตอนการดำเนินการวิจัย.....	36
	3.2 การคัดเลือกประชากรและกลุ่มตัวอย่าง.....	37
	3.3 ตัวแปรที่ใช้ในงานวิจัย.....	38
	3.4 เครื่องมือและการทดสอบเครื่องมือที่นำมาใช้ในงานวิจัย.....	38
	3.5 วิธีการเก็บรวบรวมข้อมูลสำหรับทำงานวิจัย.....	41
	3.6 หลักสถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	42
	3.7 วิธีการวิจัยโดยใช้เทคนิคเหมืองข้อมูล.....	42

## สารบัญ (ต่อ)

บทที่		หน้า
4	ผลการวิเคราะห์ข้อมูล.....	47
	4.1 การเก็บรวบรวมข้อมูล.....	47
	4.2 การเตรียมข้อมูล.....	48
	4.3 การแปลงข้อมูล.....	59
	4.4 การแบ่งกลุ่มข้อมูล.....	64
	4.5 การแปลผล.....	72
5	สรุป อภิปรายผล และข้อเสนอแนะ.....	85
	5.1 สรุปผลงานวิจัย.....	85
	5.2 อภิปรายผล.....	85
	5.3 ปัญหาที่เกิดขึ้นระหว่างการดำเนินงานวิจัย.....	91
	5.4 ข้อเสนอแนะ.....	91
	บรรณานุกรม.....	92
	ภาคผนวก ก.....	94
	ภาคผนวก ข.....	100

## สารบัญตาราง

ตารางที่		หน้า
3.1	แสดงระดับความคิดเห็น คะแนนเฉลี่ยและระดับการแปลความหมาย.....	39
3.2	แสดงค่าความเชื่อมั่นด้วยวิธีของ Cronbach's Alpha.....	41
3.3	แสดงการจํานวนนักศึกษาที่แบ่งกลุ่มในแต่ละชั้นปี.....	42
4.1	หัวข้อต่างๆ ในแบบสอบถามความพึงพอใจของนักศึกษา.....	58



มหาวิทยาลัยศรีปทุม  
SRIPATUM UNIVERSITY

# สารบัญภาพประกอบ

ภาพประกอบที่		หน้า
2.1	การแสดงผลข้อมูลสู่การตัดสินใจและปฏิบัติ .....	6
2.2	กระบวนการค้นพบความรู้จากฐานข้อมูล .....	8
2.3	การแสดงผลขบวนการตัดสินใจ .....	9
2.4	กระบวนการ Classification.....	12
2.5	โครงสร้างต้นไม้ตัดสินใจ .....	13
2.6	โครงสร้างเครือข่ายประสาท.....	13
2.7	นิเวศน์เน็ต เพื่อการวิเคราะห์การเช่าและซื้อบ้าน .....	16
2.8	ตัวอย่าง Clustering.....	18
2.9	แสดงขั้นตอนการจัดกลุ่มข้อมูลด้วยเทคนิค Clustering .....	19
3.1	แสดงขั้นตอนดำเนินการวิจัย.....	37
3.2	แสดงขั้นตอนวิจัยโดยใช้เทคนิคเหมืองข้อมูล .....	43
3.3	ค่าความแตกต่างระหว่างกลุ่ม (BCV) ควรมีค่ามาก .....	
	และค่าความแตกต่างภายในกลุ่ม (WCV) ควรมีค่าน้อย .....	45
4.1	แสดงค่า F-Test และ T-test เปรียบเทียบของการแบ่งกลุ่มนักศึกษาในชั้นปีที่ 1.....	64
4.2	แสดงการแบ่งกลุ่มจำนวน 4 กลุ่ม ของนักศึกษาชั้นปีที่ 1.....	65
4.3	แสดงค่า F-Test และ T-test เปรียบเทียบของการแบ่งนักศึกษาในชั้นปีที่ 2 .....	66
4.4	แสดงการแบ่งกลุ่มจำนวน 6 กลุ่ม ของนักศึกษาชั้นปีที่ 2.....	67
4.5	แสดงค่า F-Test และ T-test เปรียบเทียบของการแบ่งนักศึกษาในชั้นปีที่ 3 .....	68
4.6	แสดงการแบ่งกลุ่มจำนวน 5 กลุ่ม ของนักศึกษาชั้นปีที่ 3.....	69
4.7	แสดงค่า F-Test และ T-test เปรียบเทียบของการแบ่งนักศึกษาในชั้นปีที่ 4 .....	70
4.8	แสดงการแบ่งกลุ่มจำนวน 4 กลุ่ม ของนักศึกษาชั้นปีที่ 4.....	71
4.9	แสดงกราฟการแบ่งกลุ่มของนักศึกษาชั้นปีที่ 1.....	72
4.10	แสดงกราฟการแบ่งกลุ่มของนักศึกษาชั้นปีที่ 2.....	75
4.11	แสดงกราฟการแบ่งกลุ่มของนักศึกษาชั้นปีที่ 3.....	79
4.12	แสดงกราฟการแบ่งกลุ่มของนักศึกษาชั้นปีที่ 4.....	82
5.1	แสดงกราฟค่าเฉลี่ยความพึงพอใจของนักศึกษาแต่ละชั้นปี .....	86



# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันการให้บริการด้านการศึกษาของสถาบันการศึกษา มีการแข่งขันกันสูงมากขึ้น ทั้งในแง่ของการปรับปรุงคุณภาพการศึกษาให้สอดคล้องกับความต้องการของผู้เข้ารับการศึกษ และมุ่งผลิตบัณฑิตให้ตรงกับความต้องการของตลาดแรงงาน อีกทั้งตามนโยบายของภาครัฐบาลว่าพัฒนากรรมการศึกษาของประเทศได้เปลี่ยนแปลงไปในทุกระดับของการศึกษา ไม่ว่าจะเป็นระดับประถมศึกษา ระดับมัธยมศึกษา และระดับอุดมศึกษา จะเน้นเรื่องความต้องการของผู้ศึกษาเป็นหลักสำคัญ

สถาบันการศึกษาต่างๆ จึงต้องพัฒนาเพิ่มขีดความสามารถในทุกๆ ด้านของสถาบันการศึกษาให้สอดคล้อง ทันต่อการเปลี่ยนแปลงทั้งในปัจจุบันและอนาคต เช่น การให้บริการแก่ผู้เข้ารับการศึกษด้วยเทคโนโลยีสมัยใหม่ หรือ การบริหารจัดการข้อมูลของนักศึกษาให้เกิดประสิทธิภาพสูงสุด ในแง่มุมมองของการบริหารงานของสถาบันการศึกษา มีการนำแนวคิดต่างๆ เพื่อพัฒนาและบริหาร งานด้วยเทคโนโลยีสารสนเทศกันอย่างต่อเนื่อง และหนึ่งในนั้น คือการนำเอาแนวคิด ระบบบริหารความสัมพันธ์ลูกค้า(CRM) มาประยุกต์ใช้ร่วมในการบริหารงานด้านการศึกษา โดยระบบบริหารความสัมพันธ์ลูกค้า เป็นการสร้างความพึงพอใจของลูกค้าในธุรกิจนั้นๆ เพื่อรักษาลูกค้าเดิม และขยายฐานลูกค้าใหม่ให้กับธุรกิจ การสร้างความพึงพอใจนี้ ถือเป็นตรรกะความคิดที่สามารถนำมาถ่ายทอด และดัดแปลงให้เหมาะสมกับสถาบันการศึกษาได้ หากมองไปที่มุมมองของสถาบันการศึกษาแล้วความพึงพอใจของนักศึกษา ก็เปรียบเสมือนความพึงพอใจของลูกค้าเช่นกัน แม้ว่าในมุมมองของการดำเนินงานทางธุรกิจจะมุ่งเน้นให้ได้ผลกำไรสูงสุดจากการขายผลิตภัณฑ์ก็ตาม แต่การนำเอาแนวคิดบริหารลูกค้าสัมพันธ์ ถูกดัดแปลงนำมาใช้กับรูปแบบการบริหารงานของสถาบันการศึกษา ซึ่งการให้บริการที่สำคัญอย่างยิ่งของสถาบันการศึกษา คือการเพิ่มขีดความสามารถในการบริการการศึกษาให้เกิดประสิทธิภาพสูงสุด และตอบสนองความต้องการของผู้เข้ารับการศึกษ เพื่อเกิดความพึงพอใจในสถาบันการศึกษาให้มากที่สุด แนวคิดเช่นนี้ ถูกนิยามทางทฤษฎีว่า ระบบบริหารความสัมพันธ์นักศึกษา (Student Relationship Management: SRM) ในต่างประเทศมีผลงานในรูปแบบการออกแบบการทำงานวิจัยเกี่ยวกับทฤษฎีและการประยุกต์ ใช้งานระบบนักศึกษาสัมพันธ์ เพื่อใช้งานในแต่ละสถาบันการศึกษาซึ่งเป็นที่ชัดเจนว่าวัฒนธรรมการศึกษา ของแต่ละประเทศย่อมแตกต่างกัน ทั้งแนวความคิดต่อการเข้ารับการศึกษและแนวความคิดต่อการให้การศึกษา สำหรับในประเทศไทย

การออกแบบระบบสารสนเทศของสถาบันการศึกษา ส่วนใหญ่เน้นในเรื่องประโยชน์ของการใช้งานระบบเป็นหลัก ซึ่งก็คือเจ้าหน้าที่ของสถาบันการศึกษา มากกว่าจะเน้นไปที่ความพึงพอใจต่อผู้เข้ารับการศึกษา โดยมีการสร้างระบบสารสนเทศสำหรับส่วนงานย่อยๆ เพื่อให้ส่วนงานย่อยๆ นั้นสามารถทำงานได้ เช่น ส่วนงานบริการห้องสมุด ส่วนงานระบบลงทะเบียน ส่วนงานระบบรายงานผลการศึกษา แต่สำหรับระบบบริหารความสัมพันธ์นักศึกษาแล้ว รูปแบบการพิจารณา จะเริ่มจากความต้องการของนักศึกษาเป็นหลัก ซึ่งอาจจะเป็นผู้ใช้ระบบโดยตรงหรืออาจจะไม่ใช่ผู้ใช้ระบบก็ได้

จากหลักการเบื้องต้น ทำให้เกิดแนวคิดในการทำวิจัยเรื่องการสร้างโมเดลระบบบริหารความสัมพันธ์นักศึกษาโดยใช้เทคนิคเหมืองข้อมูล ซึ่งจะใช้ข้อมูลจากนักศึกษา การจัดการเรียนการสอนของหลักสูตร ข้อมูลการสอนของอาจารย์ สิ่งอำนวยความสะดวกต่างๆ ในสถาบันการศึกษาต่างๆ มาวิเคราะห์ข้อมูลโดยใช้เทคนิคเหมืองข้อมูล โดยในสถาบันการศึกษามีการเก็บข้อมูลเกี่ยวกับนักศึกษาเป็นจำนวนมาก โดยข้อมูลเหล่านี้สามารถนำมาใช้ประโยชน์ได้เป็นอย่างดี แต่ส่วนใหญ่ไม่ได้นำข้อมูลเหล่านั้นมาใช้งาน การทำเหมืองข้อมูล (Data Mining) จึงเป็นวิธีการหนึ่งที่ใช้กันอย่างแพร่หลาย เนื่องจากมีความสามารถในการจำแนกประเภท (Classification) และการแบ่งกลุ่ม (Clustering) รวมถึงการสร้างแบบจำลอง (Model) ต่างๆ มากมาย การทำเหมืองข้อมูล จะวิเคราะห์ข้อมูลที่มีขนาดใหญ่ เพื่อค้นหาความสัมพันธ์และรูปแบบ หรือกฎที่ซ่อนอยู่ และนำความสัมพันธ์ต่างๆ เหล่านั้น นำมาใช้ให้เกิดประโยชน์ อีกทั้งงานวิจัยนี้จะสร้างแบบจำลองระบบบริหารความสัมพันธ์นักศึกษาที่จะนำไปใช้กับสถาบันการศึกษาต่างๆ ในประเทศไทย เพื่อนำไปบริหารจัดการกับสถาบันการศึกษา ในการสร้างความพึงพอใจ (Satisfaction) ให้นักศึกษาต่อสถาบันนั้นๆ ให้เกิดประโยชน์สูงสุด

## 1.2 วัตถุประสงค์ของการวิจัย

ในการออกแบบแบบจำลองระบบบริหารความสัมพันธ์นักศึกษา มีการใช้เทคนิคเหมืองข้อมูล โดยใช้ Clustering ช่วยในการวิเคราะห์ข้อมูล โดยมีวัตถุประสงค์หลักดังนี้คือ

1. เพื่อศึกษาการจัดการระบบบริหารความสัมพันธ์ของนักศึกษา
2. เพื่อสร้างแบบจำลองระบบบริหารความสัมพันธ์ของนักศึกษาโดยใช้เทคนิคเหมืองข้อมูล

### 1.3 คำถามการวิจัย

1. ระบบบริหารความสัมพันธ์นักศึกษา มีลักษณะอย่างไร
2. แบบจำลองระบบบริหารความสัมพันธ์นักศึกษาโดยใช้เทคนิคเหมืองข้อมูลสามารถทำให้การจัดการความสัมพันธ์กับนักศึกษามีประสิทธิภาพมากขึ้นอย่างไร

### 1.4 สมมติฐานการวิจัย

1. นักศึกษามีระดับการยอมรับและความพึงพอใจต่อระบบบริหารความสัมพันธ์นักศึกษาอยู่ในระดับมาก
2. แบบจำลองระบบบริหารความสัมพันธ์นักศึกษาโดยใช้เทคนิคเหมืองข้อมูลที่พัฒนาขึ้น ทำให้การจัดการความสัมพันธ์กับนักศึกษามีประสิทธิภาพ อยู่ในระดับมาก

### 1.5 ขอบเขตของการวิจัย

1. ศึกษาการใช้เทคนิคต่างๆ ของเหมืองข้อมูลเพื่อประยุกต์ใช้งานบนระบบบริหารความสัมพันธ์นักศึกษา
2. ศึกษาข้อมูลนักศึกษา โดยใช้ข้อมูลจากกลุ่มตัวอย่างของนักศึกษาจากมหาวิทยาลัยรัฐบาล และเอกชน จำนวน 7 แห่ง โดยแบ่งเป็นมหาวิทยาลัยภาครัฐจำนวน 3 แห่งและมหาวิทยาลัยเอกชนจำนวน 4 แห่ง
3. ศึกษาความต้องการใช้ข้อมูลในระบบบริหารความสัมพันธ์นักศึกษา และทำการแบ่งแยก ระดับของข้อมูลให้ชัดเจน
4. วิเคราะห์เพื่อหาปัจจัยที่มีต่อการสร้างโมเดลระบบบริหารความสัมพันธ์นักศึกษา

### 1.6 นิยามศัพท์เฉพาะ

1. คลังข้อมูล (Data Warehouse) คือฐานข้อมูลขนาดใหญ่ ที่รวบรวมฐานข้อมูลจากหลายแหล่ง หลายช่วงเวลา ซึ่งอาจมีสกรีมมา (Schema) แตกต่างกัน มาไว้รวม ณ ที่เดียวกัน (และใช้ schema เดียวกัน)
2. การทำเหมืองข้อมูล (Data Mining) หรืออาจจะเรียกว่า การค้นหาความรู้ในฐานข้อมูล (Knowledge Discovery in Databases - KDD) เป็นเทคนิคเพื่อค้นหารูปแบบ (Pattern) จากข้อมูลจำนวนมากโดยอัตโนมัติ โดยใช้ขั้นตอนวิธีจากวิชาสถิติ การเรียนรู้ของเครื่อง และ การรู้จำแบบ หรือในอีกนิยามหนึ่ง การทำเหมืองข้อมูล คือ กระบวนการที่กระทำกับ

ข้อมูล เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัย หลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง และหลักคณิตศาสตร์

3. ระบบบริหารความสัมพันธ์ลูกค้า (Customer Relationship Management: CRM) เป็นกลยุทธ์ทางธุรกิจเพื่อสร้างความสัมพันธ์ระยะยาว กับลูกค้า เรียนรู้ความต้องการที่แตกต่างกันของลูกค้า และตอบสนองความต้องการของลูกค้าด้วยสินค้า หรือบริการที่เหมาะสมกับลูกค้าแต่ละคนให้มากที่สุด ความสำคัญของการบริหารงานลูกค้าสัมพันธ์ (CRM) ช่วยให้องค์กรสามารถเพิ่มความสัมพันธ์อันดีให้กับลูกค้า เพิ่มรายได้ลดค่าใช้จ่ายต่างๆ โดยเฉพาะเรื่องค่าใช้จ่ายในการแสวงหาลูกค้าและ เพิ่มความพึงพอใจของลูกค้า ( Customer Satisfaction ) โดยการสร้างกระบวนการทำงาน และพัฒนาผลิตภัณฑ์ตามความต้องการของลูกค้าจำนวนผู้ประกอบการ ที่เพิ่มขึ้นอย่างรวดเร็วในปัจจุบัน สามารถนำแนวทางการบริหารลูกค้าสัมพันธ์ เช่น แนวทางที่สำคัญในการจัดการระบบการบริหารงานและสร้างมาตรฐานการทำงานในบริษัท เช่น การรวบรวมเกี่ยวกับข้อมูลของลูกค้า การจัดการเกี่ยวกับช่องทางการสื่อสารและการพัฒนาสินค้า และบริการเพื่อสนองตอบสิ่งที่ลูกค้าต้องการ
4. ระบบบริหารความสัมพันธ์นักศึกษา (Student Relationship Management :SRM) โดยอาศัยหลักการและแนวคิดจากรูปแบบระบบบริหารความสัมพันธ์ลูกค้า แต่ในระบบบริหารความสัมพันธ์นักศึกษานั้น กล่าวคือสถาบันการศึกษาให้ความสำคัญกับการรักษาข้อมูลต่างๆ ทางด้านการศึกษา ไม่ว่าจะเป็นทะเบียนประวัติ ทะเบียนผลการศึกษา การทำกิจกรรมต่างๆ การสร้างความสัมพันธ์ระหว่างสถาบันกับนักศึกษา ไม่ว่าจะเป็นอยู่ในระหว่างการศึกษา หรือหลังจบการศึกษา ระบบบริหารความสัมพันธ์นักศึกษา ยังมีรูปแบบที่ไม่ชัดเจนนักในปัจจุบัน ในต่างประเทศมีการศึกษาและทำวิจัยในเรื่องนี้ ไม่ว่าจะเป็นที่อเมริกา หรือเยอรมัน สำหรับประเทศไทย เนื่องจากการดำเนินงานด้านการศึกษาของประเทศไทย มีทั้งส่วนเหมือนและส่วนที่แตกต่าง ไม่ว่าจะเป็นการจัดการ การควบคุม ความสัมพันธ์ระหว่างผู้สอนและลูกศิษย์ ซึ่งต้องอาศัยการทำงานวิจัยเพิ่มเติมว่าจะทำอะไรเพื่อให้ระบบบริหารความสัมพันธ์นักศึกษาเกิดขึ้นได้จริงจัง และชัดเจนในประเทศไทย
5. ความพึงพอใจ (Satisfaction) หมายถึง ความรู้สึกส่วนบุคคลในด้านความพอใจ ความยินดี หรือความผิดหวัง ซึ่งเป็นผลมาจากการเปรียบเทียบระหว่างการรับรู้เข้าใจถึงสมรรถนะ หรือผลลัพธ์ที่แสดงออกมาทางสินค้ากับความคาดหวังในตัวสินค้า

### 1.7 ประโยชน์ที่คาดว่าจะได้รับ

1. เกิดโมเดลระบบการจัดการนักศึกษาสัมพันธ์
2. สามารถจัดการข้อมูลนักศึกษาได้อย่างมีประสิทธิภาพเพิ่มมากขึ้น
3. สามารถนำโมเดลระบบบริหารความสัมพันธ์นักศึกษา ไปประยุกต์ใช้สถาบันการศึกษาอื่นๆ ได้



มหาวิทยาลัยศรีปทุม  
SRIPATUM UNIVERSITY

## บทที่ 2

### วรรณกรรมที่เกี่ยวข้อง

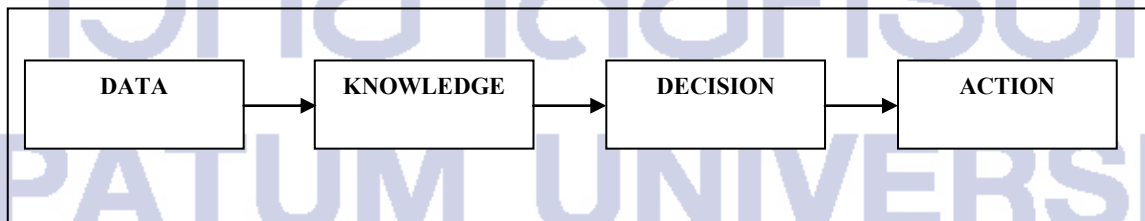
การสร้างโมเดลระบบบริหารความสัมพันธ์นักศึกษาโดยใช้เทคนิคเหมืองข้อมูล โดยนำข้อมูลต่างๆ มาใช้ในการบริหารความสัมพันธ์ของนักศึกษากับมหาวิทยาลัย เพื่อสร้างความพึงพอใจให้กับนักศึกษาในการเข้ารับการศึกษาในสถาบันการศึกษาต่างๆ จึงจำเป็นอย่างยิ่งที่ต้องศึกษาหาข้อมูลในส่วนของทฤษฎีและผลงานวิจัยต่างๆ ที่เกี่ยวข้อง ซึ่งมีรายละเอียดดังนี้คือ

#### 2.1 การทำเหมืองข้อมูล (Data Mining)

##### 2.1.1. ความหมายของการทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) หรืออาจเรียกว่าการค้นหาคำรู้ในฐานข้อมูล (Knowledge Discovery in Database - KDD) เป็นเทคนิคเพื่อค้นหารูปแบบ ของจากข้อมูลจำนวนมาก มาหาโดยอัตโนมัติ โดยใช้ขั้นตอนวิธีจากวิชาสถิติ การเรียนรู้ของเครื่องและการรู้จำแบบ หรือในอีกนิยามหนึ่งของการทำเหมืองข้อมูล คือ กระบวนการที่กระทำกับข้อมูล (โดยส่วนใหญ่จะมีจำนวนมาก) เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง และหลักคณิตศาสตร์

การทำเหมืองข้อมูล (Data Mining) คือ การค้นหาคำสัมพันธ์ และรูปแบบทั้งหมด ซึ่งมีอยู่จริงในฐานข้อมูล แต่ได้ถูกซ่อนไว้ภายในข้อมูลจำนวนมาก การทำเหมืองข้อมูลจะทำการสำรวจ และวิเคราะห์อย่างอัตโนมัติ ในปริมาณข้อมูลจำนวนมากให้อยู่ในรูปแบบที่เต็มไปด้วยความหมายและอยู่ในรูปของกฎ โดยความสัมพันธ์เหล่านี้แสดงให้เห็นถึงความรู้ต่างๆ ที่มีประโยชน์ในฐานข้อมูล



รูปที่ 2.1 การแสดงข้อมูลสู่การตัดสินใจและปฏิบัติ

ส่วนวิเคราะห์และทำนายนี้จะใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) ซึ่งเป็นกระบวนการในการขุดค้น วิเคราะห์ เพื่อที่จะค้นหารูปแบบหรือกฎเกณฑ์ที่เป็นประโยชน์ และนำเสนอใจออกมาจากข้อมูลที่มีจำนวนมากมายมหาวิทยาลัยโดยวิธีแบบอัตโนมัติ หรือ กึ่งอัตโนมัติ

**การทำเหมืองข้อมูล** จำแนกออกเป็นประเภทของรูปแบบได้ 2 รูปแบบ ดังนี้

(1) การทำเหมืองข้อมูลเพื่อหาหรือกำหนดลักษณะหรือคุณสมบัติของข้อมูลเป็นการวิเคราะห์เพื่อหาลักษณะหรือคุณสมบัติของกลุ่มข้อมูล การหาลักษณะของกลุ่มข้อมูลทั่วไปในฐานข้อมูลซึ่งจะหาคุณลักษณะ 3 แบบคือ หาคุณลักษณะที่ตรงกัน หาคุณลักษณะที่ต่างกัน และหาคุณลักษณะที่มีทั้งความต่างและเหมือนกัน

(2) การทำเหมืองข้อมูลเพื่อใช้ในการทำนายสิ่งที่จะเกิดขึ้น เป็นการวิเคราะห์หาโมเดลเพื่อทำนายแนวโน้มหรือคุณสมบัติของข้อมูลที่ยังไม่ปรากฏ เช่น การหายอดขายสินค้าของเดือนถัดไป

### **ประโยชน์ของการทำเหมืองข้อมูล**

ประโยชน์ของการทำเหมืองข้อมูล นั้นมีหลายอย่างเริ่มตั้งแต่การวิเคราะห์ผลิตภัณฑ์ การวิเคราะห์บัตรเครดิต การวิเคราะห์ลูกค้า การวิเคราะห์การขาย พาณิชย์อิเล็กทรอนิกส์ (E-Commerce)

### **กระบวนการค้นพบความรู้จากฐานข้อมูล**

กระบวนการค้นพบความรู้จากฐานข้อมูลขนาดใหญ่มาก เป็นกระบวนการสร้างแบบจำลองหรือรูปแบบกฎเกณฑ์จากกลุ่มของข้อมูล ทำให้เกิดความเข้าใจลักษณะรูปแบบความเกี่ยวข้องสัมพันธ์กันของกลุ่มข้อมูล แล้วแนวโน้มเพื่อใช้ในการทำนายข้อมูลนั้นๆ โดยมี กระบวนการตามทฤษฎี รวม 4 ขั้นตอนดังนี้

#### **(1) การเตรียมข้อมูล (Data preparation)**

ขั้นตอนการเตรียมข้อมูลเป็นขั้นตอนสำคัญ และใช้เวลามากที่สุด เนื่องจากบ่อยครั้งเลือกข้อมูลมาไม่เหมาะสม และไม่ถูกต้อง หรือการนำข้อมูลมาจากหลายแหล่งที่มาพร้อมเข้ากันเพื่อพิจารณาความสัมพันธ์ของข้อมูล ซึ่งส่งผลให้เกิดความผิดพลาด ดังนั้นขั้นตอนการเตรียมข้อมูลจึงถือเป็นหัวใจของงาน การเตรียมข้อมูลสามารถแบ่งออกได้เป็น 3 ขั้นตอนย่อย คือ

##### **- การคัดเลือกข้อมูล (Data Selection)**

จุดประสงค์หลัก คือ การระบุลักษณะข้อมูลที่ต้องการแล้ว ทำการคัดเลือกข้อมูลทำการคัดเลือกข้อมูลที่ต้องการ ซึ่งข้อมูลที่ได้จะแตกต่างกันไปตามจุดประสงค์ ของแต่ละธุรกิจ

##### **- การกลั่นกรองข้อมูล (Data Cleaning)**

จุดประสงค์ เพื่อมั่นใจว่าคุณภาพของข้อมูลที่ถูกเลือกนั้นถูกต้องและเหมาะสม เนื่องจากข้อมูลที่ถูกเลือกมาจากกระบวนการเลือกข้อมูลนั้นอาจมีข้อมูลไม่ถูกต้อง

##### **- การแปลงรูปข้อมูล (Data Transformation)**

จุดประสงค์เพื่อแปลงข้อมูลให้อยู่ในรูปแบบที่พร้อมจะนำไปวิเคราะห์ตามหลักอัลกอริทึมของการทำเหมืองข้อมูลที่ใช้ เช่น การแบ่งช่วงอายุให้เป็นกลุ่มๆ หรือ กำหนดตัวเลขให้กับแต่ละกลุ่มของข้อมูล

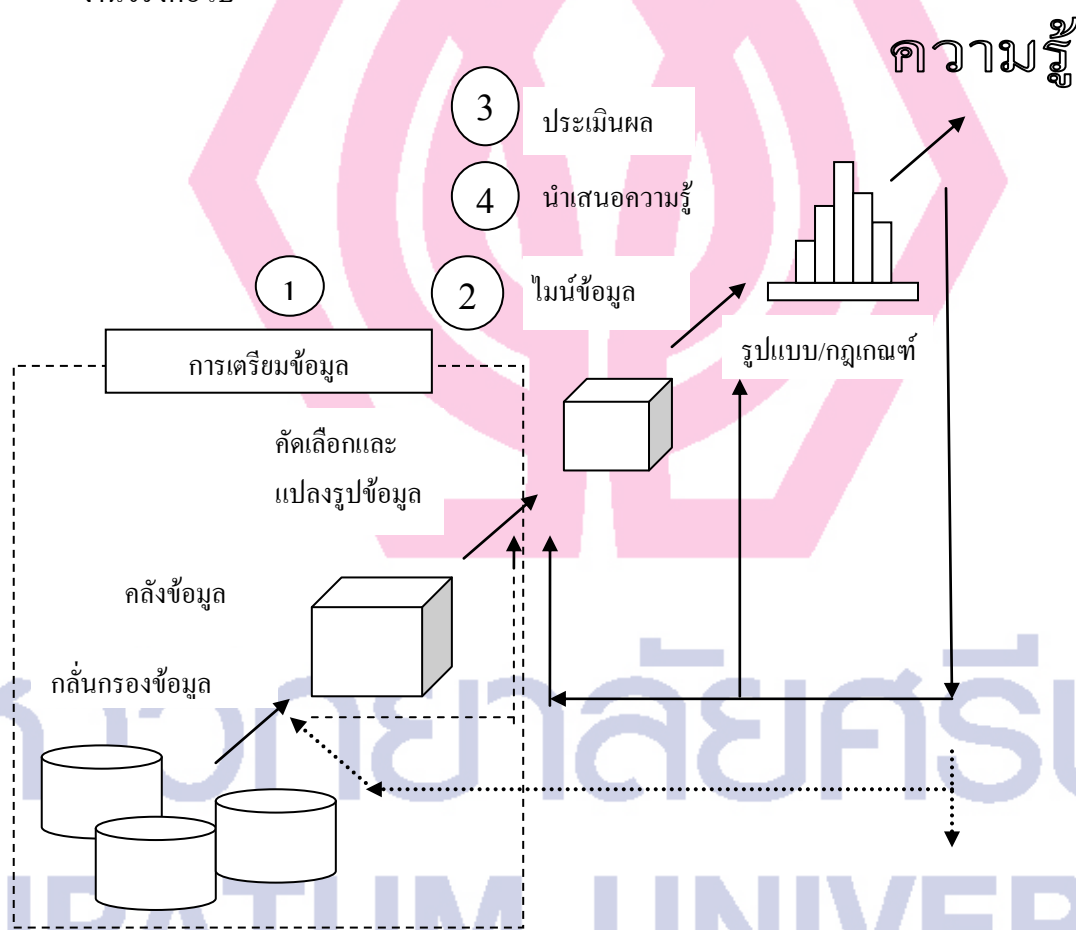
(2) การทำเหมืองข้อมูล (Data Mining)

เป็นการประมวลผลข้อมูลตามอัลกอริทึมที่ได้กำหนดไว้ ในขั้นตอนนี้จะมีความสัมพันธ์กับการวิเคราะห์ข้อมูลและขั้นตอนที่ผ่านมา โดยเมื่อทำขั้นตอนนี้แล้วอาจต้องย้อนกลับไปทำขั้นตอนการเตรียมข้อมูลใหม่

(3) การประเมินรูปแบบ/กฎเกณฑ์ที่ได้ (Pattern evaluation)

เป็นขั้นตอนการวิเคราะห์ และประเมินผลของรูปแบบหรือกฎเกณฑ์ ที่หาได้จากขั้นตอนการหาความรู้จากข้อมูล การทำงานในส่วนนี้จำเป็นต้องใช้ทักษะในการวิเคราะห์ข้อมูลทางธุรกิจเข้าช่วย

(4) การนำเสนอความรู้ (Knowledge presentation) นำความรู้ที่ค้นพบไปประยุกต์ใช้งานจริงต่อไป



รูปที่ 2.2 กระบวนการค้นพบความรู้จากฐานข้อมูล



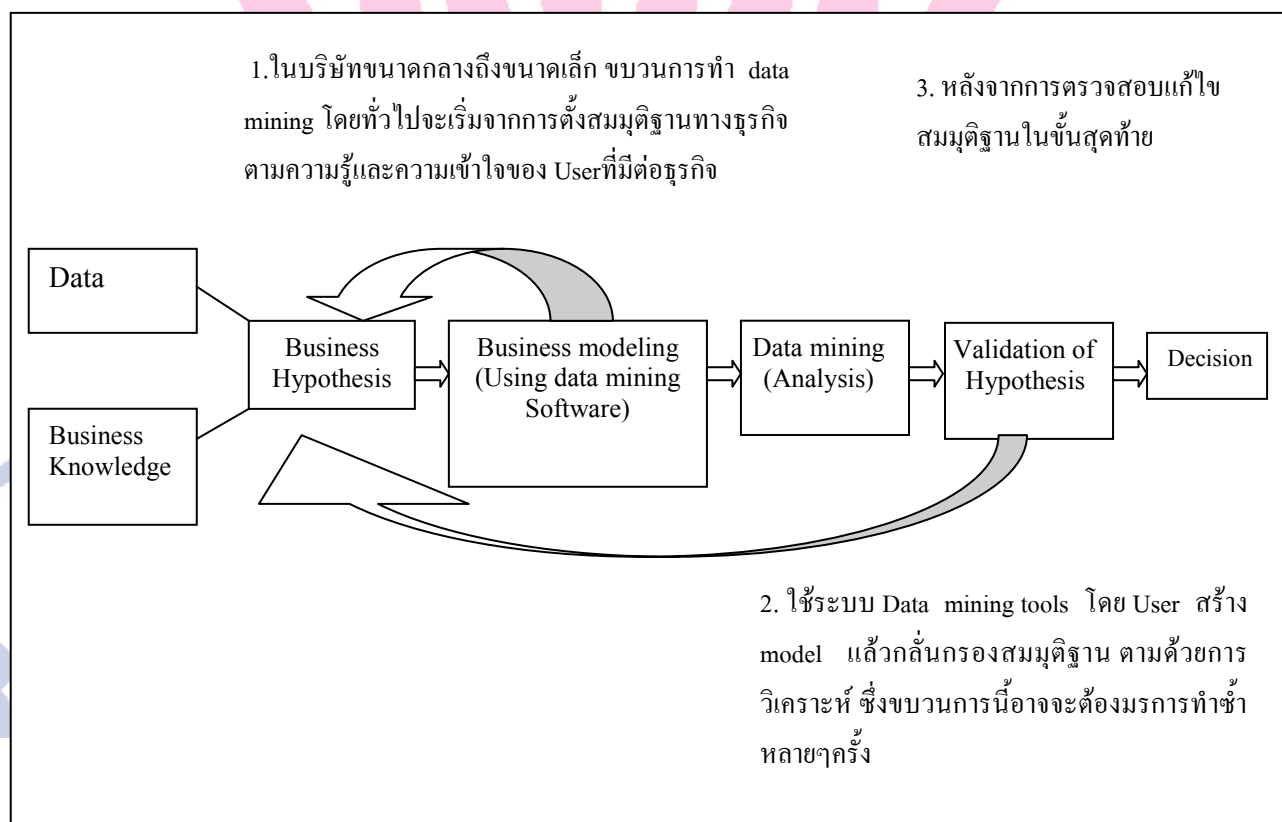
## เทคนิคในการทำเหมืองข้อมูล

เทคนิคที่ใช้ในการขุดค้นข้อมูลหรือไมน์ข้อมูลอาศัยหลักการทางสถิติในการวิเคราะห์ข้อมูล โดยมี 3 เทคนิคสำคัญที่เป็นที่รู้จักแพร่หลายในปัจจุบันดังต่อไปนี้

(1) เทคนิคการค้นหากฎความสัมพันธ์ของข้อมูล (Association Rule Discovery) เป็น การค้นหากฎความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่ที่มีอยู่ เพื่อนำไปใช้ในการวิเคราะห์หรือ ทำนายปรากฏการณ์ต่างๆ

(2) เทคนิคการจัดข้อมูลเข้ากลุ่มและการทำนาย (Classification and Prediction) เทคนิคการจัดข้อมูลเข้ากลุ่มเป็นกระบวนการการสร้าง โมเดลจัดการข้อมูลขนาดใหญ่ให้อยู่ในกลุ่มที่กำหนดมาให้ และทำการจัดข้อมูลเข้ากลุ่มตามโมเดลที่สร้างขึ้นนั้น ส่วนเทคนิคการทำนายเป็นการวิเคราะห์ทำนายค่าที่ต้องการจากข้อมูลที่มีอยู่

(3) เทคนิคการจำแนกกลุ่มข้อมูล (Clustering) เป็นการวิเคราะห์หาชั้นหรือกลุ่มของข้อมูล ให้กับข้อมูล โดยที่ไม่รู้ชื่อคลาสมาก่อนซึ่งต่างกับเทคนิคการจัดการข้อมูล และการทำนายที่วิเคราะห์หาคลาสให้กับข้อมูล โดยมีการกำหนดคลาสไว้ให้ก่อนแล้ว



รูปที่ 2.3 การแสดงขบวนการตัดสินใจ

ตัวอย่างเช่น การค้นหากฎความสัมพันธ์ (Association Rule) ของสินค้าในห้างสรรพสินค้า พบว่าลูกค้าร้อยละ 90 ที่ซื้อเบียร์ จะซื้อผ้าอ้อมเด็กด้วย ซึ่งเป็นข้อมูลให้ทางห้างสรรพสินค้าคิดรายการส่งเสริมการขายใหม่ๆ หรือธนาคารอาจพบว่าคนทั่วไปที่มีอายุ 20-29 ปี และมีรายได้ในช่วง 20,000-30,000 บาท มักซื้อเครื่องเล่น MP3 ธนาคารอาจจะเสนอให้คนกลุ่มนี้ทำบัตรเครดิตโดยแถมเครื่องเล่นดังกล่าว เป็นต้น

ในปัจจุบันองค์กรส่วนใหญ่จะเผชิญกับปัญหาของ "ข้อมูลดิบจำนวนมาก แต่ข้อมูลที่ประยุกต์ใช้ได้มีน้อย" ค่า Data Mining (Data Mining) จึงเป็นสาขาที่คาดว่าจะเป็นที่รู้จัก และนำมาใช้ประยุกต์ได้อย่างแพร่หลาย เนื่องจากการทำเหมืองข้อมูล (Data Mining) สามารถดึงความรู้จากข้อมูลจำนวนมากที่ถูกเก็บสะสมไว้

ในโลกของธุรกิจ บริษัทต่างๆ จะพยายามหาเทคนิคที่สามารถนำความสำเร็จมาสู่บริษัท เช่น ในโลกธุรกิจขนาดย่อมจะสร้างความสัมพันธ์กับลูกค้า โดยสังเกตจากความต้องการ ความชอบความสนใจของลูกค้า และอาจมีการเรียนรู้ได้จากผลสะท้อนในอดีตว่า จะทำอย่างไรให้ การบริการลูกค้ามีประสิทธิภาพดีขึ้นในอนาคต หรือบริษัทที่เป็นผู้ออกบัตรเครดิต และธนาคารต่างๆ จะมีขบวนการทำเหมืองข้อมูล (Data Mining) ให้เป็นประโยชน์ในการตัดสินใจว่ากลุ่มลูกค้าใดเป็นกลุ่มที่ดีที่ควรเข้ามาใจลูกค้าช่วยในการแยกประเภทของลูกค้า และจะทำนายกลุ่มของธุรกิจประชากรที่คาดว่าจะมาเป็นลูกค้าในอนาคต เป็นต้น อย่างไรก็ตามการเรียนรู้ที่ถูกต้องมากกว่าการเก็บสะสมข้อมูลอย่างตรงไปตรงมา ซึ่งจะทำให้การทำงานไม่เป็นประสิทธิภาพ

### 2.1.2 วัฏจักรขั้นตอนการทำงานของการทำเหมืองข้อมูล (Data Mining)

วัฏจักรขั้นตอนการทำงานของการทำเหมืองข้อมูล (Data Mining) ประกอบด้วย 4 ขั้นตอนหลักๆ

(1) การระบุโอกาสทางธุรกิจหรือการระบุปัญหาที่เกิดขึ้นกับธุรกิจ เป็นการระบุขอบเขตของข้อมูลที่จะนำมาทำการวิเคราะห์ เพื่อหาความได้เปรียบทางการตลาดเพื่อนำมาทำการแก้ไขปัญหา

(2) ส่วนของการทำเหมืองข้อมูล เป็นการนำเทคนิคของการทำเหมืองข้อมูล (Data Mining) ไปใช้ถ่ายทอด หรือทำการเปลี่ยนแปลงข้อมูลดิบให้อยู่ในรูปของข้อมูลที่จะนำไปใช้ได้จริงในทางธุรกิจ

(3) การปฏิบัติตามข้อมูล คือ การนำเอาข้อมูลที่เป็นผลลัพธ์ของส่วน Data Mining มาลองปฏิบัติจริงกับธุรกิจ

(4) การวัดประสิทธิภาพจากผลลัพธ์ การวัดประสิทธิภาพของเทคนิค ในการทำเหมืองข้อมูล (Data Mining) ที่จะนำมาใช้จากผลลัพธ์ ซึ่งสามารถตรวจสอบได้หลายทาง เช่น วัดจากส่วนแบ่งการตลาด วัดจากปริมาณลูกค้า หรือ วัดจากกำไรสุทธิ เป็นต้น จากทั้ง 4 ขั้นตอนดังกล่าว

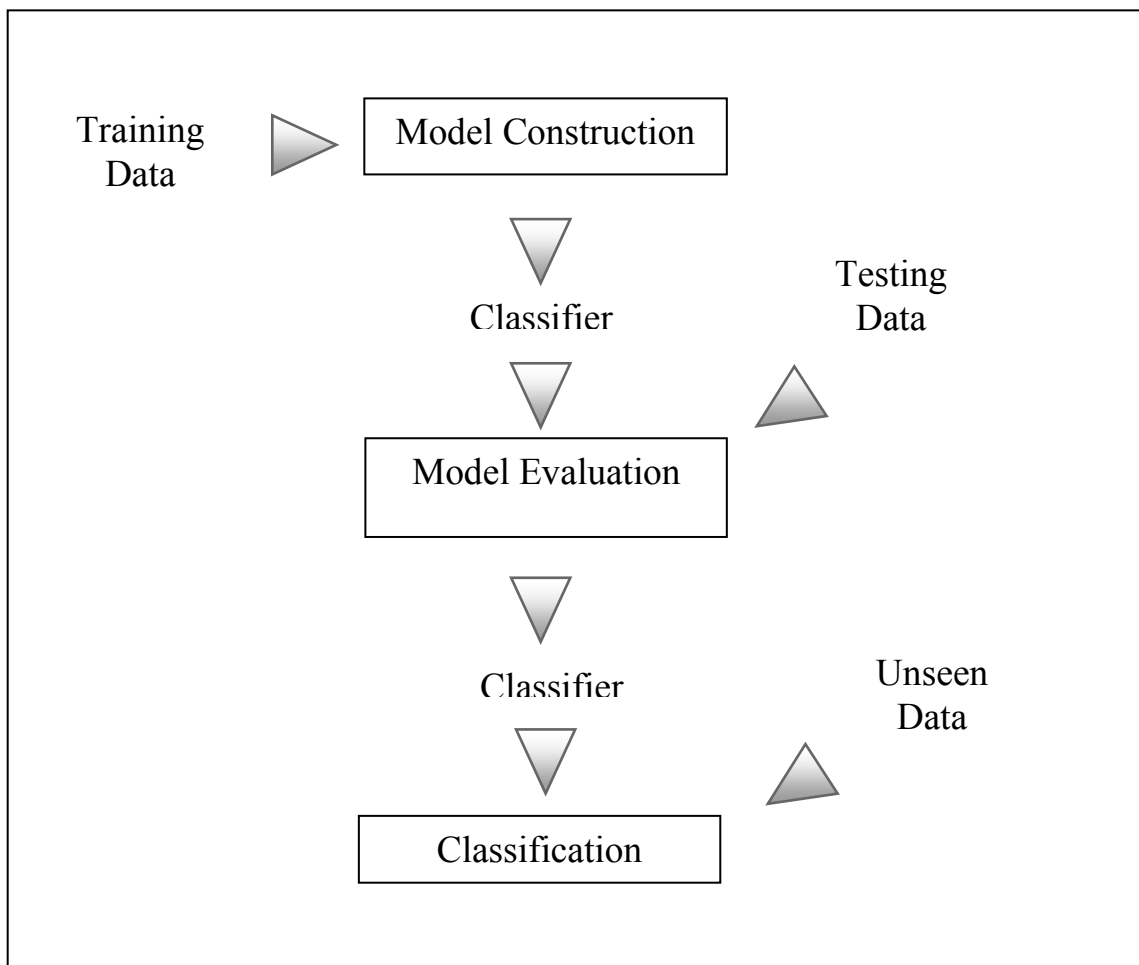
มาข้างต้น คือ การนำเอาการทำเหมืองข้อมูล (Data Mining) ไปใช้กับระบบทางธุรกิจ โดยแต่ละขั้นตอนจะพึ่งพาอาศัยกัน ผลลัพธ์จากขั้นตอนหนึ่งจะกลายมาเป็น Input จากอีกขั้นตอนต่อไปซึ่งการทำเหมืองข้อมูล (Data Mining) จะเปลี่ยนข้อมูลดิบให้เป็นข้อมูลประยุกต์ ดังนั้น การระบุแหล่งข้อมูลที่ต้องการจึงเป็นสิ่งที่สำคัญอย่างยิ่งต่อผลลัพธ์ที่ได้จากการวิเคราะห์

### 2.1.3 งานของการทำเหมืองข้อมูล (Task of data mining)

ในทางปฏิบัติการทำเหมืองข้อมูล (Data Mining) จะประสบความสำเร็จกับงานบางกลุ่มเท่านั้น และต้องอยู่ภายใต้ภาวะที่จัดปัญหาความเหมาะสมกับการใช้เทคนิคของการทำเหมืองข้อมูล (Data Mining) จะเป็นปัญหาที่ต้องใช้เหตุผลในการแก้ปัญหาเกี่ยวกับเศรษฐศาสตร์และการเงิน ซึ่งจะสามารถจัดรูปแบบของธุรกิจให้อยู่ในรูปแบบของงานทั้ง 6 งานได้ ดังนี้

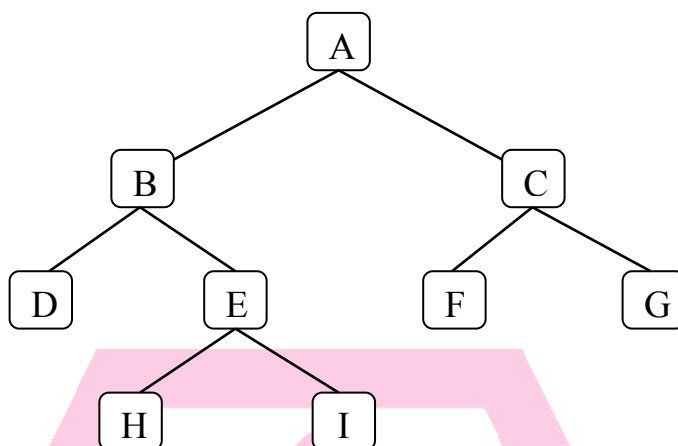
#### 1. การแบ่งประเภท (Classification)

การแบ่งประเภทถือว่าเป็นงานธรรมดาทั่วไปของการทำเหมืองข้อมูล(Data Mining) เพราะการทำความเข้าใจ และการติดต่อสื่อสารต่างๆ ที่เกี่ยวข้องกับการแบ่งเป็นประเภท การจัดแยกประเภท และการแบ่งแยกชนิดโดยการแบ่งประเภท ประกอบด้วยการสำรวจจุดเด่นของวัตถุที่ปรากฏออกมา และทำการกำหนดจุดเด่นนั้นๆ เป็นตัวที่ใช้แบ่งประเภทของงานในการแบ่งประเภท คือ การบ่งบอกลักษณะ โดยการอธิบายจุดเด่นที่เป็นที่รู้จักดีในประเภทนั้นและเทรนนิ่งเซต (Training Set) ของตัวอย่างในแต่ละประเภท ซึ่งมีภาระหน้าที่ในการสร้างโมเดลของบางชนิดที่ไม่สามารถจะจัดประเภทของข้อมูลได้ ให้สามารถจัดเป็นประเภทได้ เช่น การจัดประเภทของผู้ยื่นบัตรเครดิต เป็นระดับต่ำ ระดับกลาง และระดับสูงของความเสี่ยงที่จะได้รับดังนั้นการแบ่งประเภทข้อมูล (Data Classification) เป็นปัญหาพื้นฐานของการเรียนรู้แบบมีผู้สอนโดยปัญหา คือ การทำนายประเภทของวัตถุจากคุณสมบัติต่างๆ ของวัตถุ ซึ่งการเรียนรู้แบบมีผู้สอนจะสร้างฟังก์ชันเชื่อมโยง ระหว่างคุณสมบัติของวัตถุกับประเภทของวัตถุ จากตัวอย่างสอนแล้วจึงใช้ฟังก์ชันทำนายประเภทของวัตถุที่ไม่เคยพบ เครื่องมือหรือขั้นตอนวิธีที่ใช้สำหรับการแบ่งประเภทของข้อมูล เช่น โครงข่ายประสาทเทียม(Neural Induction) ต้นไม้ตัดสินใจ(Decision Tree)



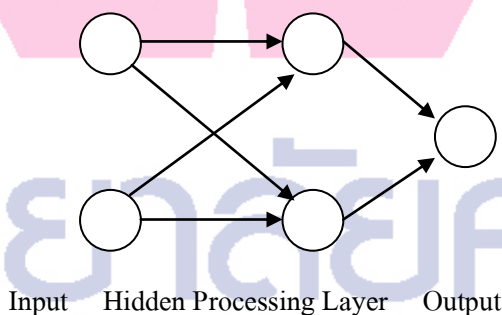
รูปที่ 2.4 กระบวนการ Classification

ต้นไม้ตัดสินใจ (Decision Tree) มีลักษณะคล้ายโครงสร้างต้นไม้ ซึ่งประกอบด้วย โหนดภายใน (Internal node) จะแสดงลักษณะ (Attribute) ของข้อมูล โดยจุดที่เริ่มต้นของต้นไม้เรียกว่า โหนดรากแต่ละกิ่งแสดงค่าของคุณลักษณะของแต่ละโหนด และลิฟโหนดแสดงกลุ่ม (Class) ซึ่งเป็นผลลัพธ์ที่สามารถแยกแยะได้



รูปที่ 2.5 โครงสร้างต้นไม้ตัดสินใจ

เครือข่ายประสาท (Neural Network) เป็นเทคโนโลยีที่มีมาจากการวิจัยด้านปัญญาประดิษฐ์ (Artificial Intelligence : AI) เพื่อใช้ในการคำนวณค่าฟังก์ชันจากกลุ่มข้อมูล วิธีการของเครือข่ายประสาทเป็นวิธีการที่ให้เครื่องเรียนรู้จากตัวอย่างต้นแบบแล้วฝึก ให้ระบบได้รู้จักคิดแก้ปัญหาที่กว้างขึ้นได้ในโครงสร้างของเครือข่ายประสาทจะประกอบด้วยโหนด (node) สำหรับ Input-Output ปลดการประมวลผล กระจายอยู่ในโครงสร้างเป็นชั้นๆ ได้แก่ input layer , output layer และ hidden layer การประมวลผลของเครือข่ายประสาทจะอาศัยการทำงานผ่าน โหนดต่างๆ ใน layer เหล่านี้ ตัวอย่างโครงสร้างแบบประสาท



รูปที่ 2.6 โครงสร้างเครือข่ายประสาท

## 2. การประเมินค่า (Estimation)

การประเมินค่าทางธุรกิจอย่างต่อเนื่อง จะก่อให้เกิดผลลัพธ์ที่มีประโยชน์กับธุรกิจ การป้อนข้อมูลที่เราได้อยู่เข้าไป เพื่อใช้ในการประเมินสิ่งต่างๆ ที่จะก่อให้เกิดประโยชน์หรือสำหรับตัวแปรที่เราไม่รู้ค่าแน่นอน เช่น รายได้จากการค้า จุดสูงสุดทางธุรกิจ หรือคุณภาพของบัตรเครดิตในทางปฏิบัติการประเมินค่าจะถูกใช้ในการทำงานการจัดหมวดหมู่ ตัวอย่างของการประเมินค่าเช่น การประเมินรายได้รวมของครอบครัว หรือการประเมินจำนวนบุตรในครอบครัว

## 3. การทำนายล่วงหน้า (Prediction)

การทำนายล่วงหน้า เป็นงานที่ลักษณะคล้ายกับการจัดหมวดหมู่ หรือ การประเมินค่า ยกเว้นเพียงแต่จะใช้สถิติการบันทึกของการจัดหมวดหมู่ในการทำนายอนาคตของพฤติกรรม หรือ การประเมินค่าที่จะเกิดขึ้นในอนาคต ตัวอย่างของการทำนายล่วงหน้า เช่น การทำนายการเปลี่ยนแปลงพฤติกรรมของตลาด หรือการทำนายจำนวนลูกค้าที่จะออกจากธุรกิจของเรา ใน 6 เดือนข้างหน้า เป็นต้น

## 4. การจัดกลุ่มโดยอาศัยความใกล้ชิด (Affinity Group)

งานในการจัดกลุ่มหรือการวิเคราะห์ตลาด คือ การตัดสินใจรวมสิ่งที่สามารถไปด้วยกันเข้าไว้ในกลุ่มเดียวกัน ตัวอย่างของการจัดกลุ่มโดยอาศัยความใกล้ชิดกัน หรือ การวิเคราะห์ของตลาด เช่น การตัดสินใจว่ามีสิ่งใดบ้างที่จะไปอยู่ด้วยกันอย่างสม่ำเสมอในรถเข็นในซูเปอร์

## 5. การแบ่งกลุ่ม (Clustering)

การแบ่งกลุ่มข้อมูล เป็นขั้นตอนเบื้องต้นของการวิเคราะห์ข้อมูล ซึ่งใช้ในการเรียนรู้ของเครื่อง การทำเหมืองข้อมูล โดยจะแบ่งชุดข้อมูลออกเป็นกลุ่ม นำข้อมูลที่มีคุณลักษณะเหมือนกันหรือคล้ายกันจัดไว้ในกลุ่มเดียวกัน ขั้นตอนวิธีที่ใช้ในการแบ่งกลุ่มจะอาศัยความเหมือน หรือความใกล้ชิด โดยคำนวณจากการวัดระยะห่างเวกเตอร์ของข้อมูลเข้า โดยใช้การวัดระยะแบบต่างๆ เช่น การวัดระยะแบบยูคลิด การวัดระยะแบบแมนฮัตตัน การวัดระยะแบบเชบิเชฟ เพื่อช่วยในการลดขนาดข้อมูล(แยกเป็นหลายๆ กลุ่ม และคัดเลือกบางกลุ่มเพื่อทำการวิเคราะห์ต่อไป หรือแยกการวิเคราะห์ออกเป็นสำหรับแต่ละกลุ่ม)ก่อนที่จะนำไปวิเคราะห์ด้วยวิธีการอื่นต่อไป

การแบ่งกลุ่มข้อมูลจะแตกต่างกันจากการแบ่งประเภทข้อมูล(Classification) โดยจะแบ่งกลุ่มข้อมูลจากความคล้าย โดยไม่มีการกำหนดประเภทของข้อมูลไว้ก่อน

ขั้นตอนวิธีการแบ่งกลุ่ม ได้แก่

- k-means clustering
- hierarchical clustering
- self-organizing map(som)

อัลกอริทึมในการแบ่งกลุ่มข้อมูล โดยทั่วไปแบ่งได้เป็น 2 ประเภทใหญ่ๆ คือ

1. การแบ่งแบบเป็นลำดับชั้น(Hierarchical) จะมีการแบ่งกลุ่มจากกลุ่มย่อยที่ถูกแบ่งไว้ก่อนหน้านั้นซ้ำหลายครั้ง

2. การแบ่งเป็นสัดส่วน(Partitional) การจัดการ การแบ่งจะทำเพียงครั้งเดียว การแบ่งแบบเป็นลำดับชั้น จะมี 2 ลักษณะ คือ

-แบบล่างขึ้นบน(bottom-up) หรือ เป็นการแบ่งแบบรวมกลุ่มจากกลุ่มย่อยให้ใหญ่ขึ้นไปเรื่อยๆ โดยเริ่มจากกลุ่มเล็กสุด คือ ในแต่ละกลุ่มมีข้อมูลเพียงตัวเดียว

-แบบบนลงล่าง(top-down) หรือเป็นการแบ่งกลุ่มจากกลุ่มใหญ่ให้ย่อยไปเรื่อยๆ โดยเริ่มจากกลุ่มใหญ่ที่สุด คือ กลุ่มเดียวมีข้อมูลทุกตัวในกลุ่ม

## 6. การบรรยาย (Description)

ในบางครั้งวัตถุประสงค์ของการทำเหมืองข้อมูล (Data Mining) คือ ต้องการอธิบายความสัมพันธ์ของฐานข้อมูลในทางที่จะเพิ่มความเข้าใจในส่วนของประชากร ผลิตภัณฑ์ หรือ ขบวนการให้มากขึ้น

เทคนิคการทำเหมืองข้อมูล (Data Mining) ส่วนใหญ่ต้องการเทรนนิ่งข้อมูลจำนวนมากที่ประด้วยหลายๆตัวอย่าง เพื่อจะสร้างกฎที่ใช้ในการจัดหมวดหมู่ กฎความสัมพันธ์คลัสเตอร์ การทำนายล่วงหน้า ดังนั้นชุดของข้อมูลขนาดเล็กจะนำไปสู่ความไม่แน่ใจของผลสรุปที่ได้

ไม่มีเทคนิคใดสามารถแก้ปัญหาของ (Data Mining) ได้ทุกปัญหา ดังนั้นความหลากหลายของเทคนิคจึงเป็นสิ่งจำเป็นในการไปสู่การแก้ปัญหาของ (Data Mining) ได้ดีที่สุด

### 2.1.4 เทคนิคของการทำเหมืองข้อมูล (Data Mining)

การแก้ปัญหาของงานชนิดต่างๆ โดยใช้วิธีการทำเหมืองข้อมูล (Data Mining) ในแต่ละงานก็จะมีเทคนิคของการทำเหมืองข้อมูล (Data Mining) ที่จะนำมาใช้ได้อย่างเหมาะสม โดยเทคนิคในการทำเหมืองข้อมูล (Data Mining) นั้นมีมากมาย ซึ่งเทคนิคที่ถูกใช้กันอย่างแพร่หลายมีดังนี้

#### 1. นิวตรอนเน็ตเวิร์ค

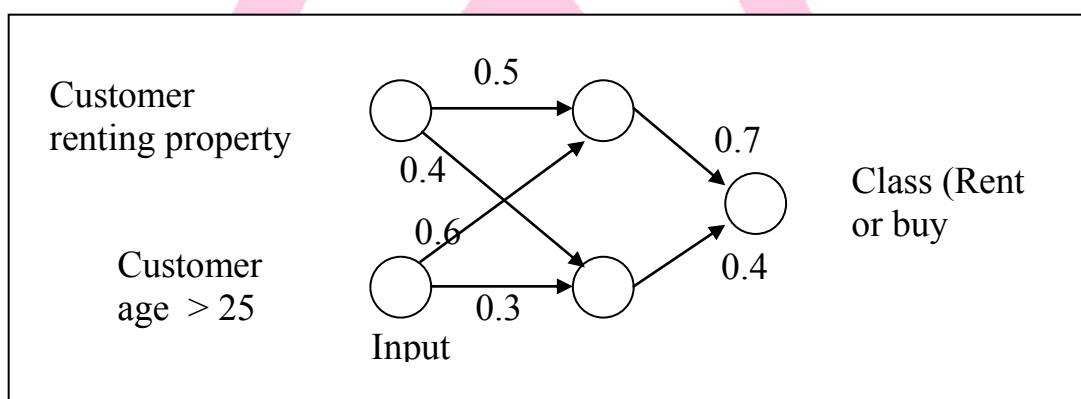
นิวตรอนเน็ตเวิร์ค คือ ระบบที่มีการประมวลผลข้อมูลซึ่งรวมคุณสมบัติของไบโโอลิจิตอล โดยถูกพัฒนาขึ้นด้วยโมเดลทางคณิตศาสตร์ของกระบวนการเรียนรู้ของมนุษย์ และเรียนรู้จากชุดข้อมูลของชุดความรู้

นิวตรอนเน็ตเวิร์ค ประกอบด้วย หน่วยความจำจำนวนมากเรียกว่า นิวตรอนเซลล์ หรือ โหนด แต่ละนิวตรอนต่อกัน โคนคอนเน็คชั่นลิงค์ (Connection Link) ที่ค่าน้ำหนักของมันอยู่โดยค่าน้ำหนักจะแสดงรายละเอียดที่เน็ตเวิร์คใช้ในการแก้ปัญหา โดยนิวตรอนเน็ตเวิร์คถูกใช้ในการแก้ปัญหาอย่างกว้างขวาง เช่น การเก็บ และการเรียกข้อมูล การแยกประเภทของข้อมูล การ

เปลี่ยนรูปแบบของ input ให้อยู่ในรูปแบบของ output ความสามารถในการตรวจสอบรูปแบบของ ข้อมูลที่คล้ายคลึงกับความคิดมนุษย์ เป็นต้น แต่นิวตรอนเน็ตเวิร์คสามารถนำไปประยุกต์ใช้กับงาน หลายๆชนิดได้อย่างมีประสิทธิภาพ แต่นิวตรอนเน็ตเวิร์คก็ยังมีข้อเสียอยู่บ้าง ดังนี้

1. เป็นวิธีที่ยากต่อการทำความเข้าใจในโมเดลที่ถูกผลิตออกมา

2. มีคุณสมบัติที่ไวต่อรูปแบบของ input ถ้าเราแทนข้อมูลด้วยรูปแบบที่แต่ต่างกันก็ จะสามารถผลิตผลลัพธ์ที่แตกต่างกันออกมา ดังนั้นการกำหนดค่าเริ่มต้นให้กับข้อมูลจึงเป็นส่วนที่มีความสำคัญส่วนหนึ่ง



รูปที่ 2.7 นิวรอลเน็ต เพื่อการวิเคราะห์การเช่าและซื้อบ้าน

## 2. จีเนติก อัลกอริทึม (Genetic Algorithm : GA)

ในยีนส์ของมนุษย์นั้น จะมีการถ่ายทอดพันธุกรรมไปยังลูกหลานได้ ซึ่งจีเนติก อัลกอริทึม จะอาศัยหลักการนี้ ข้อมูลชุดหนึ่ง ซึ่งมีกฎของตัวเอง หากมีการนำข้อมูลทั้งสองชุดมา รวมกันเป็นรูปแบบนี้ จะมีการสร้างกฎขึ้นมาโดยวิเคราะห์จากกฎที่มีอยู่ของทั้ง 2 รูปแบบข้อมูล

จีเนติก อัลกอริทึม เป็นทฤษฎีที่จำลองกระบวนการวิวัฒนาการทางธรรมชาติ คือการ คัดเลือกทางธรรมชาติ โดยอาศัยพื้นฐานความคิดทางพันธุกรรมในการถ่ายทอดลักษณะต่างๆ ไปยังรุ่น ถัดไปที่สามารถนำมาพัฒนาใช้ในการหาคำตอบที่เหมาะสมที่สุด ของแต่ละปัญหาจีเนติกอัลกอริทึม เป็นวิธีการหาคำตอบโดยการพิจารณา และดำเนินการจากกลุ่มของคำตอบของปัญหาที่ถูกสร้างขึ้นมา โดยการเข้ารหัส คือ การแปลงค่าตัวแปรหรือพารามิเตอร์ของปัญหาให้อยู่ในรูปแบบ โครงสร้างของ โครโมโซมที่กำหนด เพื่อคัดเลือกโครโมโซมที่เหมาะสมสำหรับสร้างวิวัฒนาการให้ดีขึ้นตาม กระบวนการทางพันธุศาสตร์ โดยการแลกเปลี่ยนค่าพารามิเตอร์ต่างๆ ระหว่างโครโมโซมที่ถูก คัดเลือก อันจะทำให้คำตอบของปัญหาถูกปรับปรุงให้ดีขึ้น จีเนติก อัลกอริทึม ใช้กระบวนการหลักๆ 3 กระบวนการในการหาคำตอบที่ใกล้เคียงหรือดีที่สุดของปัญหาดังนี้



1.การคัดเลือก (Selection) : คัดเลือกอันที่ดีที่สุดซึ่งจริงๆแล้วถ้าผ่านขั้นตอนต่อไป อาจจะเป็นค่าที่ใช้ไม่ได้ก็ได้

2.การครอสโอเวอร์ (Crossover) : ค่าจากขั้นตอนที่ 1 มาสับเปลี่ยนบิต(bit) เพื่อ ประเมินแล้วเลือก

3.การมิวเตชัน (Mutation) : ค่าจากขั้นตอนที่ 3 นำกลับมาบิดจาก 0 เป็น 1 เพื่อ ประเมินแล้วกลับไปขั้นตอนที่ 1 อีกครั้ง ถึงแม้ว่าในปัจจุบัน จีเนติก อัลกอริทึม ยังเป็นวิธีการที่ไม่ได้ แพร่หลาย แต่สาขาวิชาทางด้านจีเนติกอัลกอริทึม นับว่าเป็นอีกสาขาวิชาหนึ่งที่สนใจ และน่าจะเป็น วิธีที่ได้รับความนิยมในอีกไม่กี่ปีข้างหน้า เนื่องมาจากสามารถนำมาประยุกต์ใช้ได้กับหลายๆปัญหา รวมทั้งปัญหาทางการทำเหมืองข้อมูล

### 3. ดีซีชันทรี (Decision Tree)

เป็นแบบจำลองที่มีลักษณะคล้ายกับต้นไม้ จะมีการสร้างกฎต่างๆ ขึ้นเพื่อใช้ในการ ตัดสินใจ ดีซีชันทรี เป็นวิธีที่ได้รับความนิยม เนื่องจากความไม่ซับซ้อนของอัลกอริทึม ทำให้ เครื่องมือที่ใช้ในการทำวางขายกันอยู่ในท้องตลาดต่างๆ ข้อดีของวิธีนี้คือ สามารถตีความและเข้าใจ ลักษณะของรูปแบบข้อมูลได้ง่าย เพราะมีการแยกออกเป็นกฎ หรือข้อกำหนดต่างๆ แต่ก็ยังคงมีปัญหา ในเรื่องของการห้ำนน้ำหนักความน่าเชื่อถือ หรือ การให้ค่าน้ำหนักในแต่ละโหนด ซึ่งถ้าให้น้ำหนัก ผิดไป อาจจะทำให้การตีความผิดไปได้

### 4. คลัสเตอร์ลิ่ง (Clustering)

วิธีคลัสเตอร์ลิ่ง นี้เป็นวิธีที่อาจเรียกว่าเป็นการทำเหมืองข้อมูลแบบ อ้อมๆ ได้ เนื่องจากการหาผลลัพธ์ในแต่ละครั้งนั้น แม้กระทั่งผู้หายัง ไม่อาจจะทราบว่าสิ่งที่ต้องการจะหานั้นคือ อะไร จำเป็นต้องรองจนกว่าการค้นหาจะทำเสร็จสมบูรณ์จึงจะทราบข้อมูลที่ซ่อนอยู่ เปรียบเสมือนการ มีข้อมูลจำนวนมากมาอยู่ในตะกร้า แล้วจากนั้นมีเวทย์มนต์มาจัดเรียงข้อมูลหน่วยนั้นให้อยู่เป็นกลุ่ม ซึ่งทำให้สังเกตเห็นลักษณะเด่นที่ซ่อนเร้นอยู่ภายในข้อมูลจำนวนมากหน่วยนั้น

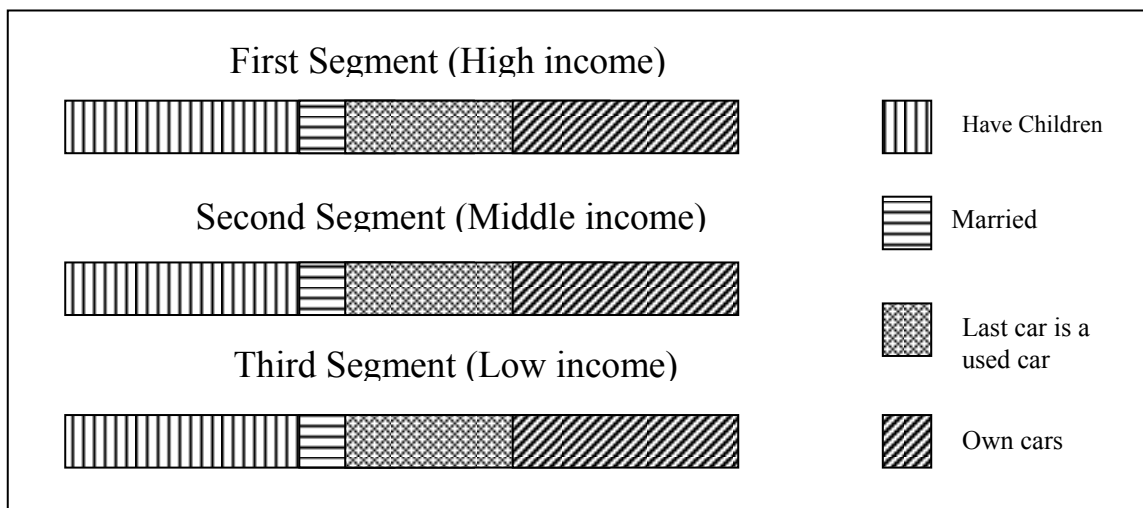
ตัวอย่าง เช่น บริษัทจำหน่ายรถยนต์ได้แยกกลุ่มลูกค้าออกเป็น 3 กลุ่ม คือ

- 1.กลุ่มผู้มีรายได้สูง (>\$80,000)
- 2.กลุ่มผู้มีรายได้ปานกลาง (\$25,000 to 80,000)
- 3.กลุ่มผู้มีรายได้ต่ำ(less than \$25,000)

และภายในแต่ละกลุ่มยังแยกออกเป็น

-Have Children

-Married



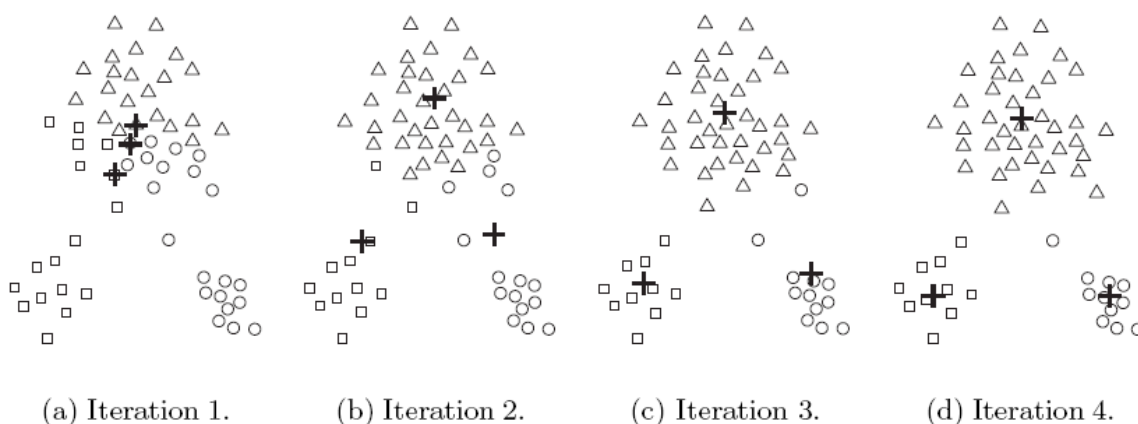
รูปที่ 2.8 ตัวอย่าง Clustering

จากข้อมูลข้างต้น ทำให้ทางบริษัทรู้ว่าเมื่อมีลูกค้าเข้ามาที่บริษัท ควรจะเสนอขายรถประเภทใด เช่น ถ้าเป็นกลุ่มผู้มีรายได้สูงควรจะเสนอรถใหม่ เป็นรถครอบครัวขนาดใหญ่พอสมควร แต่ถ้าเป็นผู้มีรายได้ค่อนข้างต่ำควรเสนอรถมือสอง ขนาดค่อนข้างเล็ก

การแบ่งกลุ่มข้อมูล (Data Clustering) เป็นศาสตร์แขนงหนึ่งทางเหมืองข้อมูล (Data Mining) เป็นวิธีการแบ่งกลุ่มข้อมูลโดยไม่อาศัยตัวอย่างในการเรียนรู้ (unsupervised learning) โดยจะทำการแบ่งกลุ่มข้อมูลออกเป็นกลุ่มย่อยๆ เรียกว่าคลัสเตอร์ (cluster) โดยพิจารณาจากความคล้ายคลึงกันของข้อมูลเป็นเกณฑ์หลักในการแบ่งกลุ่มข้อมูล ดังนั้นกลุ่มข้อมูลที่มีลักษณะคล้ายคลึงกันจะถูกจัดให้อยู่ในกลุ่มเดียวกัน และข้อมูลที่มีลักษณะแตกต่างกันจะถูกจัดไว้ต่างกลุ่มกัน วิธีการแบ่งกลุ่มนี้มีประโยชน์อย่างมากในการวิเคราะห์ความสัมพันธ์กันของชุดข้อมูล เทคนิคการแบ่งกลุ่มในปัจจุบันมีหลากหลายเทคนิค แต่ที่ใช้กันแพร่หลายได้แก่ การแบ่งกลุ่มแบบพาร์ทิชัน (partition based) การแบ่งกลุ่มแบบใช้ความหนาแน่น (density based) การแบ่งกลุ่มโดยใช้กริด (grid based) และการแบ่งกลุ่มแบบมีลำดับชั้น (hierarchical based) ซึ่งในงานวิจัยนี้ใช้วิธีการแบ่งกลุ่มแบบพาร์ทิชัน ด้วยอัลกอริทึม K-Means ซึ่งจะกล่าวถึงรายละเอียดในส่วนถัดไป

#### 4.1 K-Mean Clustering

เทคนิคการจัดกลุ่ม K-means เป็นวิธีการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) ที่ง่ายที่สุด หลักการของการจัดกลุ่มด้วย K-Means นั้น อัลกอริทึมจะทำการแบ่งข้อมูล (Partition) ออกเป็นกลุ่ม จำนวน K กลุ่ม (Cluster)ตามที่ผู้ใช้งานกำหนด โดยในแต่ละกลุ่มจะมีข้อมูลจุดศูนย์กลางประจำกลุ่ม (centroid) ซึ่งได้จากการคำนวณค่าเฉลี่ยของกลุ่ม และการพิจารณาว่าข้อมูลหนึ่งๆ ควรอยู่กลุ่มใดนั้น จะทำการพิจารณาระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางประจำกลุ่มแต่ละกลุ่ม ข้อมูลที่มีระยะห่างจากจุดศูนย์กลางประจำกลุ่มใดน้อยที่สุดจะถูกจัดอยู่ในกลุ่มนั้น เมื่อข้อมูลทุกตัวถูกพิจารณาจัดกลุ่มแล้ว อัลกอริทึมจะทำการคำนวณหาค่าเฉลี่ยจุดศูนย์กลางประจำกลุ่มใหม่ หากจุดศูนย์กลางของแต่ละกลุ่มมีการเปลี่ยนตำแหน่ง อัลกอริทึมจะทำการวนซ้ำเพื่อจัดกลุ่มข้อมูลตามจุดศูนย์กลางกลุ่มที่เปลี่ยนไป และจะหยุดกระบวนการเมื่อจุดศูนย์กลางประจำกลุ่มของทุกกลุ่มไม่มีการเปลี่ยนแปลง สุดท้ายจะได้ข้อมูลที่มีลักษณะเหมือนหรือคล้ายกันอยู่ในกลุ่มเดียวกัน และข้อมูลที่มีลักษณะต่างกันอยู่คนละกลุ่มกัน



รูปที่ 2.9 แสดงขั้นตอนการจัดกลุ่มข้อมูลด้วยเทคนิค Clustering

จากรูปที่ 2.9 แสดงการจัดกลุ่มข้อมูล โดยรูปที่ 2.9(a) เริ่มจากกำหนดผู้ใช้งานทำการกำหนดจำนวนกลุ่มที่ต้องการจัดกลุ่มข้อมูลเป็น 3 กลุ่ม จากนั้นอัลกอริทึมทำการสุ่มกำหนดจุดศูนย์กลางประจำกลุ่มเริ่มต้น แทนด้วยสัญลักษณ์ + อัลกอริทึม K-Mean ทำการคำนวณระยะห่างระหว่างข้อมูลและจุดศูนย์กลางกลุ่มแต่ละกลุ่ม หากข้อมูลห่างจากกลุ่มใดน้อยที่สุดจะถูกจัดว่าเป็นสมาชิกของกลุ่มนั้น รูปที่ 2.9 (b) และ รูปที่ 2.9 (c) อัลกอริทึมคำนวณจุดศูนย์กลางประจำกลุ่มใหม่ และพบว่ามีการเปลี่ยนแปลง จึงทำการวนซ้ำเพื่อคำนวณระยะห่างระหว่างข้อมูลและจุดศูนย์กลาง และพิจารณาการจัดกลุ่มข้อมูลใหม่ และเมื่อพบว่ากลุ่มข้อมูลไม่มีการเปลี่ยนแปลงแล้วจึงหยุดการทำงาน และได้ผลลัพธ์รูปที่ 2.9 (d)

### อัลกอริทึมการจัดกลุ่มโดย K-means

1. ผู้ใช้งานกำหนดจำนวนกลุ่มที่ต้องการจัดกลุ่มข้อมูลทั้งหมดออกเป็น K กลุ่ม
2. อัลกอริทึมทำการสุ่มกำหนดจุดศูนย์กลางประจำกลุ่มข้อมูลแต่ละกลุ่ม
3. อัลกอริทึมทำการหาระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางประจำกลุ่มทุกกลุ่ม เพื่อพิจารณาความเหมือนหรือความคล้ายกันของข้อมูล
4. ข้อมูลที่มีระยะห่างจากจุดศูนย์กลางของกลุ่มใดน้อยที่สุด จะถูกจัดอยู่กลุ่มเดียวกับกลุ่มของจุดศูนย์กลางประจำกลุ่มนั้น
5. เมื่อข้อมูลทุกตัวถูกจัดกลุ่มครบแล้ว อัลกอริทึมจะคำนวณจุดศูนย์กลางประจำกลุ่ม K กลุ่มใหม่ โดยพิจารณาค่าเฉลี่ยของข้อมูลทุกตัวที่อยู่ในกลุ่ม
6. ทำซ้ำในข้อ c จนกระทั่งจุดศูนย์กลางประจำกลุ่มไม่เปลี่ยนแปลง

### 4.2 ฟังก์ชันวัดระยะห่างระหว่างข้อมูล (Dissimilarity Function)

ฟังก์ชันวัดระยะห่างของข้อมูลใช้ในการพิจารณาความเหมือน หรือความคล้ายกันของข้อมูล โดยฟังก์ชันที่นิยม สำหรับใช้ในเทคนิคการจัดกลุ่ม ได้แก่ ฟังก์ชันระยะห่างยูคลิเดียน (Euclidean Distance) ฟังก์ชันระยะห่างแมนฮัตตัน (Manhattan Distance) หรือ ชิตีบล็อก (City-Block) เป็นต้น

#### 4.2.1 Euclidean Distance

เป็นฟังก์ชันระยะห่างที่นิยมใช้ในเทคนิค K-Mean โดยมีสูตรการคำนวณดังนี้

$$d_{Euclidean}(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}$$

โดยที่  $x$  คือข้อมูลใดๆ ที่ต้องการหาระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางกลุ่มของ Cluster K  
 $y$  คือข้อมูลที่เป็นจุดศูนย์กลางกลุ่มของ Cluster K  
 $m$  คือจำนวนแอตทริบิวต์

#### 4.2.2 Manhattan Distance หรือ City-Block Distance

เป็นฟังก์ชันวัดระยะห่างโดยพิจารณาความแตกต่างของมิติค่าเฉลี่ย โดยมีสูตรการคำนวณดังนี้

$$d_{manhattan}(x, y) = \frac{1}{m} \sum_{i=1}^m |x_i - y_i|$$

โดยที่  $x$  คือข้อมูลใดๆ ที่ต้องการหาระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางกลุ่มของ Cluster K

$y$  คือข้อมูลที่เป็นจุดศูนย์กลางกลุ่มของ Cluster K

$m$  คือจำนวนแอตทริบิวต์

### 2.1.5 เทคนิคการจำแนกประเภทข้อมูล JRip

เทคนิค JRip เป็นเทคนิคที่ใช้สำหรับจำแนกประเภทข้อมูล (Classification) ที่ถูกพัฒนามาจากเทคนิค IREP [1] ซึ่งเทคนิค JRip นี้จะประกอบด้วย 3 ขั้นตอน คือ

1. ขั้นการสร้างกฎเพื่อใช้จำแนกประเภทข้อมูล
2. ขั้นการปรับแต่งกฎ
3. ขั้นการลบกฎ

การทำงานของแต่ละขั้นมีดังต่อไปนี้

เริ่มแรกจะกำหนดให้โมเดลที่เก็บกฎหรือ RuleSet เป็นเซตว่างเสียก่อน หลังจากนั้น

**ขั้นที่ 1** จะเป็นขั้นตอนที่ทำการสร้างกฎเพื่อเพิ่มเข้าไปใน RuleSet และกฎเหล่านี้จะใช้ในการจำแนกประเภทข้อมูล โดยการทำงานของขั้นนี้จะวนรอบทำงานข้อมูลที่อยู่ในคลาส Pos (สมมุติว่าข้อมูลมี 2 คลาส คือ คลาส Pos และ คลาส Neg) ถูกนำมาสร้างกฎทุกตัวแล้วหาค่าความผิดพลาด (error rate) ของโมเดลหรือ RuleSet ที่สร้างได้ มีค่าเกินกว่า 50% ในขั้นตอนที่ 1 นี้จะประกอบด้วย 2 ขั้นตอนย่อย คือ

1. **ขั้น Grow** ในขั้นนี้จะเป็นการเพิ่มแอตทริบิวต์ต่างๆ เข้าไปในส่วนเงื่อนไขของกฎ โดยการเพิ่มเข้าไปนี้จะเป็นการเพิ่มแบบ greedy คือ ถ้าเพิ่มแอตทริบิวต์เข้าไปแล้วทำให้ค่าความถูกต้องเพิ่มมากขึ้นก็จะเพิ่มแอตทริบิวต์นั้นเข้าไปในกฎเรื่อยๆ ทำเช่นนี้จนกระทั่งได้กฎที่ดี (เช่น อาจจะได้กฎที่มีความถูกต้อง 100%) โดยค่าในแต่ละแอตทริบิวต์ที่เพิ่มเข้าไปจะต้องเป็นค่าที่ทำให้ค่า information gain มีค่ามากที่สุด

2. **ขั้น Prune** กฎที่สร้างได้จากขั้นตอนที่ผ่านมาจะถูกนำมา prune โดยจะลบบางแอตทริบิวต์ออกแล้วทำการตรวจสอบว่ากฎที่ถูก prune แล้วนี้ครอบคลุมข้อมูลของคลาส Pos และคลาส Neg เพิ่มมากขึ้นเท่าใดและการ prune นี้จะสิ้นสุดเมื่อกฎที่ prune ได้ครอบคลุมข้อมูลเป็นจำนวนลดลง หลังจากนั้นจึงนำกฎที่ prune ได้นี้เก็บเข้าไปใส่ไว้ใน RuleSet เพื่อใช้เป็นโมเดลสำหรับการจำแนก

ประเภทข้อมูลต่อไป และทำการลบข้อมูลที่กฎนี้ครอบคลุมอยู่

**ขั้นที่ 2** เป็นขั้นการปรับแต่งกฎที่อยู่ใน RuleSet ซึ่งสร้างได้จากขั้นที่ 1 ในขั้นนี้จะสร้างกฎเพิ่มขึ้นมาอีก 2 แบบคือกฎแบบ replacement และกฎแบบ revision โดยแต่ละแบบมีลักษณะดังนี้

- กฎแบบ replacement จะใช้การสร้างกฎตามขั้นที่ 1.1 แต่ในขั้นตอนการ prune กฎจะพิจารณาผลรวมความถูกต้องของทุกกฎใน RuleSet แทน
- กฎแบบ revision จะทำการเพิ่มแอตทริบิวต์เข้าไปในกฎที่สร้างได้จากขั้นที่ 1 แล้ว และในขั้นตอนการ prune กฎจะพิจารณาผลรวมความถูกต้องของทุกกฎใน RuleSet

สิ่งที่แตกต่างกันระหว่างกฎแบบ replacement และกฎแบบ revision คือขั้นตอนการเพิ่มแอตทริบิวต์ซึ่งแบบแรกจะเพิ่มเข้าไปในกฎที่ว่างๆ ในขณะที่แบบหลังจะเพิ่มเข้าไปในกฎเดิมที่มีอยู่ จากกฎทั้ง 3 แบบ (กฎที่ได้จากขั้นที่ 1, กฎแบบ replacement และกฎแบบ revision) นี้กฎที่มีค่า description length น้อยที่สุดจะถูกเลือกเพื่อใช้เป็นโมเดลต่อไป และถ้ายังมีข้อมูลอยู่ในคลาส Pos อยู่ก็จะกลับไปทำขั้นที่ 1 ต่อไปอีก

**ขั้นที่ 3** เป็นการลบกฎบางกฎซึ่งทำให้ description length ของ RuleSet เพิ่มขึ้นจึงทำการลบออกจาก RuleSet

### 2.1.6 ประเภทของข้อมูลที่สามารถทำเหมืองข้อมูล

ประเภทของข้อมูลที่สามารถทำเหมืองข้อมูล (Data Mining) มีดังต่อไปนี้

#### (1) Relational Database

เป็นฐานข้อมูลที่จัดเก็บอยู่ในรูปแบบของตาราง โดยในแต่ละตารางจะประกอบไปด้วยแถว และคอลัมน์ ความสัมพันธ์ของข้อมูลทั้งหมดสามารถแสดงได้โดยแบบจำลองฐานข้อมูลเชิงสัมพันธ์ (Entity-Relationship (E-R) Model)

#### (2) Data Warehouses

เป็นการเก็บรวบรวมข้อมูลจากหลายแหล่งมาเก็บไว้ในรูปแบบเดียวกัน และรวบรวมไว้ในที่เดียวกัน ซึ่งสะดวกต่อการนำข้อมูลไปไหนต่อไหน

#### (3) Transactional Database

ประกอบด้วยข้อมูลที่แต่ละทรานแซกชันแทนด้วยเหตุการณ์ในขณะใดขณะหนึ่ง เช่น ใบเสร็จรับเงิน จะเก็บข้อมูลในรูปแบบ ชื่อลูกค้า และรายการสินค้าที่ลูกค้ารายนั้นซื้อ เป็นต้น

#### (4) Advance Database

เป็นฐานข้อมูลที่จัดเก็บในรูปแบบอื่นๆ เช่น ข้อมูลแบบ Object-oriented ข้อมูลที่

เป็นเท็กซ์ไฟล์ ข้อมูลมัลติมีเดีย ข้อมูลในรูปแบบของเว็บ เป็นต้น

### 2.1.7 ความรู้ที่ได้จากการทำเหมืองข้อมูล

ความรู้ที่ได้จากการทำเหมืองข้อมูลมีหลายรูปแบบ ได้แก่

#### (1) กฎเชื่อมโยง (Association Rule)

กฎเชื่อมโยง แสดงความสัมพันธ์ของเหตุการณ์หรือวัตถุ ที่เกิดขึ้นพร้อมกัน กฎเชื่อมโยงมักจะใช้ในการวิเคราะห์ข้อมูลการขายสินค้าโดยเก็บข้อมูลจากระบบ ณ จุดขาย หรือระบบขายสินค้าออนไลน์ และพิจารณาสินค้าที่อยู่ในตะกร้าเดียวกัน หรือผู้ซื้อมักจะซื้อพร้อมกันเช่น ถ้าพบว่าคนที่ซื้อเทปวีดีโอ มักจะซื้อเทปกาต้มน้ำด้วย ร้านค้าอาจจะจัดร้านให้สินค้าทั้งสองอย่างอยู่ใกล้กัน เพื่อเพิ่มยอดขาย หรือ อาจพบว่าคนที่ซื้อหนังสือ เอ แล้วหลังจากนั้นมักจะซื้อหนังสือ บีสามารถนำกฎนี้ไปแนะนำผู้ที่กำลังจะซื้อหนังสือ เอ ได้

#### (2) การแบ่งประเภทข้อมูล (Data Classification)

เพื่อระบุประเภทของวัตถุจากคุณสมบัติของวัตถุ เช่น หากความสัมพันธ์ระหว่างผลการตรวจร่างกายต่างๆ กับการเกิดโรค โดยใช้ข้อมูลผู้ป่วย และการวินิจฉัยของแพทย์ที่เก็บไว้ เพื่อนำมาช่วยวินิจฉัยโรคของผู้ป่วย หรือการวิจัยทางการแพทย์ ในทางธุรกิจจะใช้เพื่อคุณสมบัตินี้ที่จะกำหนดหรือหนีเสีย เพื่อประกอบการพิจารณาการอนุมัติเงินกู้

#### (3) การแบ่งกลุ่มข้อมูล (Data Clustering)

การแบ่งกลุ่มข้อมูล แบ่งข้อมูลที่มีลักษณะคล้ายกันออกเป็นกลุ่ม แบ่งกลุ่มผู้ป่วยที่เป็นโรคเดียวกันตามลักษณะอาการ เพื่อนำไปใช้ประโยชน์ในการวิเคราะห์หาสาเหตุของโรค โดยพิจารณาจากผู้ป่วยที่มีอาการคล้ายคลึงกัน

#### (4) จินตทัศน์ (Visualization)

สร้างภาพคอมพิวเตอร์กราฟิกที่สามารถนำเสนอข้อมูลมากมายอย่างครบถ้วนแทนการใช้ข้อความนำเสนอข้อมูลที่มากมาย เราอาจพบข้อมูลที่ซ่อนเร้นเมื่อดูข้อมูลชุดนั้นด้วยจินตทัศน์ ตัวอย่างเช่น ถ้าเรามีข้อมูลชุดหนึ่งที่มีข้อมูลอยู่จำนวน 100,000 รายการ แต่เราใช้รูปแบบการนำเสนอเป็นข้อความ ผลลัพธ์คงจะไม่เหมาะสม เพราะจะทำให้เราต้องสร้างรายงานจำนวนมหาศาล กล่าวคือ ถ้าเราสามารถใช้ภาพคอมพิวเตอร์กราฟิกมาช่วยในการนำเสนอข้อมูลชุดนี้แทนการใช้ข้อความ ผลลัพธ์คือเราไม่ต้องสร้างรายงานจำนวนมหาศาล การนำเสนอข้อมูลด้วยจินตทัศน์เพียง 1 ภาพ แต่สามารถแทนค่าของชุดข้อมูลจำนวนมหาศาลได้ หรือกล่าวอีกนัยหนึ่งคือ ภาพเพียง 1 ภาพ อาจใช้แทนข้อมูลข่าวสารได้เป็น 100 คำ หรือ 1,000 คำ จึงเป็นแนวคิดของการทำจินตทัศน์ นอกจากนี้เรายังสามารถทำการโต้ตอบกับจินตทัศน์ได้อีกด้วย ซึ่งการนำเสนอข้อมูลในรูปแบบข้อความไม่สามารถโต้ตอบได้ข้อมูลที่ปริมาณมหาศาลนำเสนอให้เห็นในแบบจินตทัศน์ อาจทำให้เราพบข้อมูลที่ซ่อนเร้นอยู่ ยกตัวอย่างเช่น ถ้าเรานำชุดข้อมูล IRIS ซึ่งเป็นข้อมูลตัวเลขของดอกไม้ชนิดหนึ่งจำนวน 150

รายการ นำชุดข้อมูลชุดนี้มาแสดงให้เห็นในแบบจินตทัศน์ โดยวิธี scatter plot เราจะเห็นว่าจุด plot ที่แสดงมีการเกาะกลุ่มกันจำนวน 3 กลุ่ม บางส่วนของกลุ่มที่ 1 Intersection กับบางส่วนของกลุ่มที่ 2 ส่วนกลุ่มที่ 3 ไม่มีส่วนใด Intersection กับกลุ่มที่ 1 และ กลุ่มที่ 2 แสดงให้เห็นว่าชุดข้อมูลดอกไม้ IRIS นี้มีจำนวน 3 กลุ่มที่คล้ายคลึงกัน และบางสมาชิกของกลุ่มแรกมีความคล้ายคลึงกับบางสมาชิกของกลุ่มที่ 2 ส่วนสมาชิกของกลุ่มที่ 3 มีความแตกต่างกับสมาชิกในกลุ่มที่ 1 และกลุ่มที่ 2 อย่างมีนัยสำคัญ

### 2.1.8 ประโยชน์ของการทำเหมืองข้อมูล

การทำเหมืองข้อมูลจำเป็นต้องอาศัยบุคลากรจากหลายฝ่าย และต้องอาศัยความรู้จำนวนมากถึงจะได้รับประโยชน์อย่างแท้จริง เพราะสิ่งที่ได้จากขั้นตอนวิธีเป็นเพียงตัวเลข และข้อมูลที่อาจจะนำไปใช้ประโยชน์ได้ หรือ ใช้ประโยชน์อะไรไม่ได้เลยก็เป็นไปได้ ผู้ที่ศึกษาการทำเหมืองข้อมูลจึงควรมีความรู้รอบด้าน และต้องติดต่อกับทุกๆฝ่าย เพื่อให้เข้าใจถึงขอบเขตของปัญหาโดยแท้จริงก่อน เพื่อให้การทำเหมืองข้อมูลเกิดประโยชน์อย่างแท้จริง ซึ่งสามารถแบ่งได้ดังนี้

- (1) ค้นหาข้อมูลโดยอาศัยเทคโนโลยีของเหมืองข้อมูล
- (2) ใช้สถาปัตยกรรมแบบ Client/Server
- (3) ผู้ใช้ระบบไม่จำเป็นต้องมีทักษะในการเขียน โปรแกรม
- (4) ผู้ใช้ต้องกำหนดขอบเขต และเป้าหมายของระบบให้ชัดเจนเพื่อความรวดเร็ว และถูกต้องตามความต้องการ
- (5) การประมวลผลแบบขนานจะช่วยเพิ่มประสิทธิภาพ และความเร็วในการค้นหา

## 2.2 การสุ่มตัวอย่าง

การวิจัยทางสังคมศาสตร์จำเป็นต้องอาศัยวิธีวิทยาศาสตร์ในการคัดเลือกกลุ่มเป้าหมายที่จะทำการศึกษาซึ่งสามารถทำได้โดยการอาศัยการสุ่มตัวอย่าง (Sampling) การสุ่มตัวอย่างเป็นการคัดเลือกจากประชากรทั้งหมด โดยสุ่มตัวอย่างมาเพียงส่วนหนึ่ง เป็นตัวแทนของประชากรทั้งหมดเพื่อนำมาศึกษา

### 2.2.1 องค์ความรู้ในการสุ่มตัวอย่าง

1. ข้อมูลประชากร (Population) หมายถึง กลุ่มเป้าหมายที่ต้องการศึกษาทั้งหมด ซึ่งอาจจะเป็นคน สัตว์ พืช วัตถุ หรือปรากฏการณ์ต่างๆ เช่น ในการศึกษาความรู้ในการประกอบอาชีพด้านหม่อนไหมของเกษตรกรผู้ปลูกหม่อนเลี้ยงไหมในเขต ภาคอีสานตอนบน ประชากรในที่นี้คือ



เกษตรกร ที่มีภูมิลำเนาอยู่ในเขตจังหวัดต่างๆ ของภาคอีสานตอนบนในการวิจัยเชิงสังคมศาสตร์ ประชากรแบ่งออกได้ 2 ประเภทดังนี้

1.1 ประชากรที่มีจำนวนจำกัด (Finite Population) หมายถึงประชากรที่มีปริมาณซึ่งสามารถนับออกมาเป็นตัวเลขได้ครบถ้วนเช่น ประชากรนิสิต หรือนักศึกษา ของมหาวิทยาลัยทุกแห่ง ประชากรของเกษตรกรในภาคกลาง ฯลฯ

1.2 ประชากรที่มีจำนวนไม่จำกัด (Infinite Population) หมายถึงประชากรที่มีปริมาณซึ่งไม่สามารถนับจำนวนออกมาเป็นตัวเลขได้ครบถ้วน เช่น ประชากรเมล็ดข้าวเหลืองที่จำหน่ายในจังหวัดขอนแก่น ฯลฯ

2. ขนาดตัวอย่าง (Sample Size) ขนาดตัวอย่างต้องมากพอที่จะเป็นตัวแทนได้ วิธีการประมาณขนาดตัวอย่างโดยใช้สูตรของ TARO YAMANE ดังนี้

$$\frac{n = N}{1 + Nd^2}$$

เมื่อ  $n$  = ขนาดของหน่วยตัวอย่างกลุ่มเป้าหมาย  
 $N$  = ประชากรทั้งหมด  
 $D$  = ระดับความมีนัยสำคัญ

ตัวอย่างเช่น  $N = 1,000$  คน  
 $D = 0.05$

แทนค่า  $n = 1000 \frac{1000}{1 + 1000 (0.05)^2}$

$$n = 285.7$$

$$n = 286$$

3. ประเภทและวิธีการสุ่มตัวอย่าง ทฤษฎีการสุ่มตัวอย่าง ได้แบ่งประเภทการสุ่มตัวอย่างออกเป็น 2 ประเภทใหญ่ๆ คือ

1. การสุ่มตัวอย่างในเชิงเป็นไปได้ (Probability Sampling) การสุ่มตัวอย่างแบบนี้เราสามารถกำหนดได้ว่าทุกภาคส่วนของประชากรมีโอกาสได้รับเลือกเป็นตัวอย่างเท่ากัน การสุ่มแบบนี้มีหลายวิธีดังนี้

1.1 การสุ่มตัวอย่างแบบง่าย (Simple Random Sampling) หมายถึง การสุ่มตัวอย่างที่ประชากรทุกภาคส่วนมีโอกาสเท่าเทียมกันที่จะได้รับการคัดเลือกเป็นตัวอย่างโดยวิธีการใช้ ตารางเลขสุ่ม นำจำนวนขนาดตัวอย่างไปสุ่มในตารางสำเร็จรูปที่นักสถิติจัดทำไว้แล้ว เพียงแต่นักวิจัยกำหนดหลักที่จะใช้ว่ามีกี่หลัก และจะนับไปซ้ายขวา ขึ้นบน ลงล่างอย่างไรต้องกำหนดไว้และปฏิบัติอย่างนั้น

ตลอด สุ่มโดยการชี้ตัวเลขเริ่มต้น เมื่อชี้ตรงไหนก็บอกว่าเป็นเลขประจำตัวของประชากรหรือไม่ถ้าไม่ใช่ให้ข้ามไป ทำการคัดเลือกไปเรื่อยๆ จนได้ตามจำนวนที่ต้องการ หรือ โดยวิธีการจกฉลากโดยการเขียนหมายเลขกำกับประชากรตัวอย่าง แต่ละรายการก่อนแล้วจึงจับฉลากขึ้นมา ซึ่งวิธีการจับฉลากอาจใช้ 2 แบบคือ

ก. ไม่สุ่มประชากรที่ถูกสุ่มแล้วขึ้นมาอีก (Simple Random Sampling with out Replacement) คือหยิบแล้วเอาออกได้เลย ไม่ต้องใส่กลับลงไปอีก

ข. สุ่มประชากรที่ถูกสุ่มแล้วขึ้นมาได้อีก (Sample Random Sampling with Replacement) คือ หยิบขึ้นมาแล้วก็ใส่ลงไปใหม่เพื่อให้โอกาสแก่ประชากรทุกหน่วย มีโอกาสถูกเลือกขึ้นมาเท่าเดิม

1.2 การสุ่มตัวอย่างแบบมีระบบ (Systematic Sampling) การสุ่มแบบนี้ นักวิจัยจะต้องอาศัยบัญชีรายชื่อ เกี่ยวกับประชากรกลุ่มเป้าหมาย โดยเลือกตามเลขที่ที่กำหนดไว้ เช่น ประชากรจำนวน 1,000 นักวิจัยต้องการตัวอย่างจำนวน 100 นักวิจัยจะต้องคัดเลือกทุกหน่วยที่ 10 เป็นต้น

1.3 การสุ่มตัวอย่างแบบแบ่งชั้น (Stratified Random Sampling) การสุ่มตัวอย่างแบบนี้ต้องแยกประเภทของประชากรเป็นกลุ่มย่อยหรือชั้นก่อน แล้วจึงสุ่มตัวอย่างแยกกันคนละกลุ่มโดยวิธี Simple Random Sampling หรือ Systematic Sampling ก็ได้ กลุ่มย่อยที่มีลักษณะเป็น Homeogeneous คือมีลักษณะเหมือนกันภายในกลุ่มเช่น การแยกประเภทของประชากรตามสถานการณื เป็นสมาชิกของกลุ่มเกษตรกร

1.4 การสุ่มตัวอย่างแบบกลุ่ม (Cluster Sampling) คือการสุ่มตัวอย่างประชากรโดยแบ่งประชากรออกเป็นกลุ่มๆ ให้แต่ละกลุ่มมีความเป็น Heterogeneous กัน คือมีความแตกต่างกันภายในกลุ่ม เช่น การสุ่มตัวอย่างโดยการแบ่งตามเขตการปกครอง

1.5 การสุ่มตัวอย่างในทุกชั้นตอน (Multi Stage Sampling) เช่น ต้องการจะทำการวิจัย โดยการสุ่มตัวอย่างประชากร โดยทำการสุ่มจังหวัดที่เป็นตัวอย่างก่อน ต่อไปก็สุ่มอำเภอ ตำบล หมู่บ้าน และครัวเรือนที่เป็นตัวอย่างตามลำดับ

2. การสุ่มตัวอย่างในเชิงเป็นไปไม่ได้ (Non-probability Sampling) คือ การสุ่มตัวอย่างโดยไม่อาจกำหนดได้ว่าทุกส่วนของประชากรมีโอกาสได้รับการคัดเลือกโดยเท่ากัน ซึ่งทำให้ไม่สามารถจะคาดคะเนหรือคำนวณหาความผิดพลาดในการสุ่มเลือกตัวอย่างได้ การสุ่มแบบนี้มีหลายวิธีคือ

2.1 การสุ่มตัวอย่างแบบบังเอิญ (Accidental Sampling) เช่น พบใครก็สัมภาษณ์ตามความพอใจของผู้วิจัย เช่น สุ่มนักท่องเที่ยวที่จะเข้าประเทศไทยที่สนามบินดอนเมือง

2.2 การสุ่มตัวอย่างโดยวิธีการจัดสรรโควตา (Quota Sampling) การสุ่มตัวอย่างเหล่านี้ต้องแบ่งกลุ่มของประชากรแล้วจัดสรรโควตาตัวอย่างไปให้แต่ละกลุ่มตามสัดส่วนของปริมาณประชากรในกลุ่มนั้นๆ ที่มีอยู่จากนั้นก็ทำการสุ่มจากแต่ละกลุ่มตามโควตาที่จัดสรร ทั้งนี้เพื่อให้ได้ตัวแทนจากกลุ่มต่างๆ อย่างเหมาะสม เช่น ชาย 80 คน หญิง 80 คน

2.3 การสุ่มตัวอย่างแบบเจาะจง(Purposive Sampling) โดยจะเลือกศึกษาจากประชากรที่มีลักษณะตรงตามวัตถุประสงค์ที่จะศึกษาเช่น เกษตรกรที่ปลูกหม่อน บร.60 เป็นต้น

2.4 การสุ่มตัวอย่างพิจารณาตามความสะดวก (Convenience Sampling) โดยจะเลือกศึกษากลุ่มประชากรที่เห็นว่าง่ายต่อการศึกษา เช่น ไม่อยู่ในแดนของผู้ก่อการร้าย หรือเลือกเฉพาะผู้เป็นสมาชิกของกลุ่มทางการเมือง กลุ่มใดกลุ่มหนึ่ง

### 2.2.2 ปัจจัยที่ทำให้สำเร็จในการสุ่มตัวอย่าง (Key Success Factor)

1. ฐานข้อมูล/ประชากรต้องเป็นปัจจุบัน (Update Population)
2. วิธีการสุ่ม ต้องมีความน่าเชื่อถือ (มีแหล่งที่มาอ้างอิงได้)
3. ขนาดตัวอย่างต้องมีการกระจายตัวและครอบคลุมประชากรเพื่อให้มีความคลาดเคลื่อนน้อยที่สุด

### 2.3 ประเภทของสถิติ

นักคณิตศาสตร์ได้แบ่งสถิติในฐานะที่เป็นศาสตร์ออกเป็นสาขาใหญ่ ๆ 2 สาขาด้วยกัน คือ สถิติเชิงพรรณนา (Descriptive Statistics) และการอนุมานเชิงสถิติ หรือ สถิติเชิงอนุมาน (Inferential Statistics) ซึ่งแต่ละสาขามีรายละเอียดดังนี้

1. **สถิติพรรณนา (Descriptive Statistics)** หมายถึง การบรรยายลักษณะของข้อมูล (Data) ที่ผู้วิจัยเก็บรวบรวมจากประชากรหรือกลุ่มตัวอย่างที่สนใจ ซึ่งอาจจะแสดงในรูป ค่าเฉลี่ย มัธยฐาน ฐานนิยม ร้อยละ ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน เป็นต้น
2. **สถิติเชิงอนุมาน (Inferential Statistics)** หมายถึง สถิติที่ว่าด้วยการวิเคราะห์ข้อมูลที่รวบรวมมาจากกลุ่มตัวอย่าง เพื่ออธิบายสรุปลักษณะบางประการของประชากร โดยมีการนำทฤษฎีความน่าจะเป็นมาประยุกต์ใช้ สถิติสาขานี้ ได้แก่ การประมาณค่าทางสถิติ การทดสอบสมมติฐานทางสถิติ การวิเคราะห์การถดถอยและสหสัมพันธ์ เป็นต้น

### 2.4 ข้อมูล

ข้อมูล(Data) หมายถึง ข้อมูลหรือตัวเลขที่แสดงคุณสมบัติที่ผู้วิจัยต้องการศึกษา เช่น อายุ รายได้ ยอดขาย เป็นต้น

#### 2.4.1 ประเภทของข้อมูล

ข้อมูลที่เกี่ยวข้องกันมา ไม่ว่าจะใช้ในงานวิจัย หรือวัตถุประสงค์อื่นใดก็ตาม ย่อมมีลักษณะแตกต่างกันไป โดยการจำแนกข้อมูล อาจจำแนกตามเกณฑ์ใหญ่ ๆ ได้ 3 ประการ คือ จำแนกตามแหล่งข้อมูล จำแนกตามลักษณะของข้อมูล และจำแนกตามมาตรการวัด เป็นต้น

### จำแนกตามแหล่งข้อมูล

1. **ข้อมูลปฐมภูมิ (Primary Data)** หมายถึง ข้อมูลที่ผู้วิจัยเป็นผู้เก็บรวบรวมข้อมูลที่สนใจเอง โดยที่อาจจะใช้วิธีเก็บแบบสอบถาม แบบสัมภาษณ์ การทดลอง เป็นต้น
2. **ข้อมูลทุติยภูมิ (Secondary Data)** หมายถึง ข้อมูลที่ผู้วิจัยไม่ได้เป็นผู้เก็บรวบรวมข้อมูลที่สนใจเอง โดยนำข้อมูลจากผู้อื่น ๆ เก็บมาใช้ เช่น ข้อมูลเกี่ยวกับการจ้างงานที่กระทรวงแรงงานรวบรวมไว้ เป็นต้น

### จำแนกตามลักษณะของข้อมูล

1. **ข้อมูลเชิงปริมาณ (Quantitative Data)** หมายถึง ข้อมูลที่สามารถแสดงในรูปตัวเลขได้ เช่น น้ำหนัก อายุ คะแนน จำนวนสินค้า งบประมาณ จำนวนพนักงานในบริษัท เป็นต้น  
ข้อมูลเชิงปริมาณยังแบ่งได้ ๒ ประเภท คือ
  - 1.1. **ข้อมูลแบบต่อเนื่อง (Continuous Data)** หมายถึงข้อมูลที่มีค่าต่าง ๆ ทุกค่าต่อเนื่องกัน โดยแสดงได้ทั้งเศษส่วนหรือตัวเลขที่เป็นจำนวนเต็ม เช่น ส่วนสูง น้ำหนัก ความยาวของโต๊ะ
  - 1.2. **ข้อมูลแบบไม่ต่อเนื่อง (Discrete Data)** หมายถึงข้อมูลที่มีค่าเป็นจำนวนเต็มหรือจำนวนนับ เช่น ค่าใช้จ่าย จำนวนสินค้า งบประมาณ จำนวนพนักงานในบริษัท
2. **ข้อมูลเชิงคุณภาพ (Qualitative Data)** หมายถึง ข้อมูลที่ไม่สามารถแสดงในรูปตัวเลขได้ หรือ อาจจะแสดงในรูปตัวเลขได้แต่ไม่สามารถคำนวณในเชิงปริมาณได้ เนื่องจากตัวเลขเหล่านั้นไม่สามารถอธิบายได้ เช่น เพศ สถานภาพ วุฒิการศึกษา เป็นต้น

### จำแนกตามมาตรการวัด

1. **นามบัญญัติ (Nominal Scales)** คือระดับของข้อมูลที่เป็นกาหนดชื่อหรือแบ่งแยกประเภทของสิ่งต่าง ๆ เช่น เบอร์โทรศัพท์ ห้องเรียน อาคารเรียน บ้านเลขที่ เป็นต้น
2. **เรียงลำดับ (Ordinal Scales)** คือระดับของข้อมูลที่สามารถจัดลำดับความสำคัญของข้อมูลตามความแตกต่างได้ เช่น ชอบมาก ชอบปานกลาง ชอบน้อย ไม่ชอบ เป็นต้น
3. **อันตรภาค (Interval Scales)** คือระดับของข้อมูลที่สามารถบอกถึงปริมาณของความแตกต่างของข้อมูลได้ แต่ไม่มีศูนย์แท้ เช่น อุณหภูมิ คะแนนสอบวัดความรู้ ความสูง เป็นต้น
4. **อัตราส่วน (Ratio Scales)** คือระดับของข้อมูลที่เป็นกาบอกถึงปริมาณความแตกต่างของข้อมูลที่มีรายละเอียดมากที่สุดและมีศูนย์แท้ เช่น ระยะทาง น้ำหนัก ความเร็ว เป็นต้น

## 2.5 ค่าเฉลี่ย ( $\bar{x}$ )

$$\text{กรณีข้อมูลที่ไม่ได้แบ่งกลุ่ม} \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{หรือ} \quad \bar{x} = \frac{\sum x}{n}$$

$$\text{กรณีข้อมูลที่แบ่งกลุ่ม} \quad \bar{x} = \frac{\sum_{i=1}^n f_i x_i}{n} \quad \text{หรือ} \quad \bar{x} = \frac{\sum fx}{n}$$

## 2.6 ระบบบริหารความสัมพันธ์ลูกค้า (Customer Relationship Management)

ความสำคัญของระบบบริหารความสัมพันธ์ลูกค้า ช่วยให้องค์กรสามารถเพิ่มความสัมพันธ์อันดีให้กับลูกค้า เพิ่มรายได้ลดค่าใช้จ่ายต่างๆ โดยเฉพาะเรื่องค่าใช้จ่ายในการแสวงหาลูกค้า และเพิ่มความพึงพอใจของลูกค้า ( Customer Satisfaction ) โดยการสร้างกระบวนการทำงานและพัฒนาผลิตภัณฑ์ตามความต้องการของลูกค้า จำนวนผู้ประกอบการ ที่เพิ่มขึ้นอย่างรวดเร็วในปัจจุบันสามารถนำแนวทางการบริหารลูกค้าสัมพันธ์ เช่น แนวทางที่สำคัญในการจัดการระบบการบริหารงาน และสร้างมาตรฐานการทำงานในบริษัท เช่น การรวบรวมการเกี่ยวกับข้อมูลของลูกค้า , การจัดการเกี่ยวกับช่องทางการสื่อสารและการพัฒนาสินค้า และบริการเพื่อสนองตอบสิ่งที่ลูกค้าต้องการประโยชน์ของการบริหารลูกค้าสัมพันธ์ต่อบริษัท

1. การเพิ่มรายได้จากการขาย (Sale Revenue Increase) การมุ่งเน้นการสร้างความสัมพันธ์อันดีกับลูกค้า เพื่อให้ลูกค้าเกิดความภักดีการใช้สินค้าหรือบริการ (Customer Loyalty) การนำหลักการบริหารลูกค้าสัมพันธ์ (CRM) มีรายได้ที่เพิ่มขึ้น ปรับปรุงกระบวนการทำงานในบริษัทลดรายจ่ายในการดำเนินงาน และต้นทุนการหาลูกค้า ใหม่ๆหรือดึงลูกค้ากลับมาใช้สินค้าหรือบริการอีกครั้ง
2. การบริหารของวงจรการทำธุรกิจของลูกค้า (Customer Life Cycle Management)

2.1 การหาลูกค้าใหม่เข้าองค์กร (Customer Acquisition) โดยการสร้างความเด่น (Differentiation) ของสินค้าหรือบริการที่ใหม่ (Innovation) และเสนอความสะดวกสบาย (Convenience) ให้กับลูกค้า

2.2 การเพิ่มความพึงพอใจให้กับลูกค้าเพื่อทำการซื้อสินค้าและบริการ โดยผ่านขั้นตอนการทำงาน ที่กระชับเพื่อการสนองตอบความต้องการของลูกค้าได้รวดเร็วและถูกต้อง และการทำงานที่สนอง ตอบสิ่งที่ลูกค้าต้องการหรือเสนอสิ่งที่ลูกค้าต้องการ โดยผ่านหน่วยงาน ลูกค้าสัมพันธ์ (Customer Service)

2.3 การรักษาลูกค้า (Customer Retention) ให้อยู่กับองค์กรนานที่สุด และการดึงลูกค้าให้กลับมา ใช้สินค้าหรือบริการ โดยฟังความคิดเห็นจากลูกค้าและพนักงานในองค์กร (Listening) รวมถึงการเสนอสินค้าและบริการใหม่ (New Product)

3. การเพิ่มประสิทธิภาพในกระบวนการตัดสินใจ (Improving Integration of Decision Making Process)

การเพิ่มการประสานงานในฝ่ายต่างๆของบริษัท โดยเฉพาะการใช้ระบบฐานข้อมูลของลูกค้าร่วมกัน และผู้บริหารสามารถดึงข้อมูลจากระบบต่างๆมาประกอบการตัดสินใจ เช่น รายละเอียดของลูกค้าที่ติดต่อเข้ามาในฝ่ายลูกค้าสัมพันธ์ (Call Center) รายละเอียดของการจ่ายเงินของลูกค้าจากฝ่ายขาย(Sales) กิจกรรมทางการตลาดที่เสนอให้ลูกค้าแต่ละกลุ่มจากฝ่ายการตลาด(Marketing) และ การควบคุมปริมาณของสินค้าในแต่ละช่วงจากฝ่ายสินค้าคงคลัง (Inventory Control) เป็นต้น

4. การเพิ่มประสิทธิภาพในการดำเนินงาน (Enhanced Operational Efficiency)

การบริหารลูกค้าสัมพันธ์ (CRM) จะช่วยเพิ่มประสิทธิภาพการทำงานของฝ่ายต่างๆ ของบริษัท โดยข้อมูลต่างนั้นได้มาจากช่องทางการสื่อสาร เช่น แฟกซ์ โทรศัพท์ และ อีเมล (Email) เป็นต้น

- ฝ่ายขาย: Sales, Cross-selling และ Up-selling

ระบบเทคโนโลยีสารสนเทศที่เกี่ยวข้องกับฝ่ายขาย เช่น ในการขายสินค้าแบบ Cross-selling และ Up-selling เพิ่มความสามารถในการคาดเดาแนวโน้มการซื้อสินค้าหรือบริการ รวมถึงการใช้ข้อมูลของลูกค้า เช่น ข้อสัญญา (Contract) ระหว่างองค์กรกับลูกค้า ระบบยังช่วยระบุรายละเอียดของสินค้าหรือบริการให้เหมาะสมกับลูกค้าแต่ละราย การเก็บข้อมูลทาง ด้านการขาย และการตรวจสอบสถานะภาพของการส่งสินค้าให้กับลูกค้า

- ฝ่ายการตลาด (Marketing) ระบบการบริหารลูกค้าสัมพันธ์ (CRM) มีส่วนช่วยให้บริษัทสามารถวิเคราะห์แนวโน้มที่ควรจัดจำหน่ายสินค้าผ่านช่องทางการขาย (Sales Channels) ต่างๆ เช่น ตัวแทนการขาย (Sales Representatives) และ ผ่านทางเว็บไซต์ (Website) ระบบการบริหารลูกค้าสัมพันธ์ยังมีบทบาท สำคัญกับช่องทางการสื่อสาร (Communication Channels) เช่น ระบุช่องทางการสื่อสาร ที่เหมาะสมที่สุดสำหรับการขายสินค้าชนิดนั้นหรือลูกค้าแต่ละราย หรือ การระบุ พนักงานที่เหมาะสม ที่สุดในการให้บริการหรือติดต่อกับลูกค้ารายนั้นๆ

- ฝ่ายลูกค้าสัมพันธ์ (Customer Service) และฝ่ายสนับสนุน (Support)

ระบบการบริหารลูกค้าสัมพันธ์ในฝ่ายลูกค้าสัมพันธ์(Customer Service) และฝ่ายสนับสนุน (Support) ที่สำคัญคือด้านการดูแลลูกค้า (Customer Care Service) เช่น ระบบการจัดการเกี่ยวกับข้อมูล รายละเอียดของลูกค้าในองค์กร (Account management) และ ระบบแสดงรายละเอียดของสัญญาระหว่างองค์กรกับลูกค้า (Detail Service Agreement) นอกจากนี้แล้วระบบการจัดการด้าน

อีเมลล์ (Email Management System) ถือว่าเป็นส่วนสำคัญในการสร้างกลยุทธ์ทางการบริหารลูกค้าสัมพันธ์ (CRM) เช่น สามารถย้อนหลังดูอีเมลล์ของลูกค้าในอดีตได้ และระบุผู้แทนฝ่ายขายที่เหมาะสมที่สุดกับลูกค้ารายนั้นได้โดยข้อมูลที่ใช้อาจจะมาจากข้อมูลต่างๆที่ลูกค้าเคยติดต่อกับ

- รายละเอียดของการชำระค่าสินค้าหรือบริการให้กับลูกค้า (Customer Billing)

ธุรกิจสามารถใช้ระบบการบริหารลูกค้าสัมพันธ์ (CRM) ในออกรายละเอียดการจ่ายเงินของลูกค้า (Bill Payment) และที่ผ่านการจ่ายเงินระบบอินเทอร์เน็ต (Electronic Bill) และการให้บริการการตอบข้อสงสัยต่างๆผ่านช่องทางสื่อสารต่างๆ เช่น ในระบบออนไลน์

- การขายและให้บริการในสถานที่ที่ลูกค้าต้องการ (Field Sales and Service)

การบริหารลูกค้าสัมพันธ์ที่เกี่ยวข้องกับการขาย และให้บริการในสถานที่ที่ลูกค้าต้องการ (Field Sales and Service) ทำให้พนักงานสามารถช่วยในการดึงข้อมูลมาใช้ในขณะที่ทำการขายหรือการให้บริการกับลูกค้า โดยสามารถใช้ข้อมูลดังกล่าวร่วมข้อมูลขององค์กรร่วมกันได้ การบริหารลูกค้าสัมพันธ์ยังมีส่วนการจัดการเกี่ยวกับการทำรายงานทางการขาย การสร้างใบเสนอราคาให้กับลูกค้า และเงื่อนไขพิเศษให้กับลูกค้าแต่ละรายแบบอัตโนมัติ การเสนอสินค้าที่มีความพิเศษเฉพาะตามต้องการของลูกค้าแต่ละราย (Customized Products) ระบบสินค้าคงคลัง (Inventory System), ระบบการสั่งซื้อ (Ordering System) การส่งและรับสินค้าหรือบริการ (Logistic System), การจัดตารางให้กับพนักงานที่จะให้บริการ การออกใบแจ้งหนี้และการจัดการระบบโควต้าในการขาย

- กิจกรรมที่สร้างความภักดีและการรักษาลูกค้า (Loyalty และ Retain Program)

การบริหารลูกค้าสัมพันธ์ ที่มีประสิทธิภาพขึ้นอยู่กับความแตกต่างเหล่านี้ตามกลุ่มลูกค้า (Customer Segmentation) เช่น การจำแนกประเภทของลูกค้าออกตามความต้องการของลูกค้า ประวัติส่วนตัวของลูกค้า และประวัติการซื้อ นอกจากนี้ยังสามารถกิจกรรมลูกค้าย้อนหลัง เพื่อบริษัทจะได้นำข้อมูลเหล่านี้ไปวิเคราะห์หาข้อมูลเชิงลึก เช่น ช่องทางการสื่อสารที่เหมาะสมที่สุดของลูกค้าแต่ละราย (Effective Communication Channel) พฤติกรรมการซื้อของลูกค้า (Customer Behavior) และสินค้าที่มีความพิเศษเฉพาะตัว (Customized Product) สำหรับลูกค้าแต่ละราย

5. เพิ่มความรวดเร็วในการให้บริการ (Speed of Service) การใช้หลักการบริหารลูกค้าสัมพันธ์สามารถปรับปรุงกระบวนการทำงานโดยมุ่งเน้นที่การตอบสนองความต้องการของลูกค้า จะต้องรวดเร็วและถูกต้อง โดยเฉพาะการตอบสนองแบบให้บริการ หรือตอบสนองกับลูกค้าทันที (Real Time) เช่น ระบบการส่งสินค้ามีการเชื่อมโยงระบบต่างๆ ทั้งในฝ่ายรับการสั่งซื้อ (Order Fulfillment) ฝ่ายขาย (Sales Department) ฝ่ายบัญชี (Accounting Department) ฝ่ายสินค้าคงคลัง (Inventory) และฝ่ายที่เกี่ยวข้องกับการให้เครดิตกับลูกค้า (Credit Authorization)

6. การรวบรวมรายละเอียดต่างของลูกค้า (Gathering More Comprehensive Customer Profiles) การบริหารลูกค้าสัมพันธ์ได้ช่วยให้เพิ่มประสิทธิภาพการทำงานของฝ่ายต่างๆในบริษัท ได้

มากขึ้น เพราะว่าการบริหารลูกค้าสัมพันธ์ช่วยการจัดการเกี่ยวกับข้อมูลของลูกค้าที่มีอยู่ได้มากขึ้นทำให้ข้อมูลเก็บอย่างเป็นระบบอย่างเชื่อมโยงขึ้นบริษัทสามารถนำฐานข้อมูลนี้มาใช้ในระบบต่างๆ ได้

7. การลดต้นทุนในการขายและการจัดการ (Decrease General Sales and Marketing Administration Costs) การลดลงของต้นทุนการดำเนินงานนั้นมาจากใช้หลักการบริหารลูกค้าสัมพันธ์ เนื่องจากบริษัทมีระบบการจัดการที่เน้นในเรื่องการสร้างความสัมพันธ์อันดีกับลูกค้าเข้าใจความต้องการของลูกค้าและตอบสนองความต้องการของลูกค้าได้มากขึ้น ทำให้บริษัทไม่สูญเสียต้นทุนในการดึงลูกค้ากลับเป็นลูกค้าขององค์กรอีก และตัดกระบวนการที่ไม่จำเป็น และกิจกรรมที่ไม่ก่อให้เกิดรายได้แก่บริษัท

8. การสร้างมูลค่าเพิ่ม ( Value Added) ให้กับลูกค้าในปัจจุบันลูกค้านั้นพยายามแสวงหาความพึงพอใจสูงสุดจาก สินค้าและบริการ สิ่งที่ลูกค้าต้องการจึงไม่ใช่แค่คุณค่า (Value) อีกต่อไป แต่ต้องการคุณค่าเพิ่มที่ทำให้ลูกค้ามีความรู้สึกมากกว่าความพอใจ ซึ่งผู้ประกอบการควรสร้างคุณค่าเพิ่มในตัวสินค้าและบริการ โดยผ่าน Value Chain ทั้งในส่วนของลูกค้า (Supply Chain) และในส่วนของความต้องการของลูกค้า (Demand Chain) เพื่อทำให้เกิดการบูรณาการที่ทำให้เกิดมูลค่าเพิ่มให้กับลูกค้าอย่างครบวงจรทั้งระบบ จากหลายหน่วยงานเข้ามาเกี่ยวข้องทั้งภายในองค์กร และภายนอกองค์กร (Internal and External Organization) นับตั้งแต่ผู้จำหน่ายวัตถุดิบ (Raw Materials Suppliers) กระบวนการข่าวเศรษฐกิจที่เกี่ยวข้องกับวัตถุดิบ (Material Procurement) การออกแบบผลิตภัณฑ์ (Product Designers) การจัดหาอุปกรณ์ชิ้นส่วน(Spare Parts Suppliers) การขาย (Sales) และการตลาด (Marketing) ผู้ที่ทำการจัดจำหน่าย (Distributors) และ หน่วยงานลูกค้าสัมพันธ์ (Contact Center) เป็นต้น

ระบบบริหารความสัมพันธ์ลูกค้า (Customer Relationship Management :CRM ) คือกลยุทธ์การบริหารจัดการอย่างหนึ่ง ซึ่งถูกออกแบบมาเพื่อช่วยองค์กรให้สามารถจัดการกระบวนการต่างๆ ภายในให้ดำเนินการได้อย่างสอดคล้องและตอบสนองต่อความต้องการของลูกค้า เพื่อให้ลูกค้าเกิดความพอใจสูงสุด นำมาซึ่งความภักดีของลูกค้า รายได้ที่เพิ่มขึ้น และการทำกำไรในระยะยาว สำหรับบางคนที่เข้าใจว่า CRM เป็นซอฟต์แวร์ นั้นอาจเป็นเพราะส่วนหนึ่งของกลยุทธ์ CRM จำเป็นต้องอาศัยเทคโนโลยีเป็นเครื่องมือในการเก็บข้อมูลลูกค้า ซึ่งเป็นส่วนสำคัญที่จะทำให้องค์กรสามารถดำเนินงาน เพื่อสร้างความพึงพอใจให้ลูกค้าได้อย่างถูกต้อง มีประสิทธิภาพ และรวดเร็วขึ้น ซึ่งถ้าจะให้เห็นภาพชัดขึ้น ก็คงเปรียบเทียบฐานลูกค้าขององค์กรเป็นเหมือนน้ำที่อยู่ในถัง ถ้ามีรูรั่วที่กั้นถึงน้ำก็จะไหลออก เปรียบเทียบได้กับการที่องค์กรจะต้องสูญเสียลูกค้าออกไปอยู่ตลอดเวลา และ CRM ก็คือเครื่องมือที่จะมาลดขนาดรอยรั่วขององค์กรให้เล็กลง เท่ากับองค์กรได้ลดอัตราการสูญเสียลูกค้าให้ต่ำลงนั่นเอง



## 2.7 งานวิจัยที่เกี่ยวข้อง

- 1) อินเทลลิเวิร์ค (Intelliworks, 2004) เอกสารงานวิจัย Challenges in Marketing and Student Relationship Management in Higher Education เป็นงานวิจัยจากบริษัท Intelliworks ว่าถึงความสำคัญขององค์กรทางธุรกิจ ที่มีการประยุกต์ใช้ระบบ CRM ในการนำพองค์กรไปสู่ผลสำเร็จตามเป้าหมาย และกล่าวถึงการนำแนวคิดเดียวกันมาใช้กับองค์กรทางการศึกษา เพื่อจัดการความสัมพันธ์ของนักศึกษาภายในสถาบันการศึกษา เพื่อเพิ่มศักยภาพทางด้านการควบคุมอย่างมีประสิทธิภาพ โดยมีผลถึงการตอบรับต่อตลาดการศึกษาในองค์กรรวม นอกจากนั้นยังได้อธิบายถึงการใช้เทคโนโลยีการสื่อสารที่เหมาะสม สำหรับนำมาประกอบกับการสร้างระบบบริการการศึกษาให้กับนักศึกษา ซึ่งท้ายที่สุดได้ยกตัวอย่างการนำระบบเว็บมาใช้ เพื่อลดต้นทุนในการใช้จ่าย และเพิ่มประสิทธิภาพความรวดเร็วในการสื่อสาร อีกทั้งยังตอบรับกับการขยายตัวในอนาคตได้เป็นอย่างดี ด้วยการอิงหลักการของการใช้ผู้เรียนเป็นศูนย์กลาง ในการออกแบบและการจัดสร้าง
- 2) เนจ เชียค (Naj Shaik 2005) เอกสารงานวิจัย Service Center: A SRM Strategy to Promote Student Retention เป็นงานวิจัยที่กล่าวถึงการสร้างระบบศูนย์การบริการ เพื่อสร้างกลไกในการบริการนักศึกษา ด้วยยุทธวิธีทาง SRM โดยผลของงานวิจัยทำให้ได้มาซึ่ง ระบบการบริการนักศึกษา ผ่านระบบเว็บ โดยมีการเก็บรักษาข้อมูล ทั้งในส่วนประวัตินักศึกษา เจ้าหน้าที่ หลักสูตร และข้อมูลบริหารจัดการ เพื่อให้บริการนักศึกษาแบบเบ็ดเสร็จ วัตถุประสงค์หลักของงานชิ้นนี้คือการสร้าง แอปพลิเคชันบริการนักศึกษา เพื่อใช้งานในการจัดการ เพื่อลดความซ้ำซ้อนของข้อมูล และลดค่าใช้จ่ายที่ไม่จำเป็นออกไป แต่ยังคงและเสริมประสิทธิภาพการทำงาน ความรวดเร็ว และความถูกต้องของคุณภาพข้อมูล
- 3) อะแลน พอล (Alan Paul, 2008) เอกสารงานวิจัยชื่อ A generic student lifecycle relationship management system เป็นเอกสารที่อธิบายถึงปัญหาต่างๆ ที่ก่อให้เกิดความยุ่งยากในการบริหาร และจัดการวงจรของกรนักศึกษา ในสถาบันการศึกษาวางจร ทั้งที่มีเทคโนโลยีเพื่อก่อให้เกิดผลสำเร็จอยู่มากมาย นอกจากนั้นยังอธิบายถึงข้อมูลที่สามารถสนับสนุนการบริหารงานนักศึกษาอยู่มากมาย แต่อย่างไรก็ตามภาพและมุมมองของการผสมผสานสิ่งต่างๆก็ยังคงมีปัญหา และความซับซ้อนแฝงอยู่ภายใน งานวิจัยชิ้นนี้นำปัญหาตัวอย่างที่เกิดขึ้นจากความผิดพลาดการดำเนินงาน มาจัดสร้างเป็นโมเดล ด้วยแนวคิดในเรื่องการตัดสินใจ (Decision Making) การตรวจสอบ (Monitoring) ที่มีส่วนสำคัญในการสร้างแบบจำลองให้มีประสิทธิภาพ โดยอ้างอิงโมเดลที่ใช้อยู่ในระบบธุรกิจที่ได้รับผลสำเร็จ นำมาเทียบเคียงและเปรียบเทียบ อีกทั้งยังมีข้อเสนอแนะและตัวอย่างต่างๆ

- 4) **ชุตินา อุตมะมุณี, 2552** งานวิจัยการพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจอย่างชาญฉลาดแบบอัตโนมัติ สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา The Development of Autonomous Intelligent Decision Support Systems for Managing The Study Plan of Students in Higher Education เป็นงานวิจัยที่อธิบายการวิเคราะห์เชิงเปรียบเทียบด้วยอัลกอริทึมสำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา และพัฒนาตัวแบบการตัดสินใจอย่างชาญฉลาดแบบอัตโนมัติ ด้วยการใช้เทคนิคการทำเหมืองข้อมูลซึ่งอาศัยการประยุกต์การเรียนรู้แบบเบย์ (Bayesian Learning) ร่วมกับขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm) และกฎความสัมพันธ์ โดยมีวัตถุประสงค์เพื่อช่วยพัฒนาหลักสูตร และพัฒนาคุณภาพของนักศึกษา รวมถึงเพื่อผลักดันให้นักศึกษามีความกระตือรือร้นมากยิ่งขึ้น
- 5) **ปรีชา ยามันสะบีดิน, บุญเสริม กิจศิริกุล และ ประสงค์ ปราณิตพลกรัง ,2549** งานวิจัย เรื่อง การบริหารความสัมพันธ์กับนักศึกษาในสถาบันระดับอุดมศึกษาโดยการประยุกต์การทำเหมืองข้อมูล(Student Relationship Management in Higher Education Institutions with Application of Data Mining) งานวิจัยนี้เป็นการนำแนวความคิดของการทำเหมืองข้อมูลมาออกแบบและพัฒนาใช้สำหรับการบริหารความสัมพันธ์ (Student Relationship Management, SRM) โดยการใช้ข้อมูลทางสถิติมาใช้ในการทำเหมืองข้อมูล ซึ่งเป็นกรรมวิธีที่ช่วยให้สามารถวิเคราะห์ถึงพฤติกรรมของนักศึกษา ตลอดจนคาดคะเนถึงแนวโน้มของความเป็นไปได้ในอนาคตอันใกล้ ทั้งนี้เพื่อเป็นเหตุผลประกอบการตัดสินใจในการจัดการกับปัญหาของนักศึกษาที่มีแนวโน้มที่จะถูกคัดชื่อออก ดังนั้นการบริหารความสัมพันธ์กับนักศึกษาดังกล่าว สามารถช่วยในการรักษาสัมพันธ์ภาพที่ดีระหว่างนักศึกษาระดับอุดมศึกษา ผลลัพธ์ที่ได้คือข้อมูลทางสถิติที่เป็นประโยชน์และรูปแบบของพฤติกรรมนักศึกษา เพื่อนำไปใช้กำหนดภารกิจวัตถุประสงค์และวางแผนกลยุทธ์ของมหาวิทยาลัย
- 6) **รศ.ดร.กฤษณะ ไวยมัย และกัลยาภัทร อุดมสิน, 2550** งานวิจัยมีชื่อ คลังข้อมูลอัจฉริยะด้านการศึกษา (Intelligence Data Warehouse For Education) งานวิจัยนี้ได้จัดทำขึ้นโดยมีวัตถุประสงค์เพื่อนำข้อมูลต่างๆ ในสถานศึกษามาทำการวิเคราะห์ให้เกิดประโยชน์แก่สถานศึกษา และในภาพรวมของประเทศนับเป็นเรื่องที่สำคัญมาก งานวิจัยคลังข้อมูลอัจฉริยะด้านการศึกษามีวัตถุประสงค์ที่สำคัญ 4 ประการ คือ ประการแรกเพื่อศึกษาวางรูปแบบและออกแบบคลังข้อมูลอัจฉริยะด้านการศึกษา ประการที่สองเพื่อจัดทำคลังข้อมูลของสถานศึกษา ประการที่สามเพื่อกำหนดดัชนีชี้วัดทางการศึกษาทั้ง 4 ด้าน และประการที่สี่เพื่อกำหนดรหัสมาตรฐานในการแลกเปลี่ยนข้อมูลด้านการศึกษา ซึ่งถ้าโครงการนี้สำเร็จ ล่วงตามวัตถุประสงค์ ก็จะมีส่วนช่วยให้สถานศึกษาของประเทศไทยได้มีเครื่องมือที่ใช้ในการจัดการข้อมูลต่างๆ ที่เกี่ยวข้องกัสถานศึกษา เพื่อที่จะสามารถนำข้อมูลเหล่านี้ไปใช้ใน

การวิเคราะห์เพื่อให้เกิดประโยชน์แก่สถานศึกษา มีศูนย์กลางข้อมูลด้านการศึกษาเพื่อใช้ในการแลกเปลี่ยนข้อมูล สามารถเปรียบเทียบข้อมูลระหว่างสถานศึกษาได้ ซึ่งจะก่อให้เกิดประโยชน์แก่สถานศึกษา และในภาพรวมของประเทศ

- 7) **Hengqing Tong, Xiaochuan Lu ,2009** งานวิจัยชื่อ Consumption Psychoanalysis and Customer Relationship Management Based on Association Rules Mining การวิเคราะห์พฤติกรรมผู้บริโภคของลูกค้า ถือเป็นสิ่งสำคัญอย่างมากสำหรับองค์กรที่ต้องการปรับปรุงการผลิตและพัฒนาตัวเองขึ้นไป และดัชนีความพึงพอใจของลูกค้าก็ถือเป็นส่วนสำคัญรูปแบบหนึ่งในแง่การบริโภค ความต้องการในการสร้างสมการเพื่อบ่งชี้ถึงดัชนีความพึงพอใจ ยังคงต้องมีการปรับปรุงและพัฒนาต่อไป วิธีการปรับปรุงอัลกอริทึมสมการโครงสร้าง คือการกำหนดค่าเริ่มต้นเพื่อประมวลผลซ้ำ และหาอัลกอริทึมในการเพิ่มอัตราการใช้ความถูกต้องให้ได้มากที่สุด และเนื่องจากระบบการจัดการความสัมพันธ์ของลูกค้าเป็นปัจจัยในการระบบบริหารงานระดับใหญ่จึงมีการประยุกต์ใช้วิธีการใช้เหมืองข้อมูล ร่วมกับวิธีการหาความสัมพันธ์เป็นหากฎ (Rules) หรือรูปแบบ (Pattern) เพื่อเป็นพื้นฐานสำหรับขั้นตอนการตัดสินใจ(Decision) งานวิจัยนี้อธิบายอัลกอริทึมที่ถูกพัฒนาขึ้นเพื่อใช้ในซอฟต์แวร์ที่พร้อมสำหรับการใช้งานได้อย่างเป็นรูปธรรม
- 8) **Vasile Paul Bresfelean ,2007** งานวิจัยชื่อ "Analysis and Predictions on Students' Behavior Using Decision Trees in Weka Environment Decision Trees เป็นหลักการพื้นฐานเพื่อให้สามารถแยกแยะข้อมูล เพื่อการเรียนรู้โดยให้ผลลัพธ์ที่มีประสิทธิภาพซึ่งเป็นวิธีที่ใช้กับเหมืองข้อมูล สำหรับศึกษาข้อมูลและสร้างรูปแบบค้นไม่ร่วมกับกฎเกณฑ์เพื่อช่วยคำนวณและคาดคะเนผลลัพธ์ได้สิ่งหนึ่งที่ถูกทำทลาย คือการค้นหาคำรู้ และรูปแบบของเหมืองข้อมูล เพื่อใช้พัฒนาความสามารถในการวิเคราะห์ข้อมูลจากโมเดลทั่วไป และเพื่อความเป็นเลิศในมหาวิทยาลัยที่ต้องพร้อมรองรับ ในการปรับเปลี่ยนความต้องการของหน่วยต่างๆ คุณภาพของระบบบริหารอาชีพผู้เชี่ยวชาญ และเทคโนโลยีขั้นสูง งานวิจัยนี้แสดงวิธีการใช้เครื่องมือในการวิเคราะห์ ร่วมกับข้อมูลที่ได้จากการสำรวจจากผู้เรียนในมหาวิทยาลัย เพื่อวัตถุประสงค์ในการคาดคะเนทางเลือกแนวทางการศึกษาในระดับต่อไป หลังจากสำเร็จการศึกษาออกไป

## บทที่ 3

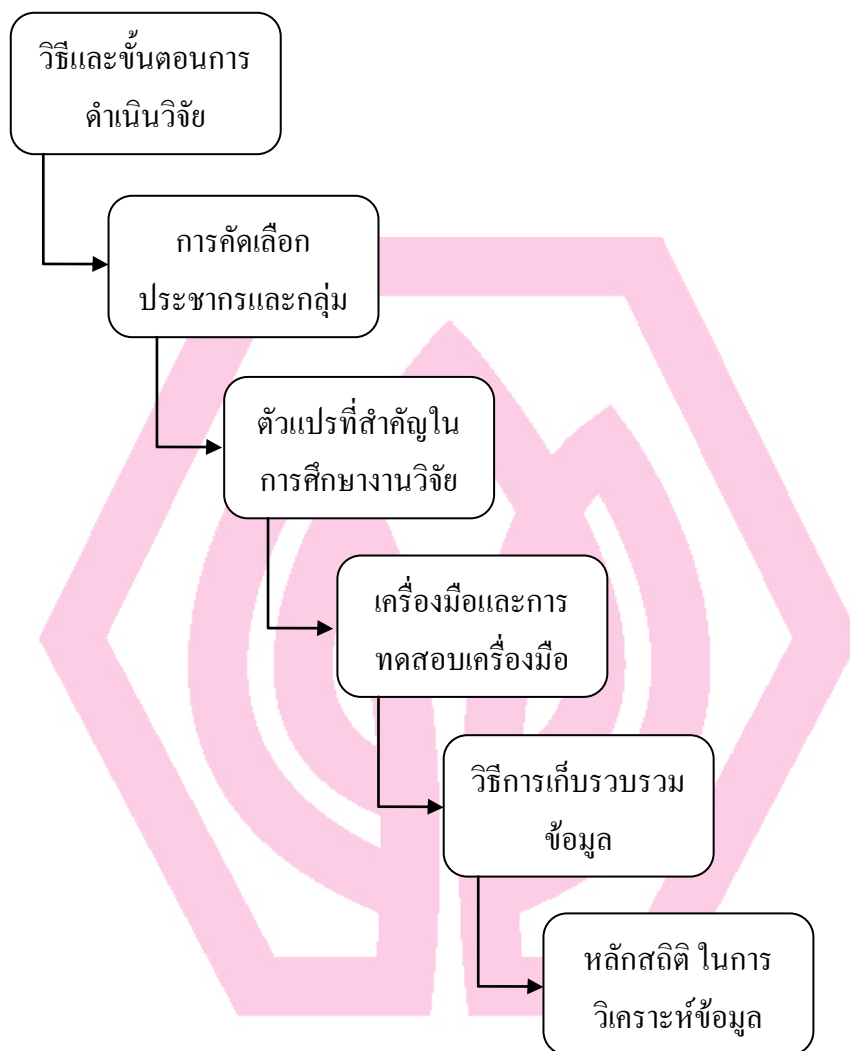
### ระเบียบวิธีการวิจัย

การดำเนินการศึกษาวิจัยเรื่องการสร้างโมเดลการจัดการระบบนักศึกษาสัมพันธ์ โดยใช้เทคนิคเหมืองข้อมูล (The Development of Student Relationship Management Model using Data Mining Technique) มีวัตถุประสงค์ เพื่อศึกษาข้อมูลเกี่ยวกับหลักสูตร อาจารย์ การบริการ ต่างๆ ของมหาวิทยาลัยต่างๆ ในการศึกษาวิจัยครั้งนี้ มีระเบียบวิธีการวิจัย ดังนี้

- 3.1 วิธีการและขั้นตอนการดำเนินการวิจัย
- 3.2 การคัดเลือกประชากรและกลุ่มตัวอย่าง
- 3.3 ตัวแปรที่สำคัญในการศึกษางานวิจัย
- 3.4 เครื่องมือและการทดสอบเครื่องมือที่นำมาใช้ในการวิจัย
- 3.5 วิธีการเก็บรวบรวมข้อมูลสำหรับทำงานวิจัย
- 3.6 หลักสถิติที่ใช้ในการวิเคราะห์ข้อมูล
- 3.7 วิธีการวิจัยโดยใช้เทคนิคเหมืองข้อมูล

#### 3.1 วิธีการและขั้นตอนการดำเนินการวิจัย

ในเนื้อหาของงานที่ทางผู้วิจัยได้ดำเนินการศึกษา เป็นรูปแบบของงานวิจัยเชิงเหตุผล (Casual Research) ซึ่งเป็นประเภทของงานวิจัยที่ระบุความสัมพันธ์ระหว่างตัวแปรเหตุและผล (Cause and Effect Relationship) โดยมีการเก็บข้อมูลจากกลุ่มตัวอย่างจากประชากร จากแบบสอบถาม ผ่านวิธีและขั้นตอน จากนั้นใช้หลักทางสถิติและพยากรณ์วิเคราะห์ข้อมูลจากงานวิจัย โดยการทำวิจัยผู้ทำการวิจัยได้วางขั้นตอนและรูปแบบการทำงาน ดังต่อไปนี้



รูปที่ 3.1 แสดงขั้นตอนดำเนินการวิจัย

### 3.2 การคัดเลือกประชากรและกลุ่มตัวอย่าง

ประชากร (Population) ที่ทางผู้วิจัยกำหนดเป็นกลุ่มที่ใช้ในการทำการวิจัยในครั้งนี้ เป็นนักศึกษาจากสถาบันการศึกษาที่เปิดสอนในระดับปริญญาตรี จำนวนทั้งสิ้น 2,731 คนจากสถาบันการศึกษาทั้ง 7 แห่ง การสุ่มตัวอย่าง (Sampling) สำหรับการเลือกตัวอย่างจากแต่ละกลุ่มผู้วิจัยใช้วิธีการตัดสินใจ ให้นำหน่วยตัวอย่างเป็นไปตามวัตถุประสงค์ของงานวิจัย จึงเลือกวิธีเลือกสุ่มตัวอย่างแบบบังเอิญ โดยได้มีการแยกกลุ่มประชากรแยกตามสถาบันการศึกษาดังนี้

1. มหาวิทยาลัยบูรพา	จำนวน 397 คน
2. มหาวิทยาลัยกรุงเทพ	จำนวน 394 คน
3. มหาวิทยาลัยหอการค้าไทย	จำนวน 392 คน
4. มหาวิทยาลัยราชภัฏราชนครินทร์	จำนวน 374 คน
5. มหาวิทยาลัยเกษตรศาสตร์	จำนวน 397 คน
6. มหาวิทยาลัยศรีปทุม	จำนวน 393 คน
7. มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ	จำนวน 384 คน

### 3.3 ตัวแปรที่ใช้ในการวิจัย

สำหรับตัวแปรที่ทางผู้วิจัยสนใจ และสัมพันธ์กับลักษณะของประชากรมีดังต่อไปนี้

1. เจตคติต่อพฤติกรรมในความต้องการต่อระบบนักศึกษาสัมพันธ์
2. ความคล้ายตามต่อสิ่งอ้างอิง
3. ความเหมาะสมของหลักสูตร
4. คุณสมบัติที่จำเป็นของอาจารย์
5. สิ่งอำนวยความสะดวกในระบบการบริการ

### 3.4 เครื่องมือและการทดสอบเครื่องมือที่นำมาใช้ในการวิจัย

สำหรับการสร้างเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลของการศึกษาวิจัย มีวัตถุประสงค์เพื่อทดสอบและวิเคราะห์ระหว่างตัวแปรต้นและตัวแปรตาม โดยเก็บข้อมูลของนักศึกษาจากมหาวิทยาลัยรัฐบาลและเอกชนต่างๆ ซึ่งข้อมูลระดับปฐมภูมินั้น ได้สร้างมาจากแบบสอบถาม โดยแบ่งเนื้อหาของแบบสอบถามเป็น 3 ส่วนดังนี้ คือ

ตอนที่ 1 เป็นแบบสอบถามเกี่ยวกับข้อมูลทั่วไปของประชากรที่ถูกสุ่มตัวอย่าง ในเรื่องต่างๆไปเกี่ยวกับภูมิลำเนา สถานภาพ การศึกษาก่อนเข้าสู่ระบบมหาวิทยาลัย ความคาดหวังที่ต้องการได้รับการศึกษา ว่ามีรูปแบบเป็นอย่างไร เพื่อวัดตัวแปรเกี่ยวกับคุณลักษณะของนักศึกษา โดยใช้วิธี check list

ตอนที่ 2 เป็นแบบสอบถามเกี่ยวกับตัวแปรที่ใช้ในการศึกษา ทั้งหมด 3 ตัวแปร ซึ่งประกอบด้วย

2.1 ด้านเนื้อหาของหลักสูตร โดยใช้วิธีการวัดแบบ Rating Scale ตามหลักการของ Likert Scale แบบ 5 ระดับ โดยมีจำนวนคำถามทั้งหมด 12 ข้อ

2.2 ด้านคุณสมบัติของอาจารย์ โดยใช้วิธีการวัดแบบ Rating Scale ตามหลักการ

ของ Likert Scale แบบ 5 ระดับ โดยมีจำนวนคำถามทั้งหมด 11 ข้อ

2.3 ด้านสิ่งอำนวยความสะดวกต่างๆ เช่น วัสดุการศึกษา ตำรา อาคาร สถานที่และระบบการให้บริการ โดยใช้วิธีการวัดแบบ Rating Scale ตามหลักการของ Likert Scale แบบ 5 ระดับ โดยมีจำนวนคำถามทั้งหมด 12 ข้อ

ตอนที่ 3 เป็นการให้ผู้ตอบแบบสอบถามแสดงความคิดเห็นหรือข้อเสนอแนะคะแนนเกี่ยวกับการแปลความหมายเป็นดังตารางที่ 3.1

ตารางที่ 3.1 แสดงระดับความคิดเห็น คะแนนเฉลี่ยและระดับการแปลความหมาย

ระดับความคิดเห็น	ค่าคะแนน	ค่าเฉลี่ยคะแนน	ระดับการแปลความหมาย
มากที่สุด	5	4.21-5.00	ระดับมากที่สุด
มาก	4	3.41-4.20	ระดับมาก
ปานกลาง	3	2.61-3.40	ระดับปานกลาง
น้อย	2	1.81-2.60	ระดับน้อย
น้อยที่สุด	1	1.00-1.80	ระดับน้อยที่สุด

เพื่อพิสูจน์ว่าเครื่องมือที่ใช้ในการทำวิจัย มีประสิทธิภาพในการนำไปใช้งาน จึงจำเป็นต้องมีการทดสอบเครื่องมือก่อนการนำไปใช้งาน ซึ่งทางผู้วิจัยได้กำหนดมาตรวัดด้วย 2 กระบวนการดังต่อไปนี้

1. การหาค่าความเที่ยงตรงเชิงเนื้อหา (Content Validity) ด้วยการให้ผู้เชี่ยวชาญพิจารณา ด้วยการหาค่าดัชนีความสอดคล้องของเครื่องมือระหว่างข้อคำถามกับวัตถุประสงค์ (Index of Item Objective Congruence: IOC) จากแบบสอบถามการสร้างโมเดลระบบนักศึกษาสัมพันธ์ โดยการหาค่าดัชนีความสอดคล้องของแบบจำลองดัชนีวัดผลความพอใจในระบบนักศึกษาสัมพันธ์ตามวิธีของโรวินลลิและแฮมบิลตัน โดยแบ่งคะแนนออกเป็น 3 ระดับดังนี้

แน่ใจว่ามีความสอดคล้องหรือวัดได้	มีระดับคะแนนเท่ากับ	1
ไม่แน่ใจว่ามีความสอดคล้องหรือวัดได้	มีระดับคะแนนเท่ากับ	0
แน่ใจว่าไม่มีความสอดคล้องหรือวัดได้	มีระดับคะแนนเท่ากับ	-1

หลังจากนั้นนำแบบสอบถามการสร้างโมเดลระบบนักศึกษาสัมพันธ์ ส่งให้ผู้ทรงคุณวุฒิที่มีความเชี่ยวชาญทางด้านการตรวจสอบแบบสอบถาม และนำมาหาค่าความสอดคล้องที่ได้รับมาใช้งาน โดยใช้สูตร

$$IOC = \frac{\sum R}{N}$$

เมื่อ

IOC หมายถึง ดัชนีความสอดคล้องมีค่าระหว่าง -1 ถึง 1

$\sum R$  หมายถึง ผลรวมของการพิจารณาของผู้เชี่ยวชาญ

N หมายถึง จำนวนผู้เชี่ยวชาญ

โดยนำผลลัพธ์ที่ได้จากสมการที่ 3.1 มาเปรียบเทียบกับเกณฑ์มาตรฐานความเที่ยงตรงของเคลเดนดอล, กรูเบอร์ และจอร์น สัน ซึ่งได้เสนอมาตรการการประเมินผลดัชนีความสอดคล้องของแบบสอบถามการสร้างโมเดลระบบนักศึกษาสัมพันธ์ดังนี้

ค่าเฉลี่ย 0.00-0.49	ความสอดคล้องของแบบสอบถามอยู่ในเกณฑ์ต่ำ
ค่าเฉลี่ย 0.50-0.69	ความสอดคล้องของแบบสอบถามอยู่ในเกณฑ์ยอมรับ
ค่าเฉลี่ย 0.70-0.79	ความสอดคล้องของแบบสอบถามอยู่ในเกณฑ์ดี
ค่าเฉลี่ย 0.80-1.00	ความสอดคล้องของแบบสอบถามอยู่ในเกณฑ์ดีมาก

2. การหาค่าความเชื่อมั่นของเครื่องมือและข้อมูลที่ใช้ในการวิจัย ด้วยวิธีหาค่าสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach alpha coefficient) ซึ่งเป็นวิธีการที่ถูกนำไปใช้กับเครื่องมือในทุกๆ ข้อคำถามที่มีคำตอบไม่ใช่แค่ 0 และ 1 เท่านั้น แต่เป็นคำตอบประเภทเรียงลำดับโดยใช้สูตร

$$\alpha = \left( \frac{k}{k-1} \right) \left( 1 - \frac{\sum s_i^2}{s_t^2} \right)$$

เมื่อ

$\alpha$  หมายถึง ค่าความเชื่อมั่นของเครื่องมือ

$k$  หมายถึง จำนวนคำถาม

$s_i^2$  หมายถึง ความแปรปรวนของคะแนนแต่ละข้อ

$s_t^2$  หมายถึง ความแปรปรวนของคะแนนทั้งหมด

การคำนวณทุกๆ ข้อคำถาม จะมีค่าความเชื่อมั่นใกล้เคียงกัน และค่าความเชื่อมั่นที่ได้จะขึ้นอยู่กับค่าความสัมพันธ์ระหว่างข้อคำถามและจำนวนคำถาม หากมีความสัมพันธ์กันสูงหรือจำนวนข้อคำถามมีมาก ค่าของความเชื่อมั่นก็จะมีค่าสูง หากค่าความเชื่อมั่นของแบบสอบถามต่ำควรที่จะเพิ่มคำถามให้มากขึ้น แต่จะต้องมีความสัมพันธ์ในทางเดียวกันกับ



ข้อคำถามเดิม แต่ข้อคำถามที่เพิ่มขึ้นมีความสัมพันธ์ในทางตรงกันข้ามกับข้อคำถามเดิม จะทำให้ค่าความเชื่อมั่นลดลงได้ ค่าเชื่อมั่นควรจะไม่น้อยกว่า 0.70

ทางผู้วิจัยได้นำแบบสอบถามที่ได้ปรับปรุงแล้วไปทำการทดสอบ (Pilot Test) เพื่อแสดงค่าความเชื่อมั่น กับกลุ่มตัวอย่างที่ใช้กลุ่มเป้าหมาย จำนวน 40 คน แล้วนำข้อมูลมาวิเคราะห์หาค่าความเชื่อมั่นด้วยวิธีครอนบาค อัลฟา ดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 แสดงค่าความเชื่อมั่นด้วยวิธีของ Cronbach's Alpha

ตัวแปร	จำนวนคำถาม	Cronbach's Alpha
ด้านความเหมาะสมของหลักสูตร	12 ข้อ	0.88
ด้านคุณสมบัติของอาจารย์	11 ข้อ	0.92
ด้านสิ่งอำนวยความสะดวกของระบบการให้บริการ	12 ข้อ	0.82

ซึ่งผลลัพธ์ที่ได้มีค่า 0.88 แสดงให้เห็นว่า ข้อคำถามที่เป็นตัววัดตัวแปรแต่ละตัวมีความเหมาะสม และเชื่อถือได้ ทำให้เชื่อมั่นได้ว่าสามารถนำเครื่องมือนี้ไปใช้เพื่อทำการวิจัยกับกลุ่มตัวอย่างได้ ซึ่งจะได้นำไปเก็บข้อมูลจากกลุ่มตัวอย่างจำนวน 2,731 ชุด

### 3.5 วิธีการเก็บรวบรวมข้อมูลสำหรับทำงานวิจัย

จากข้อมูลข้างต้นการทำงานวิจัยครั้งนี้ ซึ่งได้กำหนดประชากร กลุ่มตัวอย่าง ซึ่งก็คือนักศึกษาระดับปริญญาตรีที่กำลังศึกษาอยู่ในสถาบันการศึกษาในมหาวิทยาลัยรัฐบาลและเอกชนจำนวน 7 แห่ง โดยการใช้เครื่องมือคือแบบสอบถาม ที่ถูกสร้างขึ้นมาตามหลักการและผ่านขั้นตอนการทดสอบจนแน่ใจว่าเป็นเครื่องมือที่มีประสิทธิภาพในการทำงานวิจัย วิธีการรวบรวมข้อมูลโดยการแจกแบบสอบถามไปยังมหาวิทยาลัยทั้งรัฐและเอกชนทั้งหมด จำนวน 5000 ชุด และได้แบบสอบส่งกลับมาจำนวน 2800 ชุด จากนั้นจึงนำข้อมูลที่ได้จากการเก็บรวบรวมข้อมูล นำไปใช้ในขั้นตอนการวิเคราะห์ต่อไป

### 3.6 หลักสถิติที่ใช้ในการวิเคราะห์ข้อมูล

หลังจากที่ได้เก็บข้อมูลจากกลุ่มตัวอย่างเรียบร้อยแล้ว ขั้นตอนต่อไปซึ่งถือเป็นขั้นตอนที่สำคัญ ก็คือการนำข้อมูลที่ได้ไปใช้วิเคราะห์ตามหลักสถิติ เพื่อให้ได้ข้อมูลสรุปสำคัญเพื่อนำไปใช้งานได้ ในโอกาสต่อไป ทางผู้วิจัยได้มีการใช้ระบบคอมพิวเตอร์มาช่วยในการวิเคราะห์

โดยนำเอาโปรแกรมสำเร็จรูปที่มีความสามารถในการวิเคราะห์ข้อมูลตามหลักการทางสถิติต่างๆ โดยมีหลักการของการวิเคราะห์ดังต่อไปนี้

1. ค่าร้อยละ (Percentage) นำมาใช้ในการอธิบายข้อมูลทั่วไปสำหรับกลุ่มตัวอย่างจากประชากรผู้ตอบแบบสอบถาม
2. ค่าเฉลี่ย ( $\bar{x}$ ) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation :SD) นำมาใช้ในการวิเคราะห์และอธิบายค่า และความหมายตัวแปรแต่ละตัว

### 3.7 วิธีการวิจัยโดยใช้เทคนิคเหมืองข้อมูล

ในงานวิจัยนี้ได้ใช้หลักการ CRIPS-DM ซึ่งเป็นกระบวนการมาตรฐานในการทำงานทางด้านเหมืองข้อมูล ประกอบด้วย 6 ขั้นตอนดังรายละเอียดต่อไปนี้

#### 1. Business Understanding

วัตถุประสงค์ของงานวิจัยนี้ คือ เพื่อศึกษาปัจจัยที่มีผลต่อความพึงพอใจของนักศึกษาในด้านต่างๆ

#### 2. Data Understanding

ข้อมูลที่ใช้ในงานวิจัยนี้ได้จากการจัดทำแบบสอบถาม จำนวน 2,731 โดยลักษณะแบบสอบถามเป็นดังที่ได้กล่าวมาแล้วในข้างต้น

#### 3. Data Preparation

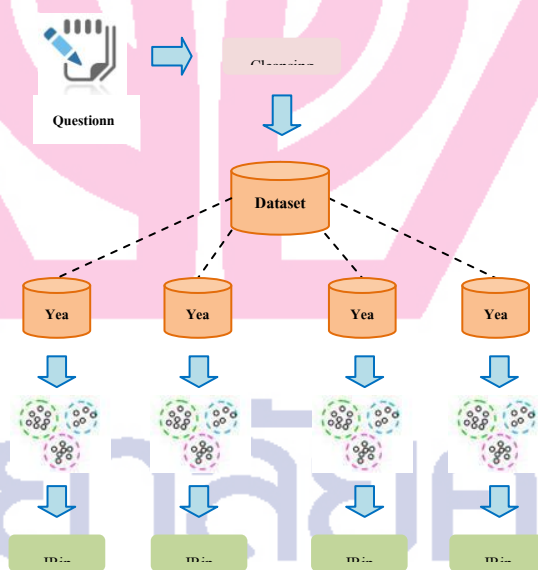
ข้อมูลที่ได้จากการตอบแบบสอบถามของนักศึกษาจำนวนทั้งหมด 2,731 คน จะถูกนำมากรองข้อมูลที่ผิดพลาดขาดหายไป (Missing Value) ก่อน โดยผู้วิจัยได้ทำการคัดกรองเฉพาะแบบสอบถามที่นักศึกษามีการทำแบบสอบถามครบทุกหัวข้อสอบถาม พบว่ามีข้อมูลจำนวน 283 แบบสอบถามที่นักศึกษาตอบแบบสอบถามไม่ครบทุกหัวข้อ ดังนั้นจึงเหลือข้อมูลจำนวน 2,517 แบบสอบถามที่จะถูกใช้ในงานวิจัยต่อไป โดยแบ่งจำนวนข้อมูลตามชั้นปีของนักศึกษาได้ดังตารางที่ 3.3

ตารางที่ 3.3 แสดงการจำนวนนักศึกษาที่แบ่งกลุ่มในแต่ละชั้นปี

ชั้นปี	ชั้นปีที่ 1	ชั้นปีที่ 2	ชั้นปีที่ 3	ชั้นปีที่ 4
จำนวนข้อมูล	974	637	567	339

#### 4. Modeling

ในงานวิจัยนี้มุ่งเน้นวิเคราะห์ผลสำรวจความพึงพอใจของนักศึกษาในหัวข้อต่างๆ เพื่อใช้ในการปรับปรุงคุณภาพการเรียนการสอน อุปกรณ์หรือสื่อการสอน รวมทั้งเครื่องอำนวยความสะดวกเพื่อส่งเสริมการศึกษาให้มีคุณภาพมากยิ่งขึ้น โดยแยกพิจารณาตามชั้นปีของนักศึกษา จากนั้นจึงใช้เทคนิคการจัดกลุ่มข้อมูล (Clustering) เพื่อช่วยในการหาความสัมพันธ์ของกลุ่มนักศึกษาที่มีความพึงพอใจในแต่ละหัวข้อคล้ายๆ กัน จากนั้นผู้วิจัยจึงทำการวิเคราะห์ลงรายละเอียดว่ากลุ่มนักศึกษาลักษณะใดที่ควรให้ความสำคัญ เพื่อทำการปรับปรุงคุณภาพตามหัวข้อที่นักศึกษาในกลุ่มนั้นต้องการได้อย่างถูกต้อง นอกจากนี้ผู้วิจัยยังได้ประยุกต์ใช้เทคนิคการจำแนกประเภท JRip เพื่อช่วยหาความสัมพันธ์ของแต่ละหัวข้อแบบสอบถามในแต่ละกลุ่มนักศึกษา โดยผลลัพธ์ที่ได้จากการใช้เทคนิค JRip นั้น สามารถช่วยวิเคราะห์การปรับปรุงคุณภาพเฉพาะด้านของหัวข้อแบบสอบถามที่อาจมีความสัมพันธ์และส่งผลกระทบต่อกันได้ สำหรับกระบวนการวิจัยแสดงดังรูปที่ 3.2



รูปที่ 3.2 แสดงขั้นตอนวิจัยโดยใช้เทคนิคเหมืองข้อมูล

ในขั้นตอนการจัดกลุ่มข้อมูลนั้น ผู้วิจัยได้ทำการจัดกลุ่มข้อมูล โดยกำหนดค่า  $K=3$  ถึง  $K=6$  จากนั้นจึงทำการเลือกจำนวนกลุ่มข้อมูลที่เหมาะสม โดยผู้วิจัยพิจารณาจากค่าทางสถิติต่างๆ ได้แก่ ค่า Sum of Square Error (SSE) ค่า Between/Within และค่า T-Test (ดูวิธีการคำนวณจากหัวข้อที่ 5)

## 5. Evaluation

ในส่วนนี้จะกล่าวถึงการวัดคุณภาพของข้อมูลที่ได้ทำการจัดกลุ่ม โดยพิจารณาจากค่า Sum of Square Error (SSE) และค่า Between/Within โดยมีสูตรการคำนวณดังนี้

### 5.1 Sum of Squared Errors (SSE)

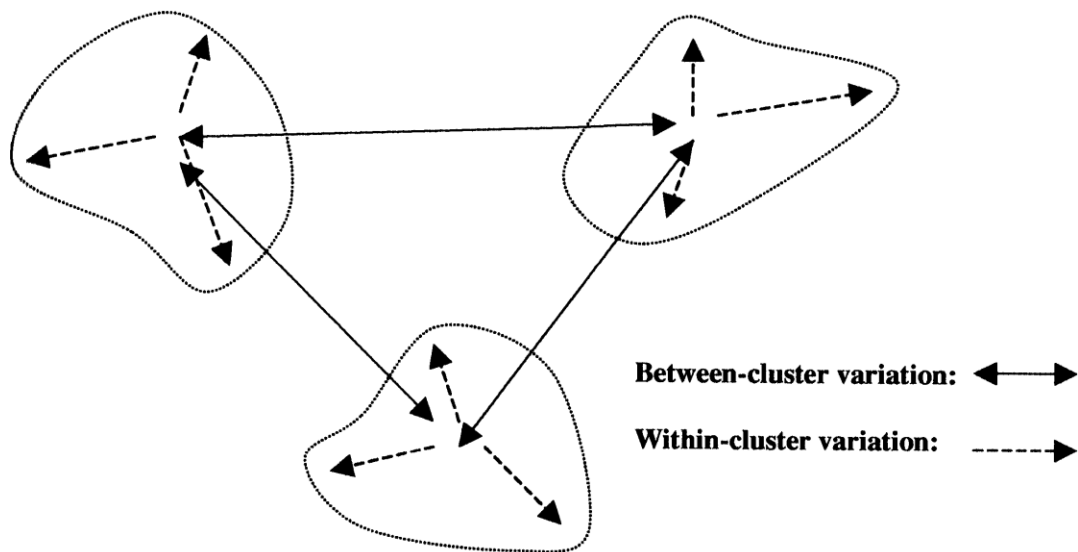
วิธีทั่วไปที่ใช้ในการวัดคุณภาพของการจัดกลุ่ม คือการหาค่าผลรวมความผิดพลาด (Error) ระหว่างจุดข้อมูลใดๆ ในกลุ่มกับจุดข้อมูลศูนย์กลางประจำกลุ่ม (Centroid of Cluster) ซึ่งค่าดังกล่าวแสดงถึงการกระจายตัวของข้อมูล โดยหากค่า Error มีค่าสูง แสดงว่าข้อมูลภายในกลุ่มมีการกระจายตัวมาก ในทางตรงกันข้ามหากพบว่า Error มีค่าต่ำ แสดงว่าข้อมูลที่ถูกจัดกลุ่มนั้น เกาะกลุ่มกันมาก หรือกล่าวอีกนัยหนึ่งคือมีความคล้ายกันหรือเหมือนกันมาก สำหรับฟังก์ชันการคำนวณค่า SSE แสดงดังสมการ

$$SSE = \sum_{i=1}^k \sum_{p \in C_i} d(p, m_i)^2$$

โดยที่  $d(p, m_i)$  คือฟังก์ชันคำนวณระยะห่างของข้อมูล  $p$  คือ จุดข้อมูลใดๆ ในกลุ่ม  $C_i$   $m_i$  คือ จุดข้อมูลศูนย์กลางประจำกลุ่ม (Centroid of Cluster)

### 5.2 Between-cluster variation and Within-cluster variation

เนื่องจากการจัดกลุ่มข้อมูลนั้น ต้องการให้ข้อมูลที่ถูกจัดอยู่ในกลุ่มเดียวกันมีลักษณะคล้ายกันหรือมีการกระจายตัวน้อย และข้อมูลที่อยู่ต่างกลุ่มกันมีความต่างของข้อมูลมาก ดังนั้นวิธีหนึ่งที่นิยมใช้ในการวัดคุณภาพคลัสเตอร์คือ การพิจารณาสัดส่วนระหว่างค่าความแตกต่างระหว่างกลุ่ม (BCV: Between-cluster variation) และค่าความแตกต่างภายในกลุ่ม (WCV: Within-cluster variation) โดยการจัดกลุ่มที่ดีควรมีค่าสัดส่วนดังกล่าวมาก ดังรูปที่ 3.3



รูปที่ 3.3 ค่าความแตกต่างระหว่างกลุ่ม (BCV) ควรมีค่ามาก  
และค่าความแตกต่างภายในกลุ่ม (WCV) ควรมีค่าน้อย

สำหรับสมการการคำนวณค่าสัดส่วนระหว่าง BCV และ WCV แสดงดังนี้

$$\frac{BCV}{WCV} = \frac{d(m_1, m_2)}{SSE}$$

โดยที่  $d(m_1, m_2)$  คือ ฟังก์ชันคำนวณระยะห่างของข้อมูล

$m_1$  และ  $m_2$  คือ จุดข้อมูลศูนย์กลางประจำกลุ่ม (Centroid of Cluster) ของคลัสเตอร์ที่พิจารณา

### 5.3 T-Test

ในการเลือกจำนวนกลุ่มข้อมูลสำหรับเทคนิค K-Means นั้น ในงานวิจัยนี้ได้ทำการปรับเปลี่ยนค่า  $k$  เพื่อทำการพิจารณาจำนวนกลุ่มที่เหมาะสม โดยใช้วิธีการสถิติช่วยในการพิจารณาร่วมกับวิธีการในหัวข้อ 5.1-5.2 โดยใช้การพิจารณาค่า T-Test เพื่อวิเคราะห์ความแตกต่างของข้อมูลในแต่ละกลุ่มว่ามีความแตกต่างอย่างมีนัยสำคัญทางหลักสถิติหรือไม่ การคำนวณ T-Test แสดงดังสมการ

TTEST(array1,array2,tails,type)

โดยที่ Array1 คือ data set ของข้อมูลชุดที่ 1

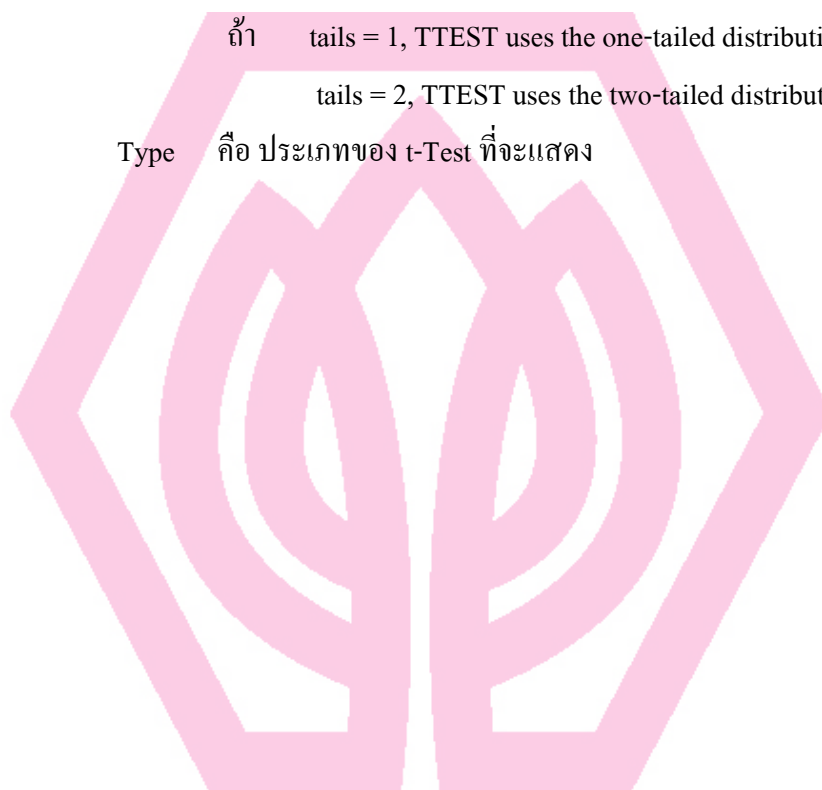
Array2 คือ data set ของข้อมูลชุดที่ 2

Tails คือ การระบุจำนวนการกระจายของ tails

ถ้า tails = 1, TTEST uses the one-tailed distribution.

tails = 2, TTEST uses the two-tailed distribution.)

Type คือ ประเภทของ t-Test ที่จะแสดง



มหาวิทยาลัยศรีปทุม  
SRIPATUM UNIVERSITY

## บทที่ 4

### ผลการวิเคราะห์ข้อมูล

สำหรับวิธีการและขั้นตอนในการทำงานของงานวิจัยชิ้นนี้ ผู้วิจัยได้วางขั้นตอนการทำงานออกเป็นส่วนๆ โดยหลังจากที่ได้อธิบายถึงวิธีการประยุกต์ใช้เทคนิคเหมืองข้อมูลซึ่งมีเนื้อหาอยู่ในบทก่อนหน้านี้ ซึ่งขั้นตอนวิธีการดำเนินงานวิจัย ได้มีการนำเอาเทคนิคเหมืองข้อมูลมาใช้ เพื่อค้นหาปัจจัยที่มีอิทธิพลต่อความพึงพอใจของนักศึกษาในด้านต่างๆ ที่สนับสนุนให้เกิดความพอใจในการศึกษาระดับอุดมศึกษา โดยเป็นการรวบรวมข้อมูลจากแบบสอบถามจากกลุ่มตัวอย่างที่ได้คัดเลือกไว้แล้ว เพื่อนำผลที่ได้ไปใช้ออกแบบหลักเกณฑ์เพื่อดำเนินนโยบายต่างๆ ให้สัมฤทธิ์ผลในด้านความพอใจในการศึกษา รวมถึงเป็นข้อมูลเบื้องต้นสำหรับฝ่ายบริหารในสถาบันการศึกษาในการใช้ประกอบในการตัดสินใจด้านการบริหารการศึกษาและกำหนดนโยบายของสถาบันต่อไป การดำเนินการผู้วิจัยได้แบ่งขั้นตอนการทำงานออกเป็น 5 ขั้นตอนหลักๆ ดังคือ

- 4.1 การเก็บรวบรวมข้อมูลและคัดเลือกข้อมูล (Data Selection)
- 4.2 การเตรียมข้อมูล (Data Preprocessing)
- 4.3 การแปลงรูปแบบข้อมูล (Data Transformation)
- 4.4 การแบ่งกลุ่มข้อมูล (Data Clustering)
- 4.5 การแปลผล (Interpretation and Evaluations)

#### 4.1 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูล เป็นขั้นตอนแรก หลังจากการออกแบบประเมินผลเสร็จ ก็เป็นขั้นตอนของการนำเอาแบบประเมินผลดังกล่าว มาใช้งาน โดยการแจกแบบสอบถาม เพื่อการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่าง ซึ่งก็คือนักศึกษาในมหาวิทยาลัยต่างๆ ทางผู้วิจัยเลือกใช้วิธีขอความร่วมมือกับสถาบันการศึกษาตามรายชื่อสถานศึกษาที่ได้กำหนดไว้และสุ่มแจกแบบสอบถามตามมหาวิทยาลัยต่างๆ การเตรียมสำเนาเอกสารแบบสอบถามตามจำนวนตัวอย่างที่ต้องการเก็บข้อมูลจากสถาบันการศึกษาจำนวน 7 แห่ง ประกอบด้วย

- |                                 |           |     |
|---------------------------------|-----------|-----|
| 1. มหาวิทยาลัยศรีปทุม           | จำนวน 393 | ชุด |
| 2. มหาวิทยาลัยกรุงเทพ           | จำนวน 394 | ชุด |
| 3. มหาวิทยาลัยเกษตรศาสตร์       | จำนวน 396 | ชุด |
| 4. มหาวิทยาลัยหัวเฉียว          | จำนวน 383 | ชุด |
| 5. มหาวิทยาลัยราชภัฏราชนครินทร์ | จำนวน 373 | ชุด |

6. มหาวิทยาลัยหอการค้าไทย	จำนวน 392	ชุด
7. มหาวิทยาลัยบูรพา	จำนวน 396	ชุด

จำนวนตัวอย่างที่ต้องการได้รับคืนจากการเก็บข้อมูล รวมทั้งสิ้น 2,727 ชุด ซึ่งผู้วิจัยได้แจกแบบสอบถามทั้งหมด 5,000 ชุด

วิธีการกระจายแบบสอบถามตามจำนวนที่กล่าวมาแล้วข้างต้น ผู้วิจัยได้ทำการแจกแบบสอบถามไปยังกลุ่มนักศึกษาแต่ละมหาวิทยาลัย ซึ่งในขณะที่เก็บข้อมูลต้องวางตัวเจ้าหน้าที่ตามความเหมาะสมและตามช่วงเวลา นอกจากงานแจกแบบสอบถามจัดเก็บแบบสอบถามแล้วผู้วิจัยได้ทำการศึกษาและทำความเข้าใจในตัวแบบสอบถาม เพื่ออธิบายข้อคำถามในกรณีที่ผู้ตอบแบบสอบถามมีความสงสัย หรือต้องการคำอธิบายความหมายของคำถามเพิ่มเติมการเก็บรวบรวมข้อมูลด้วยแบบสอบถามนี้ทางผู้วิจัยเริ่มดำเนินงานตั้งแต่ 1 มกราคม 2554 จนถึง 1 มีนาคม 2554 เป็นเวลา 2 เดือน ได้จำนวนกลุ่มตัวอย่างทั้งสิ้น 2,727 ชุด จากนั้นได้นำแบบสอบถามทั้งหมดรวบรวมไว้เพื่อใช้ในขั้นตอนการทำงานในขั้นตอนต่อไป

#### 4.2 การเตรียมข้อมูล

หลังจากที่เก็บรวบรวมแบบสอบถามจากการเก็บข้อมูลตามมหาวิทยาลัยต่างๆ เรียบร้อยแล้ว ถือว่าเป็นข้อมูลดิบในรูปแบบเอกสาร ซึ่งข้อมูลดิบเหล่านี้จำเป็นต้องถูกนำมาจัดเตรียมข้อมูลก่อน เนื่องจากแบบสอบถามบางส่วน (จากผู้ตอบแบบสอบถาม) อาจจะไม่สมบูรณ์ที่จะนำข้อมูลไปวิเคราะห์ ในการนำไปใช้งานสำหรับขั้นตอนต่อไป ทางผู้วิจัยจะต้องมีการคัดแยกแบบสอบถามที่มีสภาพไม่สมบูรณ์ ตัวอย่างเช่น แบบสอบถามที่ไม่ได้มีการบันทึกคำตอบใดๆเลย หรือแบบสอบถามที่อยู่ในสภาพเสียหายฉีกขาดหรือตกหล่นในบางหน้า จะเหลือแบบสอบถามที่นักศึกษากรอกครบถ้วนทุกข้อ จำนวน 2,517 ชุดเท่านั้น จากนั้นขั้นตอนต่อไปจะต้องแปลงข้อมูลที่อยู่ในแบบสอบถามให้เป็นข้อมูลที่อยู่ในคอมพิวเตอร์ ซึ่งก็คือการแปลงข้อมูลดิบที่อยู่ในรูปของเอกสาร ให้เป็นข้อมูลดิจิทัลที่อยู่ในไฟล์คอมพิวเตอร์ เพื่อพร้อมสำหรับการประมวลผลในขั้นตอนต่อไป วิธีที่ผู้วิจัยเลือกใช้เพื่อนำข้อมูลเหล่านี้เข้าสู่ขบวนการทางคอมพิวเตอร์

ในขั้นตอนนี้เป็นขั้นแรกของการนำข้อมูลจากแบบสอบถามสู่ข้อมูลทางคอมพิวเตอร์ ผู้วิจัยได้เลือกใช้โปรแกรมไมโครซอฟต์เอ็กเซล (Microsoft Excel) ในการบันทึกข้อมูลค่าตัวเลขคำตอบจากแบบสอบถาม ในแต่ละข้อของแต่ละกลุ่มตัวอย่างลงในแต่ละเซลล์ โดย



ได้ทำการบันทึกข้อมูลค่าคำตอบถูกแทนด้วยตัวเลขในแต่ละข้อ แบบสอบถามในตอนที 1 ซึ่งเป็นคำถามที่เป็นการสอบถามข้อมูลเกี่ยวกับตัวแปรต้นตามรายละเอียดดังต่อไปนี้

#### สถาบันการศึกษา(Institute)

- |         |                              |
|---------|------------------------------|
| คำตอบ 1 | มหาวิทยาลัยศรีปทุม           |
| 2       | มหาวิทยาลัยกรุงเทพ           |
| 3       | มหาวิทยาลัยเกษตรศาสตร์       |
| 4       | มหาวิทยาลัยหัวเฉียว          |
| 5       | มหาวิทยาลัยราชภัฏราชนครินทร์ |
| 6       | มหาวิทยาลัยหอการค้าไทย       |
| 7       | มหาวิทยาลัยบูรพา             |

#### วิทยาเขต(institute\_branch)

- |         |            |
|---------|------------|
| คำตอบ 1 | บางเขน     |
| 2       | บางเขน     |
| 3       | รังสิต     |
| 4       | กล้วยน้ำไท |
| 5       | บางพลี     |
| 6       | บางคล้า    |
| 7       | วิภาวดี    |
| 8       | บางแสน     |
| 9       | จันทบุรี   |

#### คณะ (institute\_faculty)

- |         |                   |
|---------|-------------------|
| คำตอบ 1 | บัญชี             |
| 2       | ศิลปศาสตร์        |
| 3       | บริหารธุรกิจ      |
| 4       | เทคโนโลยีสารสนเทศ |
| 5       | นิเทศศาสตร์       |
| 6       | สถาปัตยกรรมศาสตร์ |

- 7 นิติศาสตร์
- 8 วิศวกรรมศาสตร์
- 9 เศรษฐศาสตร์
- 10 ดิจิทัลมีเดีย
- 11 วิทยาศาสตร์และเทคโนโลยี
- 12 มนุษย์ศาสตร์
- 13 สาธารณสุขศาสตร์และสิ่งแวดล้อม
- 14 พยาบาลศาสตร์
- 15 ภาษาและวัฒนธรรมจีน
- 16 เกษตรศาสตร์
- 17 เทคนิคการแพทย์
- 18 ครุศาสตร์
- 19 มนุษย์ศาสตร์และสังคมศาสตร์
- 20 วิทยาการจัดการ
- 21 เทคโนโลยีอุตสาหกรรม
- 22 วิทยาการสารสนเทศ
- 23 วิทยาศาสตร์
- 24 ศึกษาศาสตร์
- 25 วิทยาศาสตร์และศิลปศาสตร์
- 26 การจัดการและการท่องเที่ยว
- 27 ศิลปกรรมศาสตร์
- 28 รัฐศาสตร์และนิติศาสตร์

สาขาวิชา (institute\_program)

- |         |                                 |
|---------|---------------------------------|
| คำตอบ 1 | การบัญชี                        |
| 2       | การจัดการโรงแรมและการท่องเที่ยว |
| 3       | การตลาด                         |
| 4       | คอมพิวเตอร์ธุรกิจ               |
| 5       | โฆษณา                           |
| 6       | การจัดการ                       |

- 7 วิทยุกระจายเสียง
- 8 วิศวกรรมคอมพิวเตอร์
- 9 ออกแบบภายใน
- 10 การเงินและการธนาคาร
- 11 นิติศาสตร์
- 12 การจัดการทรัพยากรมนุษย์
- 13 การจัดการธุรกิจค้าปลีกและเฟรนไชส์
- 14 ภาษาอังกฤษสื่อสารธุรกิจ
- 15 การจัดการลอจิสติกส์และห่วงโซ่อุปทาน
- 16 เศรษฐศาสตร์ธุรกิจ
- 17 เทคโนโลยีและการสื่อสาร
- 18 การประชาสัมพันธ์
- 19 เกมและแอนิเมชัน
- 20 วิศวกรรมโยธา
- 21 วิศวกรรมไฟฟ้า
- 22 วิศวกรรมหุ่นยนต์
- 23 บริหารทั่วไป
- 24 การจัดการธุรกิจระหว่างประเทศ
- 25 การจัดการธุรกิจสมัยใหม่
- 26 เทคโนโลยีสารสนเทศ
- 27 วิศวกรรมซอฟต์แวร์
- 28 วิทยาการคอมพิวเตอร์
- 29 วิทยุกระจายเสียงและวิทยุโทรทัศน์
- 30 ศิลปะการแสดง
- 31 การสื่อสารแบรนด์
- 32 วารสารศาสตร์
- 33 การจัดการธุรกิจสายการบิน
- 34 ภาษาจีนเพื่อการท่องเที่ยวและโรงแรม
- 35 การสอนคณิตศาสตร์
- 36 การสอนวิทยาศาสตร์

- 37 ไฟฟ้าเครื่องการผลิต  
38 วิศวกรรมเครื่องกล  
39 วิศวกรรมสิ่งแวดล้อม  
40 วิศวกรรมโยธาและทรัพยากรน้ำ  
41 วิศวกรรมเคมี  
42 วิศวกรรมสำรวจและสารสนเทศภูมิศาสตร์  
43 วิศวกรรมไฟฟ้าเครื่องกลการผลิต  
44 วิศวกรรมอุตสาหการ  
45 วิศวกรรมวัสดุ  
46 วิศวกรรมโยธาและสารสนเทศภูมิศาสตร์  
47 วิศวกรรมการบินและอวกาศ  
48 Software & Knowledge  
49 คณิตศาสตร์  
50 การจัดการสิ่งแวดล้อม  
51 สถิติ  
52 พื้นพิภพ  
53 การจัดการการผลิต  
54 การเงิน  
55 ภาษาอังกฤษ  
56 ปรัชญาและศาสนา  
57 ภาษต่างประเทศ  
58 ภาษาเยอรมัน  
59 ภาษาไทย  
60 การโรงแรม  
61 ศิลปะชีพ  
62 การท่องเที่ยว  
63 รัฐศาสตร์  
64 เอเชียตะวันออกเฉียงใต้  
65 บริหารรัฐกิจ  
66 จิตวิทยา

- 67 ภูมิศาสตร์
- 68 เทคโนโลยีชีวภาพ
- 69 พัฒนาผลิตภัณฑ์
- 70 เทคโนโลยีการบรรจุและพัสดุ
- 71 วิทยาศาสตร์และเทคโนโลยีการอาหาร
- 72 ประมง
- 73 วนศาสตร์ทั่วไป
- 74 เทคโนโลยีเชื้อและกระดาษ
- 75 วิทยาศาสตร์และเทคโนโลยีทางไม้
- 76 วิทยาศาสตร์และเทคโนโลยีทางน้ำ
- 77 การจัดการโรงพยาบาล
- 78 อาชีวอนามัยและความปลอดภัย
- 79 อนามัยสิ่งแวดล้อม
- 80 ภาษาจีน
- 81 เทคโนโลยีศึกษา
- 82 สังคมศึกษา
- 83 วิทยาศาสตร์ทั่วไป
- 84 การศึกษาปฐมวัย
- 85 ฟิสิกส์
- 86 รัฐประศาสนศาสตร์
- 87 การพัฒนาชุมชน
- 88 นานุกรณวิทยาคศิลป์ไทย
- 89 ศิลปกรรม
- 90 ภาษาญี่ปุ่น
- 91 ดนตรีสากล
- 92 ดนตรีไทย
- 93 นิเทศ
- 94 เคมี่
- 95 สาธารณสุขชุมชน
- 96 วิทยาการความปลอดภัย

- 97 ชีววิทยา  
98 การเงินและการบัญชี  
99 นิเทศศาสตร์  
100 อุตสาหกรรมท่องเที่ยว  
101 ออกแบบผลิตภัณฑ์  
102 E-Office  
103 การจัดการอุตสาหกรรม  
104 เศรษฐศาสตร์ระหว่างประเทศ  
105 ทฤษฎีเศรษฐศาสตร์  
106 ภาษาไทยเพื่อการสื่อสาร  
107 วิศวกรรมโลจิสติกส์  
108 วิศวกรรมคอมพิวเตอร์และมัลติมีเดีย  
109 จุลชีววิทยา  
110 ชีวเคมี  
111 ระบบสารสนเทศทางธุรกิจ  
112 การสอนนาฏยสังคีต  
113 การสอนภาษาอังกฤษ  
114 การสอนภาษาไทย  
115 สังคมศึกษาศาสนาและวัฒนธรรม  
116 การสอนศิลปะ  
117 การสอนฟิสิกส์  
118 การสอนสุขศึกษาและพลศึกษา  
119 การสอนชีววิทยา  
120 ภาษาเกาหลี  
121 ภาษาฝรั่งเศส  
122 เศรษฐศาสตร์  
123 ดนตรีและการแสดง  
124 เซรามิกส์  
125 นิเทศศิลป์

เพศ (gender)

- คำตอบ 1 ชาย  
2 หญิง

ภูมิลำเนาที่อยู่ (homeland)

- คำตอบ 1 ภาคกลาง  
2 ภาคเหนือ  
3 ภาคใต้  
4 ภาคตะวันออก  
5 ภาคตะวันตก  
6 ภาคตะวันออกเฉียงเหนือ

ปัจจุบันเป็นนักศึกษา (study\_year)

- คำตอบ 1 ระดับชั้นปีที่ 1  
2 ระดับชั้นปีที่ 2  
3 ระดับชั้นปีที่ 3  
4 ระดับชั้นปีที่ 4  
5 ระดับชั้นปีที่ 5 ขึ้นไป

ท่านกำลังศึกษาใน (study\_course)

- คำตอบ 1 ภาคนกดิ  
2 ภาคนกดิต่อเนื่อง  
3 ภาคนกดิ  
4 ภาคนเสาร้-อาทิตย้  
5 อื่นๆ

ปัจจุบันนักศึกษามีเครื่องคอมพิวเตอร์หรือไม่ (computer\_owner)

- คำตอบ 1 มี  
2 ไม่มี

ลักษณะสถานศึกษาก่อนเข้าศึกษาระดับปริญญาตรีจากที่ใด (pre\_study\_kind)

- คำตอบ 1     รัฐบาล  
          2     เอกชน

วุฒิการศึกษาจากสถานศึกษาก่อนเข้าศึกษาระดับปริญญาตรี (pre\_study\_level)

- คำตอบ 1     มัธยมศึกษาตอนปลายหรือเทียบเท่า  
          2     ปวช.  
          3     ปวส.  
          4     อื่นๆ

เกรดเฉลี่ยสะสมที่ท่านสำเร็จก่อนเข้าศึกษาระดับปริญญาตรี (pre\_study\_gpa)

- คำตอบ 1     เกรดเฉลี่ย 1.00-1.50  
          2     เกรดเฉลี่ย 1.51-2.00  
          3     เกรดเฉลี่ย 2.01-2.50  
          4     เกรดเฉลี่ย 2.51-3.00  
          5     เกรดเฉลี่ย 3.01-3.50  
          6     เกรดเฉลี่ย 3.51-4.00

ท่านเข้ามาศึกษาระดับปริญญาตรีด้วยวิธี (entrance\_by)

- คำตอบ 1     Entrance / O-Net / A-Net  
          2     รับสมัครตรง  
          3     โควต้าพิเศษ  
          4     ทุนการศึกษา

เกรดเฉลี่ยสะสมในขณะนี้ (current\_gpa)

- คำตอบ 1     เกรดเฉลี่ย 1.00-1.50  
          2     เกรดเฉลี่ย 1.51-2.00  
          3     เกรดเฉลี่ย 2.01-2.50  
          4     เกรดเฉลี่ย 2.51-3.00  
          5     เกรดเฉลี่ย 3.01-3.50



6 เกรดเฉลี่ย 3.51-4.00

ท่านคิดว่าเข้ามาศึกษาในสาขาวิชานี้เพื่ออะไร (study\_target)

- คำตอบ 1      มีโอกาสได้งานสูง  
 2                หลักสูตรน่าเรียน  
 3                เรียนตามเพื่อน  
 4                ชื่อเสียงสถาบัน  
 5                เรียนตามเพื่อน  
 6                บิดามารดาชี้แนะ  
 7                เทคโนโลยีทันสมัย  
 8                อื่นๆ

ท่านเคยถูกพักการเรียนหรือไม่ (study\_time\_out)

- คำตอบ 1      เคย  
 2                ไม่เคย

ท่านกู้เงินกองทุนเรียนหรือไม่ (study\_loan)

- คำตอบ 1      กู้  
 2                ไม่กู้

สำหรับคำตอบของคำถามของแต่ละข้อในส่วนที่ 2 ซึ่งเป็นข้อมูลในส่วนของตัวเองแปรตามที่ใช้ประเมินผลความพึงพอใจในหัวข้อต่างๆ ดังตารางที่ 4.1 ค่าคะแนนระดับความคิดเห็น คำตอบถูกแทนด้วยตัวเลข 1 – 5 ตามระดับความพอใจจากน้อยไปมาก เนื่องจากคำถามต้องการจะวัดระดับความคิดเห็นของผู้ตอบแบบถามเป็นหลัก ดังนั้นการบันทึกค่าตัวเลขคำตอบ จึงใช้ระดับค่าความพึงพอใจเป็นข้อมูลในการบันทึก ดังนี้คือ

- 1                แทนความพึงพอใจน้อยที่สุด  
 2                แทนความพึงพอใจน้อย  
 3                แทนความพึงพอใจ  
 4                แทนความพึงพอใจมาก  
 5                แทนความพึงพอใจมากที่สุด

ตารางที่ 4.1 หัวข้อต่างๆ ที่ใช้สอบถามความพึงพอใจของนักศึกษา

ภาพรวมเนื้อหาหลักสูตร	
Course_1	ส่งเสริมให้นัก.มีความรู้ความเข้าใจในหลักสูตร
Course_2	เป็นความรู้ที่นัก.สามารถนำไปใช้ในการประกอบอาชีพได้
Course_3	เป็นพื้นฐานในการศึกษาชั้นสูงหรือศึกษาหาความรู้ใหม่
Course_4	เรียงลำดับการจัดการเรียนการสอนตามความยากง่าย
Course_5	ได้ฝึกทักษะการปฏิบัติงานในด้านต่างๆ
Course_6	จำนวนอาจารย์ที่ปรึกษาสอดคล้องกับจำนวนนศ.
Course_7	อาจารย์ที่ปรึกษาให้เวลาเพียงพอในการให้คำปรึกษา
Course_8	ระยะเวลาที่กำหนดไว้ในแผนการเรียนเหมาะสม
Course_9	เปิดโอกาสให้นัก.ลงทะเบียนวิชาต่างๆ
Course_10	ความเหมาะสมของจำนวนหน่วยกิตที่สถาบันกำหนด
Course_11	การจัดตารางเรียนสอดคล้องกับความต้องการเหมาะสม
Course_12	เชื่อมต่อการเรียนในระบบสองปริญญาหรือมากกว่า
ภาพรวมของคุณสมบัติอาจารย์	
Lecture_1	อาจารย์มีความรู้และประสบการณ์ตรงตามรายวิชาที่สอน
Lecture_2	สามารถถ่ายทอดและมีเทคนิคการสอนให้เข้าใจได้โดยง่าย
Lecture_3	ส่งเสริมให้นัก.คิดวิเคราะห์อย่างเป็นระบบ
Lecture_4	รับฟังและแลกเปลี่ยนความคิดเห็นกับนักศึกษา
Lecture_5	ส่งเสริมให้นัก.ได้ศึกษาค้นคว้าและการทำงานร่วมกัน
Lecture_6	นำความรู้และประสบการณ์ในงานวิจัยมาประยุกต์สอน
Lecture_7	ใช้แหล่งการเรียนรู้การเรียนการสอนรูปแบบอิเล็กทรอนิกส์
Lecture_8	ตำราหรือเอกสารการสอนที่อาจารย์พัฒนาขึ้นเอง
Lecture_9	ความเหมาะสมการวัดผลก่อนและหลังเรียน
Lecture_10	ส่งเสริมให้นัก.ให้ร่วมกิจกรรมเสริมหลักสูตรต่างๆ
Lecture_11	สอนสอดแทรกคุณธรรมจริยธรรมและจรรยาบรรณในวิชาชีพ
ภาพรวมของสิ่งอำนวยความสะดวก	
Facilities_1	มีวัสดุในการศึกษาในการเรียนเพียงพอ
Facilities_2	อุปกรณ์ห้องเรียนและห้องปฏิบัติการมีจำนวนเพียงพอและทันสมัย

ตารางที่ 4.1 (ต่อ)

Facilities_3	มีพื้นที่และวัสดุในการทำกิจกรรมต่างๆให้กับศ.อย่างเหมาะสม
Facilities_4	ความพร้อมทางด้านเทคโนโลยีสำหรับเชื่อมต่อเข้าระบบเครือข่าย
Facilities_5	แหล่งการเรียนรู้และการสืบค้นผ่านเครือข่ายมีจำนวนเพียงพอ
Facilities_6	มีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุด
Facilities_7	ความพร้อมของห้องสมุดอิเล็กทรอนิกส์ในการสืบค้นหนังสือ
Facilities_8	ความพร้อมของหน่วยงานดูแลและให้บริการจัดตั้งชุมนุมต่างๆ
Facilities_9	ความเหมาะสมของหน่วยงานให้คำแนะนำช่วยเหลือปัญหาต่างๆ
Facilities_10	มีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม
Facilities_11	มีระบบสืบค้นและจัดเก็บประวัติการเรียนของนักศึกษา
Facilities_12	มีระบบอิเล็กทรอนิกส์สนับสนุนการลงทะเบียนออนไลน์

### 3. การแปลงรูปแบบข้อมูล (Data transformation)

เมื่อเสร็จขั้นตอนการบันทึกข้อมูล เป็นข้อมูลทางคอมพิวเตอร์ ผ่านโปรแกรมไมโครซอฟต์เอ็กเซลแล้ว และเนื่องจากโปรแกรมไมโครซอฟต์เอ็กเซล เป็นโปรแกรมในกลุ่มของตารางคำนวณ ที่สามารถทำงานตามหลักการสถิติเบื้องต้น เราจึงสามารถหาค่าสถิติเบื้องต้นกับข้อมูลดังกล่าวได้ ซึ่งจะกล่าวต่อไป และเพื่อให้ข้อมูลทางคอมพิวเตอร์ สามารถถูกนำไปใช้ในการประมวลผลด้วยวิธีเหมืองข้อมูล จึงต้องทำการแปลงข้อมูลจากโปรแกรมไมโครซอฟต์เอ็กเซล ไปเป็นข้อมูลที่พร้อมสำหรับการประมวลผลเหมืองข้อมูลก่อน จากรายละเอียดที่ผู้วิจัยได้ระบุถึงซอฟต์แวร์สำหรับประมวลผลเหมืองข้อมูลมาแล้วนั้น ได้เลือกใช้โปรแกรมที่ชื่อว่า Weka ในการประมวลผล และเนื่องจาก Weka ต้องการรูปแบบข้อมูลเฉพาะเพื่อป้อนเข้าสู่กระบวนการทำงาน แม้ว่าตัวโปรแกรมเองจะสามารถเชื่อมโยงหรือเข้าถึงข้อมูลได้หลายหลาย ไม่ว่าจะเป็นข้อมูลที่อ่านจากเว็บผ่านทางอ้างอิงข้อมูล URL รวมถึงข้อมูลที่บันทึกไว้ในระบบฐานข้อมูล และอีกแบบหนึ่งก็คือข้อมูลตัวอักษรที่อยู่ในฟอร์แมต ARFF อันเป็นฟอร์แมตของข้อมูลที่ผู้วิจัยได้เลือกใช้งานวิจัยชิ้นนี้ และเนื่องจากข้อมูลที่ทำวิจัยได้ถูกบันทึกเข้าสู่โปรแกรมไมโครซอฟต์เอ็กเซลอยู่แล้ว จึงต้องมีการแปลงข้อมูลไปเป็นตัวอักษรเสียก่อน โปรแกรมไมโครซอฟต์เอ็กเซลมีความสามารถในการบันทึกข้อมูลแปลงไปเป็นไฟล์ข้อมูลตัวอักษร (เท็กซ์ไฟล์) ในรูปแบบ CSV (Comma-Separated Values) หรือไฟล์ตัวอักษรที่ถูกแบ่งแยกข้อมูลแต่ละชุดด้วยตัวอักษรแบ่ง (Delimiter Character) ดังนั้นผู้วิจัยจึงนำเอาคุณสมบัติในความสามารถ ของโปรแกรมไมโครซอฟต์เอ็กเซลนี้มาใช้งานกล่าวโดยสรุปก็คือ ใช้ซอฟต์แวร์ไมโครซอฟต์เอ็กเซลบันทึกไฟล์เนื้อข้อมูล

เพื่อมาวางในไฟล์ ARFF โดยเลือกแปลงฟอร์แมตไปเป็นไฟล์ตัวอักษร CSV จากนั้นนำไฟล์ตัวอักษรดังกล่าวมาปรับเปลี่ยนให้เข้าสู่รูปแบบฟอร์แมต ARFF (Attribute-Relation File Format) ที่เป็นรูปแบบที่รองรับการทำงานของโปรแกรม Weka

Excel (.xlsx)	CSV (.csv)	ARFF (.arff)
------------------	---------------	-----------------

สำหรับข้อมูลตัวอักษรในฟอร์แมต ARFF มีการแบ่งแยกข้อมูลเป็นสองส่วนหลักคือข้อมูลส่วนหัว (Head Data) และส่วนเนื้อข้อมูล (Content Data) ข้อมูลส่วนหัวของไฟล์ ARFF วางไว้ในส่วนต้นของไฟล์ ประกอบด้วยข้อมูล และคำอธิบายทั่วไป เช่น ผู้จัดทำ วันที่จัดทำ รายละเอียดของการทำงาน

```
% 1. Title: SRM
%
% 2. Sources:
% (a) Creator: Paralee Maneerat
% (b) Donor: SPU
% (c) Date: Jan, 2011
%
@RELATION srm
```

นอกจากนั้นแล้วข้อมูลส่วนหัว ยังใช้บอกลักษณะของเนื้อข้อมูล (ข้อมูลที่แปลงมาจากไมโครซอฟต์เอ็กเซล) เช่น ลำดับข้อมูล ชื่อข้อมูล และรูปแบบข้อมูลแบบใด เช่น ตัวเลข ตัวอักษร วันที่ ต่อไปนี้เป็นชื่อข้อมูล และรูปแบบข้อมูลทั้งหมด ที่ใช้ในงานวิจัย (อ้างอิงตามส่วน a) ดังต่อไปนี้

```
@ATTRIBUTE institute          STRING
@ATTRIBUTE institute_branch  STRING
@ATTRIBUTE institute_faculty  STRING
@ATTRIBUTE institure_program  STRING
```

@ATTRIBUTE gender	NUMERIC
@ATTRIBUTE homeland	NUMERIC
@ATTRIBUTE study_year	NUMERIC
@ATTRIBUTE study_course	NUMERIC
@ATTRIBUTE computer_owner	NUMERIC
@ATTRIBUTE pre_study_kind	NUMERIC
@ATTRIBUTE pre_study_level	NUMERIC
@ATTRIBUTE pre_study_gpa	NUMERIC
@ATTRIBUTE entrance_by	NUMERIC
@ATTRIBUTE current_gpa	NUMERIC
@ATTRIBUTE study_target	NUMERIC
@ATTRIBUTE study_time_out	NUMERIC
@ATTRIBUTE study_loan	NUMERIC
@ATTRIBUTE course_1	{0,1,2,3,4,5}
@ATTRIBUTE course_2	{0,1,2,3,4,5}
@ATTRIBUTE course_3	{0,1,2,3,4,5}
@ATTRIBUTE course_4	{0,1,2,3,4,5}
@ATTRIBUTE course_5	{0,1,2,3,4,5}
@ATTRIBUTE course_6	{0,1,2,3,4,5}
@ATTRIBUTE course_7	{0,1,2,3,4,5}
@ATTRIBUTE course_8	{0,1,2,3,4,5}
@ATTRIBUTE course_9	{0,1,2,3,4,5}
@ATTRIBUTE course_10	{0,1,2,3,4,5}
@ATTRIBUTE course_11	{0,1,2,3,4,5}
@ATTRIBUTE course_12	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_1	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_2	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_3	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_4	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_5	{0,1,2,3,4,5}

@ATTRIBUTE lecturer_6	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_7	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_8	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_9	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_10	{0,1,2,3,4,5}
@ATTRIBUTE lecturer_11	{0,1,2,3,4,5}
@ATTRIBUTE facilities_1	NUMERIC
@ATTRIBUTE facilities_2	NUMERIC
@ATTRIBUTE facilities_3	NUMERIC
@ATTRIBUTE facilities_4	NUMERIC
@ATTRIBUTE facilities_5	NUMERIC
@ATTRIBUTE facilities_6	NUMERIC
@ATTRIBUTE facilities_7	NUMERIC
@ATTRIBUTE facilities_8	NUMERIC
@ATTRIBUTE facilities_9	NUMERIC
@ATTRIBUTE facilities_10	NUMERIC
@ATTRIBUTE facilities_11	NUMERIC
@ATTRIBUTE facilities_12	NUMERIC

ส่วนสุดท้ายของไฟล์ฟอร์แมต ARFF คือส่วนของเนื้อข้อมูล หมายถึง เนื้อข้อมูลที่แต่ละชุดแยกกันด้วยตัวอักษรแบ่ง (ใช้ตัวอักษร , แบ่งข้อมูล) เนื้อข้อมูลก็คือข้อมูลดิบที่ได้แปลงจากข้อมูลเอกสารมาเป็นข้อมูลดิจิทัล ซึ่งก่อนหน้านี้ได้บันทึกลงในไฟล์เอกสาร ด้วยโปรแกรมไมโครซอฟต์เอ็กเซล ส่วนเนื้อข้อมูลนี้ทำได้โดยการแปลงมาจากไฟล์ข้อมูลของโปรแกรมไมโครซอฟต์เอ็กเซล มาเป็นเท็กซ์ไฟล์ (ไฟล์ตัวอักษร) ตามรูปแบบ CSV โดยข้อมูลนี้ต้องตรงตามลำดับ ตามที่ได้ระบุไว้ในข้อมูลส่วนหัวดังรายละเอียดตามฟอร์แมต ARFF ต่อไปนี้

@DATA

1,1,1,1,2,6,1,2,1,2,3,5,2,6,2,2,1,5,4,5,4,3,3,2,5,5,5,4,2,5,5,5,4,4,4,5,4,5,5,5,5,4,4,3,3,4,4,4,5,4,5  
 1,1,1,1,2,1,1,2,1,1,3,5,2,6,1,2,1,5,5,5,4,5,4,3,4,5,4,4,4,5,4,4,3,5,4,5,4,4,5,5,3,4,4,5,4,3,3,4,3,5,5,5  
 1,1,1,1,2,1,1,2,1,2,3,5,2,4,1,2,2,5,5,4,4,5,5,5,5,4,4,4,4,5,5,5,5,4,4,5,4,5,5,5,5,4,4,5,5,5,4,4,4,5  
 1,1,1,1,2,3,3,2,1,1,3,6,2,5,1,2,1,3,4,4,4,2,2,1,3,2,4,4,4,3,3,4,3,3,3,3,3,1,3,3,4,3,2,3,3,4,1,3,4  
 1,1,1,1,2,1,1,2,2,2,3,5,2,4,1,2,2,5,5,5,4,3,5,5,5,4,3,3,3,5,5,4,4,4,5,5,5,5,4,4,5,5,5,4,4,4,4,3,4,4  
 1,1,1,1,2,1,1,2,1,2,3,5,2,4,1,2,2,5,5,4,4,5,4,5,5,5,5,5,5,5,5,5,4,5,5,5,5,4,5,5,5,5,4,5,5,3,5,5  
 1,1,1,1,2,1,3,2,1,2,3,5,2,5,1,2,1,3,3,4,4,4,3,3,5,5,5,5,5,5,4,4,4,4,4,4,4,3,3,4,4,4,4,4,4,3,5,5  
 1,1,1,1,2,1,3,2,1,2,3,5,2,5,1,2,1,5,4,4,3,4,3,2,4,2,3,1,3,5,4,4,4,4,4,4,4,4,4,3,4,4,4,4,4,4,3,5,5  
 1,1,1,1,2,3,1,2,2,1,3,3,2,3,1,2,2,4,4,4,4,4,5,3,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,3,3,3,4,4,4,4,4,4,4,4  
 1,1,1,1,2,3,1,2,1,1,3,5,2,6,1,2,1,4,4,4,5,4,4,4,5,4,3,4,5,5,5,5,5,4,5,5,5,4,5,4,5,4,5,5,5,5,4,4  
 1,1,1,1,2,2,1,2,1,2,3,6,2,6,1,2,1,4,5,5,5,5,3,5,3,4,5,4,5,5,4,5,5,4,3,3,4,5,5,5,4,5,5,4,5,5,5,5  
 1,1,1,1,2,1,1,2,1,1,3,4,2,4,2,2,2,4,5,5,4,3,5,4,4,3,4,4,4,5,5,5,4,4,4,4,4,5,5,4,5,4,4,4,4,4,4,4,4  
 1,1,1,1,2,4,1,2,1,2,3,5,2,4,8,2,1,4,4,4,3,3,3,3,5,4,4,4,4,3  
 1,1,1,1,2,3,1,2,1,1,3,4,2,4,1,2,1,4,5,5,4,3,5,3,3,4,3,3,2,4,3,4,3,4,4,4,3,4,3,3,4,3,4,3,4,3,4,2,3,4  
 1,1,1,1,2,6,3,2,2,1,3,5,2,5,1,2,1,4,4,4,3,4,3,3,4,4,3,3,3,4,4,4,3,4,3,4,3,3,4,4,3,4,4,3,4,4,3,3,3  
 1,1,1,1,2,3,1,2,2,2,3,0,2,3,1,2,1,3,4,3,3,3,3,3,4,3,3,4,4,4,3,4,3,4,4,3,4,4,4,4,4,4,4,4,4,4,4,4  
 1,1,1,1,2,3,3,2,1,1,3,5,2,5,1,2,2,3,4,4,4,4,4,4,4,4,4,4,4,5,5,4,5,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4

จากข้อมูลตัวอย่างด้านบนคือเนื้อข้อมูลบางส่วนที่ตัดมาใช้ประกอบคำอธิบายหนึ่งบรรทัดคือหนึ่งข้อมูลของแบบสอบถามหนึ่งชุด หรือหนึ่งกลุ่มตัวอย่างนั่นเอง ข้อมูลทั้งสองส่วนทั้งข้อมูลส่วนหัว และส่วนเนื้อข้อมูล ต้องบันทึกไว้ในรูปแบบไฟล์ตัวอักษรที่มีนามสกุล .arff เมื่อได้ไฟล์ตามข้อกำหนด ARFF แล้วก็พร้อมที่จะนำไปใช้ในขั้นตอนต่อไป หากไฟล์ไม่สมบูรณ์ เช่น เนื้อหาข้อมูลไม่ครบ หรือไม่ตรงตามข้อกำหนดที่ระบุไว้ในส่วนหัว โปรแกรม Weka จะแสดงข้อผิดพลาด ถึงตำแหน่งบรรทัดที่มีปัญหา ซึ่งสามารถกลับมาแก้ไขข้อมูลให้ถูกต้อง และนำกลับไปใช้ประมวลผลต่อไปได้

#### 4.4 การแบ่งกลุ่มข้อมูล (Data Clustering)

ในการจัดกลุ่มข้อมูลนี้ใช้โปรแกรม Weka 3.6.5 จากข้อมูลทั้งหมด เมื่อนำมาแบ่งกลุ่มข้อมูล ตามระดับชั้นปีของนักศึกษา (Study\_Year) ที่ศึกษา โดยเลือกเอททริบิวต์ที่สำคัญ เพื่อให้ตรงตามวัตถุประสงค์ของวิจัย ผลจากการแบ่งกลุ่มตามชั้นปี เป็นดังนี้คือ

##### 4.4.1 การแบ่งกลุ่มนักศึกษาตามชั้นปี 1

จากการทดลองจัดกลุ่มข้อมูลตามชั้นปี 1 มีจำนวนนักศึกษาเท่ากับ 974 คน โดยพิจารณาค่า Sum Square Error และค่า Between/Within ร่วมด้วย เมื่อดูแนวโน้มของ 2 ค่านี้แล้วไม่ค่อยแตกต่างกัน พบว่าที่  $K = 3$  และ 4 เมื่อพิจารณาที่ค่า T-Test แล้วนั้น การจัดกลุ่มทั้ง 4 กลุ่มนี้มีความแตกต่างอย่างมีนัยสำคัญ (ระดับความเชื่อมั่น 95% และค่าความน่าจะเป็นน้อยกว่า 0.05) ดังรูปที่ 4.1

1 vs 2		2 vs 3		3 vs 4	
F-Test	T-Test	F-Test	T-Test	F-Test	T-Test
0.86437185	3.9224E-19	0.1770415	1.7519E-47	0.53272916	1.9797E-22
		1 vs 3		1 vs 4	
		F-Test	T-Test	F-Test	T-Test
		0.23756082	8.4483E-36	0.57527632	4.1E-16
				2 vs 4	
				F-Test	T-Test
				0.46502776	2.718E-34

รูปที่ 4.1 แสดงค่า F-Test และ T-test เปรียบเทียบของการแบ่งกลุ่มนักศึกษาในชั้นปีที่ 1

แต่เมื่อพิจารณาที่ค่า SSE และ BCV/WCV แล้ว พบว่า  $K=4$  เป็นจำนวนกลุ่มที่เหมาะสมกว่า แสดงดังรูปที่ 4.2



SSE =	945.32676				
between/within =	0.0302992				
Attribute	Full Data	1	2	3	4
	974	266	152	229	327
course_1	3.9456	3.6992	3.2697	4.5284	4.052
course_2	4.1027	3.8985	3.4671	4.6419	4.1865
course_3	4.0359	3.8195	3.3421	4.6288	4.1193
course_4	3.8398	3.6316	3.1974	4.4105	3.9083
course_5	3.7474	3.4887	3.0724	4.3275	3.8654
course_6	3.6704	3.4211	2.8355	4.3843	3.7615
course_7	3.5678	3.3571	2.625	4.3493	3.63
course_8	3.9836	3.7857	3.25	4.5983	4.055
course_9	3.8378	3.7068	2.9671	4.3624	3.9817
course_10	3.7546	3.5263	3.0395	4.3581	3.8502
course_11	3.5842	3.3496	2.9079	4.1965	3.6606
course_12	3.4887	3.2895	2.7961	4.0655	3.5688
lecturer_1	4.2146	4.0489	3.4276	4.8079	4.2997
lecturer_2	3.8799	3.5977	3.0329	4.6245	3.9817
lecturer_3	3.845	3.6504	2.9539	4.5153	3.948
lecturer_4	3.8162	3.5789	2.9539	4.524	3.9144
lecturer_5	3.8789	3.6053	3.0789	4.5371	4.0122
lecturer_6	3.847	3.5752	3.0789	4.4891	3.9755
lecturer_7	3.9322	3.6767	3.0526	4.6157	4.0703
lecturer_8	3.9333	3.6729	3.1579	4.6114	4.0306
lecturer_9	3.8316	3.5714	3.0526	4.5066	3.9327
lecturer_10	3.8121	3.5414	2.9145	4.5415	3.9388
lecturer_11	3.8696	3.5789	2.9934	4.5677	4.0245
facilities_1	3.923	3.6241	3.1382	4.5502	4.0917
facilities_2	3.8573	3.5188	2.9276	4.5808	4.0581
facilities_3	3.7813	3.3571	2.9408	4.4585	4.0428
facilities_4	3.9066	3.4774	3.1053	4.5764	4.159
facilities_5	3.8162	3.4135	2.9868	4.4803	4.0642
facilities_6	3.7967	3.3835	2.9211	4.5066	4.0428
facilities_7	3.8563	3.4699	2.9013	4.5633	4.1193
facilities_8	3.7053	3.2782	2.8684	4.393	3.9602
facilities_9	3.6776	3.3759	2.8618	4.3362	3.841
facilities_10	3.732	3.3835	2.9342	4.4498	3.8838
facilities_11	3.884	3.5489	3.0921	4.559	4.052
facilities_12	4.0092	3.6805	3.2303	4.655	4.1865
<b>Means</b>		<b>3.5595029</b>	<b>3.03928</b>	<b>4.4943257</b>	<b>3.9791171</b>
<b>Stdev</b>		<b>0.1727896</b>	<b>0.177965</b>	<b>0.1407969</b>	<b>0.1568359</b>

รูปที่ 4.2 แสดงการแบ่งกลุ่มจำนวน 4 กลุ่ม ของนักศึกษาชั้นปีที่ 1

#### 4.4.2 การแบ่งกลุ่มนักศึกษาตามชั้นปี 2

จากการทดลองจัดกลุ่มข้อมูลตามชั้นปี 2 มีจำนวนนักศึกษาเท่ากับ 637 คน โดยพิจารณาค่า Sum Square Error และค่า Between/Within ร่วมด้วย เมื่อดูแนวโน้มของ 2 ค่านี้แล้วไม่ค่อยแตกต่างกันนัก พบว่าที่  $K = 3, 4$  และ 6 เมื่อพิจารณาที่ค่า T-Test แล้วนั้น การจัดกลุ่มทั้ง 6 กลุ่มนี้มีความแตกต่างอย่างมีนัยสำคัญ(ระดับความเชื่อมั่น 95% และค่าความน่าจะเป็นน้อยกว่า 0.05) ดังรูปที่ 4.3

1 vs 2		2 vs 3		3 vs 4	
F-Test	T-Test	F-Test	T-Test	F-Test	T-Test
4.8669E-19	0.04786859	1.1002E-16	0.35457678	0.00014495	8.5959E-34
		1 vs 3		1 vs 4	
		F-Test	T-Test	F-Test	T-Test
		0.33059191	3.3478E-19	0.00372607	2.178E-24
				2 vs 4	
				F-Test	T-Test
				2.238E-26	7.3183E-06

รูปที่ 4.3 แสดงค่า F-Test และ T-test เปรียบเทียบของการแบ่งนักศึกษาในชั้นปีที่ 2

เมื่อพิจารณาที่ค่า SSE และ BCV/WCV แล้ว พบว่า  $K=6$  เป็นจำนวนกลุ่มที่เหมาะสมกว่า ดังแสดงดังรูปที่ 4.4

SSE =		629.11174					
between/within =		0.14950765					
Attribute	Full Data	1	2	3	4	5	6
	637	171	1	38	162	114	151
course_1	3.8289	3.7778	4	2.7895	4.1049	4.4474	3.3841
course_2	4.022	4.0351	4	3.1842	4.2716	4.5702	3.5364
course_3	3.9294	3.9123	4	3.0263	4.142	4.5614	3.4702
course_4	3.7567	3.655	5	3.0263	4.0062	4.3333	3.3444
course_5	3.7645	3.5731	2	3.0263	4.0802	4.4035	3.3576
course_6	3.719	3.5673	3	2.7105	4.0494	4.3333	3.3311
course_7	3.6201	3.345	2	2.9737	3.9568	4.2807	3.245
course_8	3.9325	3.8421	4	3.0263	4.2716	4.4737	3.4901
course_9	3.617	3.386	5	2.2895	4.0247	4.4123	3.1656
course_10	3.7598	3.655	1	2.5526	4.0741	4.4649	3.3311
course_11	3.4474	3.1053	2	2.0263	3.7963	4.3158	3.1722
course_12	3.3469	2.8538	4	2.2368	3.7346	4.1754	3.1391
lecturer_1	4.1743	4.2047	4	3.4737	4.3704	4.7281	3.6887
lecturer_2	3.8603	3.7135	4	2.8947	4.1235	4.5614	3.457
lecturer_3	3.8085	3.6491	3	2.6842	4.1543	4.5789	3.3245
lecturer_4	3.7928	3.7018	3	2.6579	4.0926	4.5175	3.3179
lecturer_5	3.8556	3.7368	3	3.1053	4.1358	4.5175	3.3841
lecturer_6	3.8179	3.5731	3	3.1579	4.2037	4.5263	3.3179
lecturer_7	3.8681	3.8947	5	2.4474	4.216	4.5526	3.298
lecturer_8	3.8728	3.7778	5	2.8158	4.2469	4.5789	3.3046
lecturer_9	3.8022	3.6667	2	2.6053	4.1605	4.5	3.3576
lecturer_10	3.7755	3.5497	2	2.6579	4.1852	4.5263	3.3179
lecturer_11	3.8477	3.731	3	3	4.216	4.5702	3.2583
facilities_1	3.8571	3.807	2	2.6842	4.142	4.6316	3.3311
facilities_2	3.7316	3.7076	5	2.6316	3.9568	4.5175	3.1921
facilities_3	3.6075	3.5146	2	2.5263	3.8951	4.5263	2.9934
facilities_4	3.8038	3.9006	3	2.5263	3.9691	4.6228	3.2252
facilities_5	3.7488	3.807	2	2.6579	3.9383	4.6053	3.1192
facilities_6	3.7221	3.7836	2	2.7105	3.9383	4.5877	3.0331
facilities_7	3.73	3.8012	2	2.5526	3.963	4.5088	3.1192
facilities_8	3.5981	3.4678	2	2.3947	3.8951	4.5088	3.053
facilities_9	3.5667	3.4269	2	2.2632	3.9074	4.4737	3.0132
facilities_10	3.6578	3.5906	3	2.1316	4	4.5175	3.106
facilities_11	3.7771	3.7135	5	2.6053	4.0926	4.5088	3.245
facilities_12	3.9419	4.0234	5	2.7368	4.1481	4.7193	3.3377
<b>Means</b>		<b>3.67</b>	<b>3.2</b>	<b>2.70827</b>	<b>4.07037</b>	<b>4.5045</b>	<b>3.2789</b>
<b>Stdev</b>		<b>0.2537</b>	<b>1.2078</b>	<b>0.32068</b>	<b>0.14208</b>	<b>0.1153</b>	<b>0.154</b>

รูปที่ 4.4 แสดงการแบ่งกลุ่มจำนวน 6 กลุ่ม ของนักศึกษาชั้นปีที่ 2

### 4.4.3 การแบ่งกลุ่มนักศึกษาตามชั้นปี 3

จากการทดลองจัดกลุ่มข้อมูลตามชั้นปี 3 มีจำนวนนักศึกษาเท่ากับ 567 คน โดยพิจารณาค่า Sum Square Error และค่า Between/Within ร่วมด้วย เมื่อดูแนวโน้มของ 2 ค่านี้แล้วไม่ค่อยแตกต่างกันนัก และพบว่าที่  $K = 3, 4$  และ 5 เมื่อพิจารณาที่ค่า T-Test แล้วนั้น การจัดกลุ่มทั้ง 5 กลุ่มนี้มีความแตกต่างอย่างมีนัยสำคัญ (ระดับความเชื่อมั่น 95% และค่าความน่าจะเป็นน้อยกว่า 0.05) ดังรูปที่ 4.5

1 vs 2		2 vs 3		3 vs 4		4 vs 5	
F-Test	T-Test	F-Test	T-Test	F-Test	T-Test	F-Test	T-Test
0.172625883	9.83012E-16	0.359295794	5.31185E-09	0.04217694	2.27439E-20	0.001227351	2.5187E-33
		1 vs 3		1 vs 4		1 vs 5	
		F-Test	T-Test	F-Test	T-Test	F-Test	T-Test
		0.651811107	0.00032935	0.013632659	3.56322E-12	0.417324623	3.08663E-26
				2 vs 4		2 vs 5	
				F-Test	T-Test	F-Test	T-Test
				0.258326698	5.35031E-34	0.030958954	6.50676E-13
						3 vs 5	
						F-Test	T-Test
						0.208054713	7.17534E-22

รูปที่ 4.5 แสดงค่า F-Test และ T-test เปรียบเทียบของการแบ่งนักศึกษาในชั้นปีที่ 3

เมื่อพิจารณาที่ค่า SSE และ BCW/WCV แล้ว พบว่า  $K=5$  เป็นจำนวนกลุ่มที่เหมาะสมกว่า แสดงดังรูปที่ 4.6

Attribute	Full Data	1	2	3	4	5
SSE =	585.457013					
between/within =	0.08425654					
	567	136	139	128	107	57
course_1	3.8095	4.1029	3.4604	3.7813	4.3458	3.0175
course_2	4.0653	4.3529	3.8345	3.9297	4.514	3.4035
course_3	3.97	4.2353	3.7266	3.8438	4.486	3.2456
course_4	3.7266	4.0662	3.4532	3.5859	4.271	2.8772
course_5	3.746	4.1176	3.4245	3.6328	4.3925	2.6842
course_6	3.6102	4.0735	3.1223	3.5625	4.1963	2.7018
course_7	3.5273	4.0074	2.9856	3.6094	4.0467	2.5439
course_8	3.9224	4.2206	3.6331	3.8047	4.3645	3.3509
course_9	3.6208	4.1103	3.259	3.3594	4.2991	2.6491
course_10	3.7407	4.0735	3.446	3.5547	4.3458	2.9474
course_11	3.418	3.9632	3.1439	2.9531	4.1495	2.4561
course_12	3.3492	3.8529	3.0432	3.0469	3.9065	2.5263
lecturer_1	4.1623	4.5147	3.8058	4.0078	4.7477	3.4386
lecturer_2	3.8536	4.2574	3.5108	3.7344	4.5047	2.7719
lecturer_3	3.8166	4.1912	3.4604	3.7578	4.3925	2.8421
lecturer_4	3.8765	4.1691	3.5252	3.7813	4.4766	3.1228
lecturer_5	3.8571	4.1176	3.5827	3.7031	4.4673	3.1053
lecturer_6	3.7989	4.2132	3.4245	3.6094	4.4486	2.9298
lecturer_7	3.8589	4.0956	3.6115	3.7422	4.5327	2.8947
lecturer_8	3.8624	4.2279	3.4892	3.7734	4.514	2.8772
lecturer_9	3.8025	4.1397	3.4317	3.7422	4.514	2.7018
lecturer_10	3.7848	4.0956	3.3669	3.7813	4.4673	2.7895
lecturer_11	3.8501	4.0735	3.5252	3.8984	4.4579	2.8596
facilities_1	3.7549	3.75	3.3237	3.9844	4.5981	2.7193
facilities_2	3.716	3.75	3.2302	4.0625	4.5701	2.4386
facilities_3	3.6138	3.6544	3.2158	3.8594	4.5234	2.2281
facilities_4	3.7319	3.7353	3.2518	4.0078	4.6262	2.5965
facilities_5	3.6561	3.5882	3.223	3.9844	4.5981	2.3684
facilities_6	3.5873	3.4926	3.1655	3.8438	4.5421	2.4737
facilities_7	3.6878	3.5956	3.2302	4.0156	4.5514	2.6667
facilities_8	3.5697	3.5147	3.1583	3.75	4.4766	2.5965
facilities_9	3.5256	3.5294	3.1079	3.6563	4.4486	2.5088
facilities_10	3.5979	3.6103	3.2086	3.7031	4.5421	2.5088
facilities_11	3.7619	3.75	3.3237	3.8984	4.5327	3.1053
facilities_12	3.9295	3.9118	3.5036	4.0625	4.7196	3.2281
<b>Mean</b>		<b>3.9758</b>	<b>3.37738</b>	<b>3.7435343</b>	<b>4.4448571</b>	<b>2.8050171</b>
<b>Stdev</b>		<b>0.2666</b>	<b>0.2104</b>	<b>0.2466633</b>	<b>0.1730222</b>	<b>0.3068547</b>

รูปที่ 4.6 แสดงการแบ่งกลุ่มจำนวน 5 กลุ่ม ของนักศึกษาชั้นปีที่ 3

#### 4.4.4 การแบ่งกลุ่มนักศึกษาตามชั้นปี 4

จากการทดลองจัดกลุ่มข้อมูลตามชั้นปี 4 มีจำนวนนักศึกษาเท่ากับ 339 คน โดยพิจารณาค่า Sum Square Error และค่า Between/Within ร่วมด้วย เมื่อดูแนวโน้มของ 2 ค่านี้แล้วไม่ค่อยแตกต่างกันนัก พบว่าที่  $K = 3$  เมื่อพิจารณาที่ค่า T-Test แล้วนั้น การจัดกลุ่มทั้ง 3 กลุ่มนี้มีความแตกต่างอย่างมีนัยสำคัญ(ระดับความเชื่อมั่น 95% และค่าความน่าจะเป็นน้อยกว่า 0.05 ) ดังรูปที่ 4.7

1 vs 2		2 vs 3		1 vs 3	
F-Test	T-Test	F-Test	T-Test	F-Test	T-Test
4.5753E-05	3.0616E-25	0.31092589	1.5201E-22	0.00162351	1.0057E-13

รูปที่ 4.7 แสดงค่า F-Test และ T-test เปรียบเทียบของการแบ่งนักศึกษาในชั้นปีที่ 4

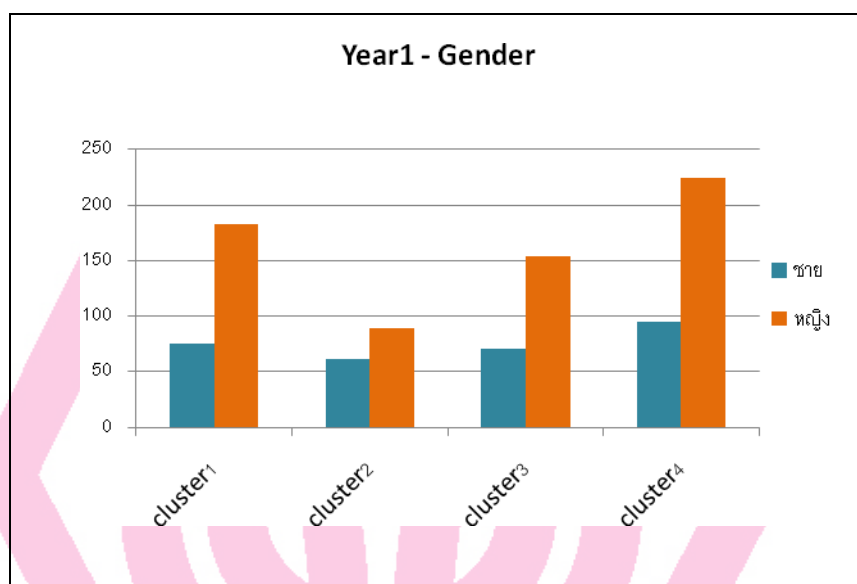
เมื่อพิจารณาที่ค่า SSE และ BCV/WCV แล้ว พบว่า  $K=3$  เป็นจำนวนกลุ่มที่เหมาะสมกว่า แสดงดังรูปที่ 4.8

SSE =	408.007482			
between/within =	0.03648255			
Attribute	Full Data	1	2	3
	339	94	92	153
course_1	3.8673	3.3936	4.3587	3.8627
course_2	4.0826	3.6809	4.5652	4.0392
course_3	4.0147	3.5745	4.4457	4.0261
course_4	3.8024	3.3298	4.2283	3.8366
course_5	3.7906	3.3298	4.3696	3.7255
course_6	3.646	3.1064	4.1739	3.6601
course_7	3.5664	2.9362	4.1304	3.6144
course_8	4.0029	3.5532	4.413	4.0327
course_9	3.5044	2.6383	4.3043	3.5556
course_10	3.7493	3.1702	4.3696	3.732
course_11	3.41	2.734	4.1413	3.3856
course_12	3.2684	2.5851	4.087	3.1961
lecturer_1	4.2183	3.7553	4.6739	4.2288
lecturer_2	3.9056	3.383	4.5761	3.8235
lecturer_3	3.885	3.4149	4.4783	3.817
lecturer_4	3.8584	3.4043	4.3913	3.817
lecturer_5	3.9351	3.5319	4.4783	3.8562
lecturer_6	3.9204	3.4574	4.5543	3.8235
lecturer_7	3.9587	3.4362	4.6304	3.8758
lecturer_8	3.941	3.4255	4.5326	3.902
lecturer_9	3.8378	3.3404	4.4239	3.7908
lecturer_10	3.8614	3.2766	4.4891	3.8431
lecturer_11	3.9174	3.2872	4.5543	3.9216
facilities_1	3.7316	3.0745	4.3696	3.7516
facilities_2	3.7198	3.0106	4.3261	3.7908
facilities_3	3.5752	2.7979	4.2609	3.6405
facilities_4	3.767	2.9149	4.4674	3.8693
facilities_5	3.6991	2.8085	4.3804	3.8366
facilities_6	3.6578	2.883	4.3696	3.7059
facilities_7	3.705	2.9681	4.3804	3.7516
facilities_8	3.5074	2.6489	4.2283	3.6013
facilities_9	3.4956	2.7021	4.0217	3.6667
facilities_10	3.5634	2.8191	4.2174	3.6275
facilities_11	3.7434	3.0213	4.3913	3.7974
facilities_12	3.8053	2.9787	4.4674	3.915
<b>Mean</b>		<b>3.1534943</b>	<b>4.3785714</b>	<b>3.7805743</b>
<b>Stdev</b>		<b>0.325734</b>	<b>0.1564361</b>	<b>0.1864476</b>

รูปที่ 4.8 แสดงการแบ่งกลุ่มจำนวน 4 กลุ่ม ของนักศึกษาชั้นปีที่ 4

#### 4.5 การแปลผล (Interpretation and Evaluation)

##### 4.5.1 จากผลการทดลองแบ่งกลุ่มนักศึกษาชั้นปี 1 ดังรูปที่ 4.9 แบ่งออกเป็น 4 กลุ่ม



รูปที่ 4.9 แสดงกราฟการแบ่งกลุ่มของนักศึกษาชั้นปีที่ 1

กลุ่มที่ 1 มีจำนวนนักศึกษาเท่ากับ 226 คน เท่ากับ 23.20% ที่ค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 3.56 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ดังนี้คือ

- อาจารย์ที่ปรึกษาใช้เวลาเพียงพอในการให้คำปรึกษานักศึกษา (Course 7)
- การจัดตารางเรียน ตารางสอบของแต่ละวิชาสอดคล้องกับความต้องการเหมาะสม (Course 11)
- เชื้อต่อการเรียนในระบบสองปริญญาหรือมากกว่า (Course 12)
- มีพื้นที่และวัสดุในการทำกิจกรรมต่างๆ ให้กับนักศึกษาอย่างเหมาะสม (Facilities 3 )
- มีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุดเพียงพอและทันสมัย(Facilities 6)
- ความพร้อมของหน่วยงานดูแลและให้บริการจัดตั้งชุมนุมต่างๆ(Facilities 8)
- ความเหมาะสมของหน่วยงานบริการให้คำแนะนำช่วยเหลือปัญหาต่างๆ(Facilities 9)
- มีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม (Facilities 10)

กลุ่มที่ 2 มีจำนวนนักศึกษาเท่ากับ 152 คน เท่ากับ 15.60% ที่มีค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 3.04 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ในทุกหัวข้อของ



แบบสอบถาม (แบบสอบถามในภาคผนวก ก ในด้านเนื้อหาหลักสูตร ตั้งแต่ข้อ 1-12 ด้านคุณสมบัติอาจารย์ ตั้งแต่ข้อ 1-11 และด้านสิ่งอำนวยความสะดวก ตั้งแต่ข้อ 1-12)

กลุ่มที่ 3 มีจำนวนนักศึกษาเท่ากับ 229 คน เท่ากับ 23.51% ที่มีค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 4.49 ซึ่งซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมากที่สุด (4.21-5.00) ไม่ต้องปรับปรุงในหัวข้อต่างๆ

กลุ่มที่ 4 มีจำนวนนักศึกษาเท่ากับ 327 คน เท่ากับ 33.57% ค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 3.98 ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก (3.41-4.20) หรือมากที่สุด(4.21-5.00) ไม่ต้องปรับปรุงในหัวข้อต่างๆ

การวิเคราะห์เพิ่มเติม ในกลุ่มที่ค่าเฉลี่ยต่ำกว่า 3.41 เพื่อมุ่งเน้นปรับปรุงความพึงพอใจให้กับนักศึกษากลุ่มนั้นคือกลุ่มที่ 1 และ กลุ่มที่ 2

- ลักษณะของนักศึกษาในกลุ่มที่ 1 มีลักษณะดังนี้คือ เป็นเพศหญิง 18.79% เพศชาย 7.7% มีช่วงเกรดเฉลี่ยส่วนใหญ่อยู่ในเกณฑ์ระหว่าง 2.01-3.51 เท่ากับ 21.46%
- ลักษณะของนักศึกษาในกลุ่มที่ 2 มีลักษณะดังนี้คือ เป็นเพศหญิง 9.14% เพศชาย 6.26% มีช่วงเกรดเฉลี่ยส่วนใหญ่อยู่ในเกณฑ์ระหว่าง 2.01-3.00 เท่ากับ 8.62%

จากลักษณะข้อมูลของนักศึกษาทั้งกลุ่มที่ 1 และกลุ่มที่ 2 เป็นกลุ่มที่ควรที่จะเข้าไปปรับปรุงในที่หัวข้อต่างๆ ที่สรุปมาแล้วข้างต้น เนื่องจากมีค่าระดับความพึงพอใจในหลายหัวข้อระดับปานกลางและน้อย

เมื่อนำกลุ่มที่ 1 มาวิเคราะห์กฎความสัมพันธ์ ด้วยวิธี JRip เกิดความสัมพันธ์หลายกฎ แต่เมื่อพิจารณากฎความสัมพันธ์ต่างๆ ที่ได้ นั้น มีความสัมพันธ์ที่ควรพิจารณา ดังนี้

1. (lecturer\_3 <= 3) and (facilities\_1 <= 3) and (course\_4 <= 3) and (lecturer\_1 <= 3)  
=> Cluster=cluster1 (55.0/3.0)

มีความต้องการให้อาจารย์ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น และมีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ และมีเนื้อหาของหลักสูตร เรียงลำดับการจัดการเรียนการสอนตามความยากง่ายและเนื้อหา และอาจารย์มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอน

2. (lecturer\_10 <= 3) and (lecturer\_3 <= 3) and (facilities\_5 <= 3) and (lecturer\_5 <= 3) and (course\_7 <= 3) and (lecturer\_2 <= 3) => Cluster=cluster1 (28.0/0.0)

มีความต้องการให้อาจารย์มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอนและให้อาจารย์ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น และมีแหล่งการเรียนรู้และการสืบค้นผ่านเครือข่ายมีจำนวนเพียงพอต่อความต้องการ และให้อาจารย์ ส่งเสริมให้นักศึกษาได้ศึกษาค้นคว้าและการทำงานร่วมกัน และมีระบบอาจารย์ที่ปรึกษา ให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา และอาจารย์สามารถถ่ายทอดและมีเทคนิคการสอนให้เข้าใจได้โดยง่าย

3. (course\_7 <= 2) and (facilities\_8 <= 3) and (facilities\_9 <= 3) and (course\_6 <= 2) => Cluster=cluster1 (14.0/2.0)

มีความต้องการอาจารย์ที่ปรึกษา ให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา และมี ความพร้อมของหน่วยงานดูแลและให้การบริการจัดตั้งชุมชนต่างๆ และมี ความเหมาะสมของหน่วยงานบริการให้คำแนะนำช่วยเหลือปัญหาต่างๆ และมีระบบจำนวนอาจารย์ที่ปรึกษา สอดคล้องกับจำนวนนักศึกษา

4. (lecturer\_9 <= 3) and (facilities\_7 <= 3) and (course\_7 <= 3) and (lecturer\_3 <= 3) and (Instance\_number >= 454) => Cluster=cluster1 (10.0/2.0)

มีความต้องการระบบความเหมาะสมการวัดผลก่อนและหลังเรียน การติดตามผล และการสื่อสารกับนักศึกษา และมีความพร้อมของห้องสมุดอิเล็กทรอนิกส์ ในการสืบค้นหนังสือ เอกสารวิชาการต่างๆ และมีระบบอาจารย์ที่ปรึกษา ให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา และอาจารย์ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น

เมื่อนำกลุ่มที่ 2 มาวิเคราะห์ทฤษฎีความสัมพันธ์ ด้วยวิธี JRip เกิดความสัมพันธ์หลายกฎ แต่เมื่อพิจารณาจากความสัมพันธ์ต่างๆ ที่ได้ นั้น มีความสัมพันธ์ที่ควรพิจารณา ดังนี้

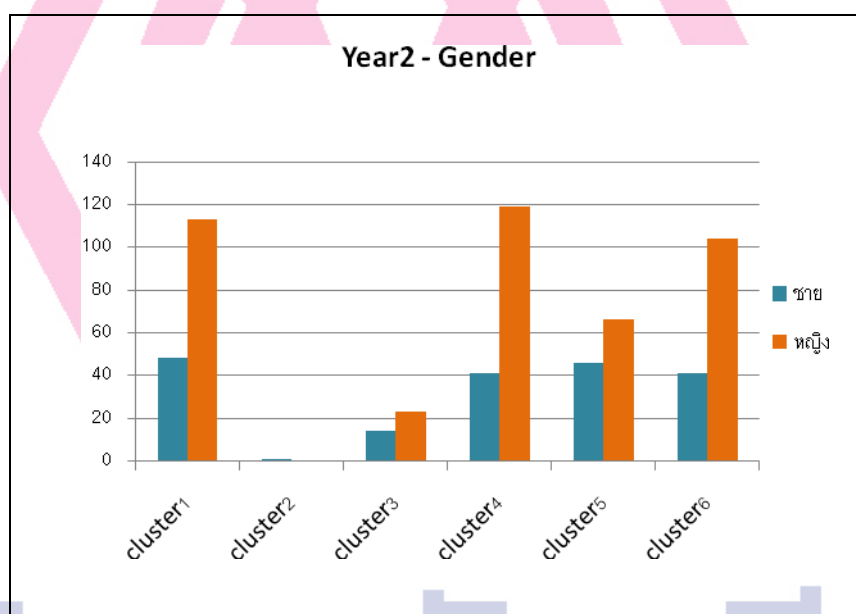
1. (facilities\_3 <= 3) and (course\_2 <= 4) and (facilities\_4 <= 4) and (facilities\_10 <= 3) => Cluster=cluster0 (46.0/2.0)

มีความต้องการพื้นที่และวัสดุในการทำกิจกรรมต่างๆ ให้นักศึกษาอย่างเหมาะสม และหลักสูตร เป็นความรู้ที่นักศึกษาสามารถนำไปใช้ในการประกอบอาชีพและดำรงชีวิตได้ และมีความพร้อมทางด้านเทคโนโลยีสำหรับเชื่อมต่อเข้าระบบเครือข่ายของมหาวิทยาลัย และ มีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ

2. (lecturer\_6 <= 3) and (facilities\_1 <= 4) and (course\_9 <= 4) and (facilities\_12 <= 4)  
=> Cluster=cluster 0(260/6)

มีความต้องการหลักสูตรที่อาจารย์นำความรู้และประสบการณ์ในงานวิจัย หรือระบบงานธุรกิจมาประยุกต์สอน และมีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ และหลักสูตรเปิดโอกาสให้นักศึกษาลงทะเบียนวิชาต่างๆ ได้ตามความเหมาะสม และมีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ

#### 4.5.2 จากผลการทดลองแบ่งกลุ่มนักศึกษาชั้นปี 2 ดังรูปที่ 4.10 ออกเป็น 6 กลุ่ม



รูปที่ 4.10 แสดงกราฟการแบ่งกลุ่มของนักศึกษาชั้นปีที่ 2

กลุ่มที่ 1 มีจำนวนนักศึกษาเท่ากับ 171 คน เท่ากับ 26.84% ที่ค่าระดับความพอใจเฉลี่ย ( $\bar{X}$ ) เท่ากับ 3.67 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ดังนี้

- เนื้อหาการเรียนในระบบสองปริญญาหรือมากกว่าพร้อมกัน (Course 12)

กลุ่มที่ 2 มีจำนวนนักศึกษาเท่ากับ 1 คน เท่ากับ 0.15% ที่มีค่าระดับความพอใจเฉลี่ย ( $\bar{X}$ ) เท่ากับ 3.2 แต่จำนวนนักศึกษาในกลุ่มนี้มีเพียง 1 คนเท่านั้น ถือว่าข้อมูลค่าผิดปกติ (Outlier) ไม่สามารถนำมาวิเคราะห์และแปลผลได้ ดังนั้นผู้วิจัยจึงไม่กล่าวถึงข้อมูลในกลุ่มนี้

กลุ่มที่ 3 มีจำนวนนักศึกษาเท่ากับ 38 คน เท่ากับ 5.96% ที่มีค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 2.70 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ในทุกหัวข้อของแบบสอบถาม (แบบสอบถามในภาคผนวก ก ในด้านเนื้อหาหลักสูตร ตั้งแต่ข้อ 1-12 ด้านคุณสมบัติอาจารย์ ตั้งแต่ข้อ 1-11 และด้านสิ่งอำนวยความสะดวก หลักสูตร ตั้งแต่ข้อ 1-12) ยกเว้นด้านคุณสมบัติของอาจารย์ ในเรื่องอาจารย์มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอน (lec1)

กลุ่มที่ 4 มีจำนวนนักศึกษาเท่ากับ 162 คน เท่ากับ 25.43% ค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 4.07 ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก (3.41-4.20) หรือมากที่สุด(4.21-5.00) ไม่ต้องปรับปรุงในหัวข้อต่างๆ

กลุ่มที่ 5 มีจำนวนนักศึกษาเท่ากับ 114 คน เท่ากับ 17.89% ค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 4.5 ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก (3.41-4.20) หรือมากที่สุด(4.21-5.00) ไม่ต้องปรับปรุงในหัวข้อต่างๆ

กลุ่มที่ 6 มีจำนวนนักศึกษาเท่ากับ 151 คน เท่ากับ 23.70% ค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 3.2 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ดังนี้

- ส่งเสริมให้นักศึกษามีความรู้ ความเข้าใจ ในหลักสูตร เช่น การประชาสัมพันธ์ ผ่านทางการแจกแผ่นพับ หรือ ทางอินเทอร์เน็ต และ การแจกคู่มือหลักสูตร(course 1)
- เรียงลำดับการจัดการเรียนการสอนตามความยากง่ายและเนื้อหา(course 4)
- ได้ฝึกทักษะการปฏิบัติงานในด้านต่างๆ เช่น ฝึกงาน สหกิจศึกษา,(course 5)
- จำนวนอาจารย์ที่ปรึกษา สอดคล้องกับจำนวนนักศึกษา(course 6)
- อาจารย์ที่ปรึกษาใช้เวลาเพียงพอในการให้คำปรึกษานักศึกษา (course 7)
- เปิดโอกาสให้นักศึกษาลงทะเบียนวิชาต่างๆ ได้ตามความต้องการ (course 9)
- ความเหมาะสมของจำนวนหน่วยกิตที่สถาบันกำหนดในแต่ละภาคเรียน (course 10)
- การจัดการเรียน ตารางสอบของแต่ละวิชาสอดคล้องกับความต้องการเหมาะสม (course 11)
- เชื้อต่อการเรียนในระบบสองปริญญาหรือมากกว่าพร้อมกัน (course 12)
- ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น(lecture 3)
- รับฟังและแลกเปลี่ยนความคิดเห็นกับนักศึกษา ระหว่างการสอน(lecture 4)
- ส่งเสริมให้นักศึกษาได้ศึกษาค้นคว้าและการทำงานร่วมกัน(lecture 5)

- อาจารย์นำความรู้และประสบการณ์ในงานวิจัย หรือระบบงานธุรกิจมาประยุกต์สอน (lecture 6)
- ใช้แหล่งการเรียนรู้การเรียนการสอนรูปแบบอิเล็กทรอนิกส์ ที่หลากหลาย รูปแบบ เช่น โต้ตอบ Onlineบทเรียนผ่านทางเว็บ,, Power point(lecture 7)
- ตำราหรือเอกสารการสอนที่อาจารย์พัฒนาขึ้นเอง สอดคล้องตามเนื้อหาวิชาและการเปลี่ยนแปลงทางวิชาการ(lecture 8)
- ความเหมาะสมการวัดผลก่อนและหลังเรียน การติดตามผล และการสื่อสารกับนักศึกษา (lecture 9)
- ส่งเสริมให้นักศึกษาเข้าร่วมกิจกรรมเสริมหลักสูตรต่างๆ เพื่อพัฒนาตนเองและสังคม (lecture 10)
- อาจารย์สอนสอดแทรกคุณธรรม จริยธรรมและจรรยาบรรณในวิชาชีพ (lecture 11)
- มีวัสดุในการศึกษา (เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ) ในการเรียนเพียงพอ (Facilities 1)
- อุปกรณ์ ห้องเรียนและห้องปฏิบัติการมีจำนวนเพียงพอและทันสมัย (Facilities 2)
- มีพื้นที่และวัสดุในการทำกิจกรรมต่างๆ ให้กับนักศึกษาอย่างเหมาะสม(Facilities 3)
- ความพร้อมทางด้านเทคโนโลยีสำหรับเชื่อมต่อเข้าระบบเครือข่ายของมหาวิทยาลัย (Facilities 4)
- แหล่งการเรียนรู้และการสืบค้นผ่านเครือข่ายมีจำนวนเพียงพอต่อความต้องการ (Facilities 5)
- มีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุดเพียงพอและทันสมัย(Facilities 6)
- ความพร้อมของห้องสมุดอิเล็กทรอนิกส์ในการสืบค้นหนังสือ เอกสารวิชาการต่างๆ (Facilities 7)
- ความพร้อมของหน่วยงานดูแลและให้บริการจัดตั้งชุมนุมต่างๆ (Facilities 8)
- ความเหมาะสมของหน่วยงานบริการให้คำแนะนำช่วยเหลือปัญหาต่างๆ(Facilities 9)
- มีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม(Facilities 10)
- มีระบบสืบค้นและจัดเก็บประวัติการเรียนของนักศึกษา (E Student) ซึ่งสามารถตรวจสอบได้ด้วยตัวเอง(Facilities 11)
- มีระบบอิเล็กทรอนิกส์สนับสนุนการลงทะเบียนออนไลน์(Facilities 12)

การวิเคราะห์เพิ่มเติม ในกลุ่มที่ค่าเฉลี่ยต่ำกว่า 3.41 เนื่องจากมีค่าระดับความพึงพอใจในหลายหัวข้อระดับปานกลางและน้อย เพื่อมุ่งเน้นปรับปรุงความพึงพอใจให้กับนักศึกษา กลุ่มที่ 3 และ กลุ่มที่ 6

- ลักษณะของนักศึกษาในกลุ่มที่ 3 มีลักษณะดังนี้คือ เป็นเพศหญิง 7.01% เพศชาย 2.2% มีช่วงเกรดเฉลี่ยส่วนใหญ่อยู่ในเกณฑ์ระหว่าง 2.01-3.51 เท่ากับ 5.03%
- ลักษณะของนักศึกษาในกลุ่มที่ 6 มีลักษณะดังนี้คือ เป็นเพศหญิง 16.33% เพศชาย 6.44% มีช่วงเกรดเฉลี่ยส่วนใหญ่อยู่ในเกณฑ์ระหว่าง 2.01-3.00 เท่ากับ 16.79%

จากลักษณะข้อมูลของนักศึกษาทั้งกลุ่มที่ 3 และกลุ่มที่ 6 เป็นกลุ่มที่ควรที่จะเข้าไปปรับปรุงในที่หัวข้อต่างๆ ที่สรุปมาแล้วข้างต้น เนื่องจากมีค่าระดับความพึงพอใจในหลายหัวข้อระดับปานกลางและน้อย

เมื่อนำกลุ่มที่ 3 มาวิเคราะห์กฎความสัมพันธ์ ด้วยวิธี JRip เกิดความสัมพันธ์หลายกฎ แต่เมื่อพิจารณากฎความสัมพันธ์ต่างๆ ที่ได้ นั้น มีความสัมพันธ์ที่ควรพิจารณา ดังนี้

1. (facilities\_10 <= 2) and (course\_8 <= 3) and (facilities\_2 <= 3) => Cluster=cluster2 (15.0/1.0)

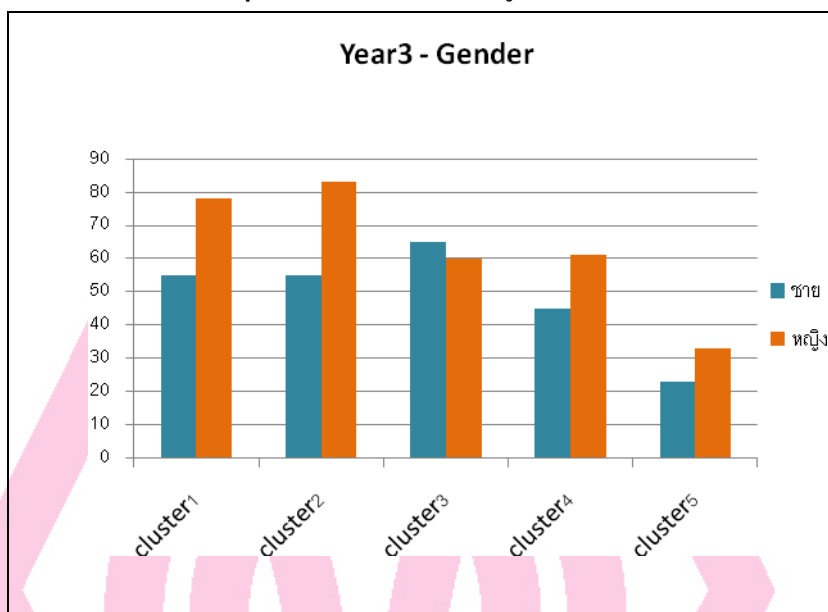
มีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม และ ระยะเวลา (จำนวนปี) ที่กำหนดไว้ในแผนการเรียนเหมาะสม และ อุปกรณ์ ห้องเรียนและห้องปฏิบัติการมีจำนวนเพียงพอและทันสมัย

เมื่อนำกลุ่มที่ 6 มาวิเคราะห์กฎความสัมพันธ์ ด้วยวิธี JRip เกิดความสัมพันธ์หลายกฎ แต่เมื่อพิจารณาความสัมพันธ์ต่างๆ ที่ได้ นั้น มีความสัมพันธ์ที่ควรพิจารณา ดังนี้

1. (facilities\_7 <= 3) and (lecturer\_11 <= 3) and (course\_10 <= 3) and (course\_4 <= 3) => Cluster=cluster5 (38.0/1.0)

มีความพร้อมของห้องสมุดอิเล็กทรอนิกส์ในการสืบค้นหนังสือ เอกสารวิชาการต่างๆ และอาจารย์สอนสอดแทรกคุณธรรม จริยธรรมและจรรยาบรรณในวิชาชีพ และความเหมาะสมของจำนวนหน่วยกิตที่สถาบันกำหนดในแต่ละภาคเรียน และเรียงลำดับการจัดการเรียนการสอนตามความยากง่ายและเนื้อหา

#### 4.5.3 จากผลการทดลองแบ่งกลุ่มนักศึกษาชั้นปี 3 ดังรูปที่ 4.11 แบ่งออกเป็น 5 กลุ่ม ดังนี้คือ



รูปที่ 4.11 แสดงกราฟการแบ่งกลุ่มของนักศึกษาชั้นปีที่ 3

กลุ่มที่ 1 มีจำนวนนักศึกษาเท่ากับ 136 คน เท่ากับ 23.98% ที่ค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 3.97 ไม่ต้องปรับปรุงในหัวข้อต่างๆ

กลุ่มที่ 2 มีจำนวนนักศึกษาเท่ากับ 139 คน เท่ากับ 24.51% ที่มีค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 3.37 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ดังนี้

- จำนวนอาจารย์ที่ปรึกษา สอดคล้องกับจำนวนนักศึกษา(Course 6)
- อาจารย์ที่ปรึกษาให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา(Course 7)
- เปิดโอกาสให้นักศึกษาลงทะเบียนวิชาต่างๆ ได้ตามความต้องการ(Course 9)
- การจัดการเรียน ตารางสอบของแต่ละวิชาสอดคล้องกับความต้องการเหมาะสม (Course 11)
- เอื้อต่อการเรียนในระบบสองปริญญาหรือมากกว่าพร้อม(Course 12)
- 
- ส่งเสริมให้นักศึกษาเข้าร่วมกิจกรรมเสริมหลักสูตรต่างๆ เพื่อพัฒนาตนเองและสังคม (lecture 10)
- มีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ (facilities 1)

- อุปกรณ์ ห้องเรียนและห้องปฏิบัติการมีจำนวนเพียงพอและทันสมัย (facilities 2)
- มีพื้นที่และวัสดุในการทำกิจกรรมต่างๆ ให้กับนักศึกษาอย่างเหมาะสม(facilities 3)
- ความพร้อมทางด้านเทคโนโลยีสำหรับเชื่อมต่อเข้าระบบเครือข่ายของมหาวิทยาลัย (facilities 4)
- แหล่งการเรียนรู้และการสืบค้นผ่านเครือข่ายมีจำนวนเพียงพอต่อความต้องการ (facilities 5)
- มีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุดเพียงพอและทันสมัย(facilities 6)
- ความพร้อมของห้องสมุดอิเล็กทรอนิกส์ในการสืบค้นหนังสือ เอกสารวิชาการต่างๆ (facilities 7)
- ความพร้อมของหน่วยงานดูแลและให้บริการจัดตั้งชุมนุมต่างๆ (facilities 8)
- ความเหมาะสมของหน่วยงานบริการให้คำแนะนำช่วยเหลือปัญหาต่างๆ(facilities 9)
- มีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม(facilities 10)
- มีระบบสืบค้นและจัดเก็บประวัติการเรียนของนักศึกษา (E Student) ซึ่งสามารถตรวจสอบได้ด้วยตัวเอง(facilities 11)

กลุ่มที่ 3 มีจำนวนนักศึกษาเท่ากับ 128 คน เท่ากับ 22.57% ที่มีค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 3.74 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ดังนี้

- เปิดโอกาสให้นักศึกษาลงทะเบียนวิชาต่างๆ ได้ตามความต้องการ(Course 9)
- การจัดการเรียน ตารางสอนของแต่ละวิชาสอดคล้องกับความต้องการเหมาะสม (Course 11)
- เชื่อมต่อการเรียนในระบบสองปริญญาหรือมากกว่าพร้อม(Course 12)

กลุ่มที่ 4 มีจำนวนนักศึกษาเท่ากับ 107 คน เท่ากับ 18.87% ค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 4.44 ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก (3.41-4.20) และมากที่สุด(4.21-5.00) ไม่ต้องปรับปรุงในหัวข้อต่างๆ

กลุ่มที่ 5 มีจำนวนนักศึกษาเท่ากับ 57 คน เท่ากับ 10.05% ค่าระดับความพอใจเฉลี่ย( $\bar{x}$ ) เท่ากับ 2.80 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ในทุกหัวข้อของแบบสอบถาม (แบบสอบถามในภาคผนวก ก ในด้านของเนื้อหาหลักสูตร ตั้งแต่ข้อ 1-12 ด้าน



คุณสมบัติอาจารย์ หลักสูตร ตั้งแต่ข้อ 1-11 และด้านสิ่งอำนวยความสะดวก หลักสูตร ตั้งแต่ข้อ 1-12)

การวิเคราะห์เพิ่มเติม ในกลุ่มที่ค่าเฉลี่ยต่ำกว่า 3.41 เนื่องจากมีค่าระดับความพึงพอใจในหลายหัวข้อระดับปานกลางและน้อย เพื่อมุ่งเน้นปรับปรุงความพึงพอใจให้กับนักศึกษากลุ่มที่ 2 และ กลุ่มที่ 5

- ลักษณะของนักศึกษาในกลุ่มที่ 2 มีลักษณะดังนี้คือ เป็นเพศหญิง 13.03% เพศชาย 8.63% มีช่วงเกรดเฉลี่ยส่วนใหญ่อยู่ในเกณฑ์ระหว่าง 2.01-3.50 เท่ากับ 18.37%
- ลักษณะของนักศึกษาในกลุ่มที่ 5 มีลักษณะดังนี้คือ เป็นเพศหญิง 9.14% เพศชาย 6.26% มีช่วงเกรดเฉลี่ยส่วนใหญ่อยู่ในเกณฑ์ระหว่าง 2.01-3.00 เท่ากับ 5.49%

จากลักษณะข้อมูลของนักศึกษาทั้งกลุ่มที่ 2 และกลุ่มที่ 5 เป็นกลุ่มที่ควรที่จะเข้าไปปรับปรุงในที่หัวข้อต่างๆ ที่สรุปมาแล้วข้างต้น เนื่องจากมีค่าระดับความพึงพอใจในหลายหัวข้อระดับปานกลางและน้อย

เมื่อนำกลุ่มที่ 2 มาวิเคราะห์กฎความสัมพันธ์ ด้วยวิธี JRip เกิดความสัมพันธ์หลายกฎ แต่เมื่อพิจารณากฎความสัมพันธ์ต่างๆ ที่ได้ นั้น มีความสัมพันธ์ที่ควรพิจารณา ดังนี้

1. (lecturer\_3 <= 3) and (facilities\_1 <= 3) and (course\_2 <= 3) and (course\_8 <= 3) and (facilities\_12 <= 3) => Cluster=cluster4 (9.0/2.0)

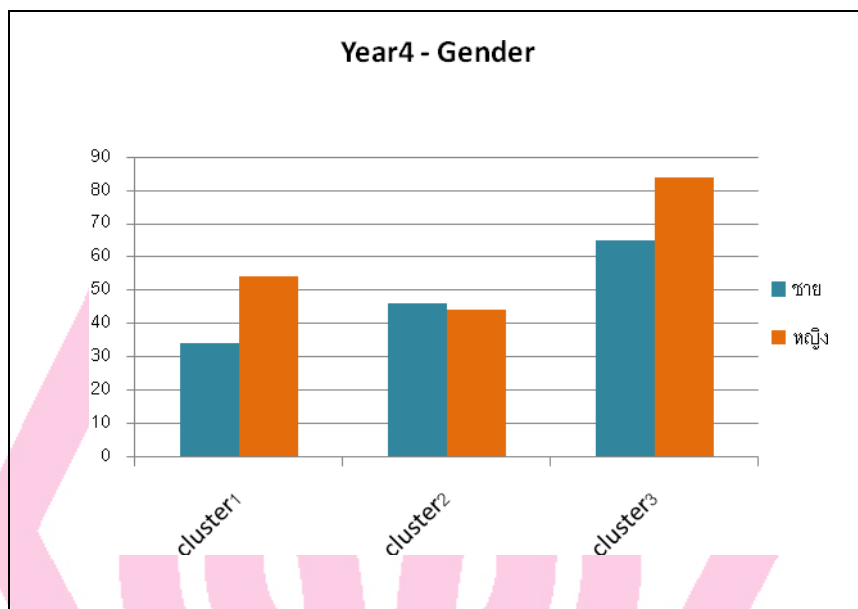
ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น และมีวัสดุในการศึกษา (เช่น กระดาษ หมีกิมพ์ และวัสดุอื่นๆ) ในการเรียนเพียงพอ และเป็นความรู้ที่นักศึกษาสามารถนำไปใช้ในการประกอบอาชีพและดำรงชีวิตได้ และระยะเวลา (จำนวนปี) ที่กำหนดไว้ในแผนการเรียนเหมาะสม และมีระบบอิเล็กทรอนิกส์สนับสนุนการลงทะเบียนออนไลน์

2. (lecturer\_5 <= 3) and (facilities\_6 <= 2) and (computer\_owner = 1) and (facilities\_10 <= 3) => Cluster=cluster4 (7.0/1.0)

ส่งเสริมให้นักศึกษาได้ศึกษาค้นคว้าและการทำงานร่วมกัน และมีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุดเพียงพอและทันสมัย และมีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม

เมื่อนำกลุ่มที่ 5 มาวิเคราะห์กฎความสัมพันธ์ ด้วยวิธี JRip ไม่เกิดความสัมพันธ์ใดๆ เลย

#### 4.5.4 จากผลการทดลองแบ่งกลุ่มนักศึกษาชั้นปี 4 ดังรูปที่ 4.12 ออกเป็น 4 กลุ่ม ดังนี้คือ



รูปที่ 4.12 แสดงกราฟการแบ่งกลุ่มของนักศึกษาชั้นปีที่ 4

กลุ่มที่ 1 มีจำนวนนักศึกษาเท่ากับ 94 คน เท่ากับ 27.72% ที่ค่าระดับความพอใจเฉลี่ย( $\bar{x}$ ) เท่ากับ 3.15 ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ดังนี้

- ส่งเสริมให้นักศึกษามีความรู้ ความเข้าใจ ในหลักสูตร เช่น การประชาสัมพันธ์ ผ่านทางการแจกแผ่นพับ หรือ ทางอินเทอร์เน็ต และ การแจกคู่มือหลักสูตร(course 1)
- เรียงลำดับการจัดการเรียนการสอนตามความยากง่ายและเนื้อหา(course 4)
- ได้ฝึกทักษะการปฏิบัติงานในด้านต่างๆ เช่น ฝึกงาน นักศึกษาสหกิจ (course 5)
- จำนวนอาจารย์ที่ปรึกษา สอดคล้องกับจำนวนนักศึกษา(course 6)
- อาจารย์ที่ปรึกษาให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา (course 7)
- เปิดโอกาสให้นักศึกษาลงทะเบียนวิชาต่างๆ ได้ตามความต้องการ (course 9)
- ความเหมาะสมของจำนวนหน่วยกิตที่สถาบันกำหนดในแต่ละภาคเรียน (course 10)
- การจัดการเรียน ตารางสอบของแต่ละวิชาสอดคล้องกับความต้องการเหมาะสม (course 11)
- เชื้อต่อการเรียนในระบบสองปริญญาหรือมากกว่าพร้อมกัน (course 12)
- รับฟังและแลกเปลี่ยนความคิดเห็นกับนักศึกษา ระหว่างการสอน(lecture 4)

- ความเหมาะสมการวัดผลก่อนและหลังเรียน การติดตามผล และการสื่อสารกับนักศึกษา (lecture 9)
- ส่งเสริมให้นักศึกษาเข้าร่วมกิจกรรมเสริมหลักสูตรต่างๆ เพื่อพัฒนาตนเองและสังคม (lecture 10)
- อาจารย์สอนสอดแทรกคุณธรรม จริยธรรมและจรรยาบรรณในวิชาชีพ (lecture 11)
- มีวัสดุในการศึกษา (เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ) ในการเรียนเพียงพอ (Facilities 1)
- อุปกรณ์ ห้องเรียนและห้องปฏิบัติการมีจำนวนเพียงพอและทันสมัย (Facilities 2)
- มีพื้นที่และวัสดุในการทำกิจกรรมต่างๆ ให้กับนักศึกษาอย่างเหมาะสม(Facilities 3)
- ความพร้อมทางด้านเทคโนโลยีสำหรับเชื่อมต่อเข้าระบบเครือข่ายของมหาวิทยาลัย (Facilities 4)
- แหล่งการเรียนรู้และการสืบค้นผ่านเครือข่ายมีจำนวนเพียงพอต่อความต้องการ (Facilities 5)
- มีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุดเพียงพอและทันสมัย(Facilities 6)
- ความพร้อมของห้องสมุดอิเล็กทรอนิกส์ในการสืบค้นหนังสือ เอกสารวิชาการต่างๆ (Facilities 7)
- ความพร้อมของหน่วยงานดูแลและให้บริการจัดตั้งชุมนุมต่างๆ (Facilities 8)
- ความเหมาะสมของหน่วยงานบริการให้คำแนะนำช่วยเหลือปัญหาต่างๆ(Facilities 9)
- มีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม(Facilities 10)
- มีระบบสืบค้นและจัดเก็บประวัติการเรียนของนักศึกษา (E Student) ซึ่งสามารถตรวจสอบได้ด้วยตัวเอง (Facilities 11)
- มีระบบอิเล็กทรอนิกส์สนับสนุนการลงทะเบียนเรียนออนไลน์(Facilities 12)

กลุ่มที่ 2 มีจำนวนนักศึกษาเท่ากับ 92 คน เท่ากับ 27.13% ที่มีค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 4.37 ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมากที่สุด(4.21-5.00) ไม่ต้องปรับปรุงในหัวข้อต่างๆ

กลุ่มที่ 3 มีจำนวนนักศึกษาเท่ากับ 153 คน เท่ากับ 45.13% ที่มีค่าระดับความพอใจเฉลี่ย ( $\bar{x}$ ) เท่ากับ 3.78 ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก (3.41-4.20) ควรปรับปรุงในหัวข้อต่างๆ ที่มีเกณฑ์คะแนนเฉลี่ยต่ำกว่า 3.41 ดังนี้

- การจัดการเรียน ตารางสอบของแต่ละวิชาสอดคล้องกับความต้องการเหมาะสม (course 11)
- เชื้อต่อการเรียนในระบบสองปริญญาหรือมากกว่าพร้อมกัน (course 12)

การวิเคราะห์เพิ่มเติม ในกลุ่มที่ค่าเฉลี่ยต่ำกว่า 3.41 เนื่องจากมีค่าระดับความพึงพอใจในหลายหัวข้อระดับปานกลางและน้อย เพื่อมุ่งเน้นปรับปรุงความพึงพอใจให้กับนักศึกษาในกลุ่มที่ 1

- ลักษณะของนักศึกษาในกลุ่มที่ 1 มีลักษณะดังนี้คือ เป็นเพศหญิง 8.48% เพศชาย 5.34% มีช่วงเกรดเฉลี่ยส่วนใหญ่อยู่ในเกณฑ์ระหว่าง 2.01-3.00 เท่ากับ 9.89%

จากลักษณะข้อมูลของนักศึกษาทั้งกลุ่มที่ 1 เป็นกลุ่มที่ควรที่จะเข้าไปปรับปรุงในที่หัวข้อต่างๆ ที่สรุปมาแล้วข้างต้น เนื่องจากมีค่าระดับความพึงพอใจในหลายหัวข้อระดับปานกลางและน้อย

**เมื่อนำกลุ่มที่ 1 มาวิเคราะห์กฎความสัมพันธ์ ด้วยวิธี JRip เกิดความสัมพันธ์หลายกฎ แต่เมื่อพิจารณากฎความสัมพันธ์ต่างๆ ที่ได้นั้น มีความสัมพันธ์ที่ควรพิจารณา ดังนี้**

1.  $(course\_1 \leq 3) \text{ and } (lecturer\_10 \leq 3) \text{ and } (lecturer\_6 \leq 3) \Rightarrow Cluster=cluster0$   
(7.0/1.0)

มีความต้องการส่งเสริมให้นักศึกษาวางใจ ในหลักสูตร เช่น การประชาสัมพันธ์ ผ่านทางการแจกแผ่นพับ หรือทางอินเทอร์เน็ต และ การแจกคู่มือหลักสูตร และให้ อาจารย์ มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอน และให้อาจารย์นำความรู้และประสบการณ์ในงานวิจัย หรือระบบงานธุรกิจมาประยุกต์สอน

## บทที่ 5

### สรุป อภิปรายผล และข้อเสนอแนะ

#### 5.1 สรุปผลการวิจัย

ในงานวิจัยนี้เป็นการเสนอข้อมูลเพื่อศึกษาตัวแปรที่มีต่อความพึงพอใจของนักศึกษาในมหาวิทยาลัย เพื่อนำมาสร้างแบบจำลองความสัมพันธ์ของนักศึกษา โดยศึกษาข้อมูลในอดีต เพื่อนำข้อมูลมาวิเคราะห์ โดยใช้เทคนิคเหมืองข้อมูล ซึ่งใช้เทคนิคการแบ่งกลุ่มข้อมูล โดยใช้ อัลกอริทึม K mean โดยใช้โปรแกรม Weka (Waikato Environment Knowledge Analysis) ในการสร้างแบบจำลอง หลังจากนั้นได้นำผลลัพธ์ที่ดีที่สุดจากเทคนิคการจัดกลุ่ม (Cluster Analysis) มาสร้างต้นแบบในขั้นสุดท้าย และได้เปรียบเทียบความถูกต้องในการจัดกลุ่มด้วยวิธี Sum of Squared Errors (SSE) , Between-cluster variation and Within-cluster variation และค่า T -Test และใช้เทคนิค JRip ช่วยในการวิเคราะห์กฎความสัมพันธ์ในแต่ละกลุ่มที่สนใจ โดยที่งานวิจัยนี้ได้ใช้ข้อมูลด้านการเรียนของนักศึกษาจากมหาวิทยาลัยรัฐและเอกชน จำนวน 7 แห่ง

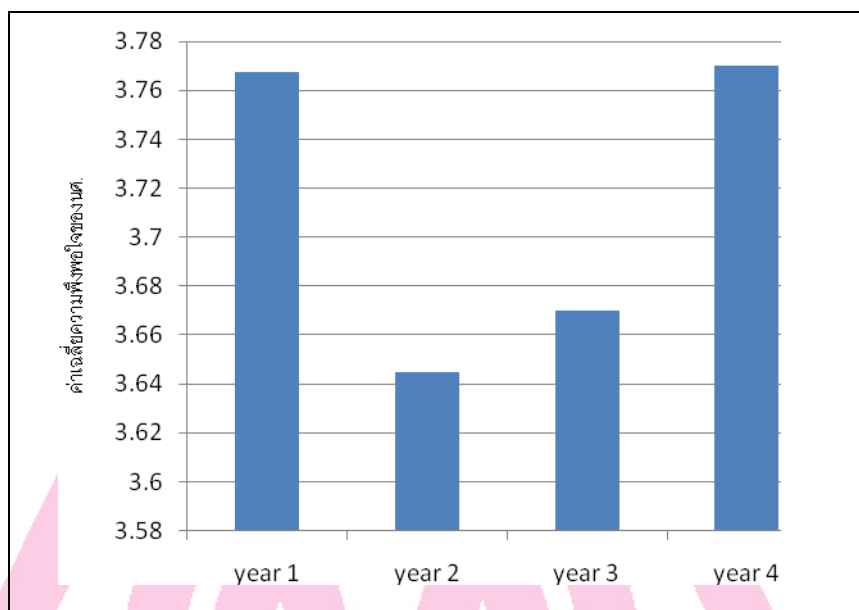
ผลการวิจัยทำให้ทราบถึงตัวแปรที่มีผลต่อความพึงพอใจของนักศึกษาในด้านต่างๆ

1. ด้านการเรียนการสอน
2. ด้านอาจารย์
3. ด้านอุปกรณ์และสิ่งอำนวยความสะดวก

จะเห็นว่าตัวแบบความน่าจะเป็นโดยใช้เทคนิคเหมืองข้อมูล โดยวิธีการแบ่งกลุ่ม อัลกอริทึม K mean สามารถแบ่งกลุ่มข้อมูลของนักศึกษาในชั้นปีที่ 1,2,3 และ 4 โดยที่แต่ละชั้นปี แบ่งแยกกลุ่มย่อยต่างๆ ลงไปอีก เพื่อวิเคราะห์ได้ชัดเจนว่าแต่ละกลุ่มในแต่ละชั้นปี มีความพึงพอใจของนักศึกษาในด้านต่างๆ

#### 5.2 อภิปรายผล

นอกจากงานวิจัยนี้จะได้ทราบถึงปัจจัยต่างๆ ที่มีผลต่อความพึงพอใจของนักศึกษาของนักศึกษาแต่ละชั้นปี แสดงได้ดังรูปที่ 5.1 และควรปรับปรุง ในแต่ละกลุ่มย่อยๆ ดังนี้



รูปที่ 5.1 แสดงกราฟค่าเฉลี่ยความพึงพอใจของนักศึกษาแต่ละชั้นปี

1. กลุ่มนักศึกษาชั้นปี 1 มีค่าความพึงพอใจเฉลี่ย ( $\bar{X}$ ) เท่ากับ 3.7675 โดยแบ่งออกเป็น 4 กลุ่มคือ

กลุ่มที่ 1 จำนวนนักศึกษาเท่ากับ 23.20% ควรปรับปรุงในหัวข้อต่างๆ ดังที่กล่าวไว้ในบทที่ 4

กลุ่มที่ 2 จำนวนนักศึกษาเท่ากับ 15.60% ต้องการให้ปรับปรุงในหัวข้อต่างๆ ในทุกหัวข้อของแบบสอบถาม (แบบสอบถามในภาคผนวก ก ในด้านเนื้อหาหลักสูตร ตั้งแต่ข้อ 1-12 ด้านคุณสมบัติอาจารย์ ตั้งแต่ข้อ 1-11 และด้านสิ่งแวดล้อมความสะดวก ตั้งแต่ข้อ 1-12)

กลุ่มที่ 3 จำนวนนักศึกษา เท่ากับ 23.51% ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมากที่สุด ไม่ต้องปรับปรุงในหัวข้อต่างๆ

กลุ่มที่ 4 จำนวนนักศึกษา เท่ากับ 33.57% ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก หรือมากที่สุดไม่ต้องปรับปรุงในหัวข้อต่างๆ

ดังนั้น นักศึกษาชั้นปีที่ 1 57.08% มีความพึงใจในด้านต่างๆ ทุกข้ออยู่ในระดับมากและมากที่สุด แต่นักศึกษา 42.92% มีความพึงพอใจในด้านต่างๆ น้อย โดยจะต้องปรับปรุงแก้ไข ตามที่กล่าวไว้แล้วในบทที่ 4

โดยงานวิจัยนี้ให้ความสำคัญในการปรับปรุงประสิทธิภาพความพึงพอใจให้กับนักศึกษา ดังนั้นจึงวิเคราะห์ข้อมูลในกลุ่มที่มีความพึงใจน้อย คือ กลุ่มที่ 1 และ กลุ่มที่ 2 จึงได้วิเคราะห์ถึงความสัมพันธ์ต่างๆ ดังนี้

- มีความต้องการให้อาจารย์ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบ และกล้าแสดงความคิดเห็น และมีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ และมีเนื้อหาของหลักสูตร เรียงลำดับการจัดการเรียนการสอนตามความยากง่ายและเนื้อหา และอาจารย์มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอน

- มีความต้องการให้อาจารย์มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอน และให้อาจารย์ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น และมีแหล่งการเรียนรู้และการสืบค้นผ่านเครือข่ายมีจำนวนเพียงพอต่อความต้องการ และให้อาจารย์ส่งเสริมให้นักศึกษาได้ศึกษาค้นคว้าและการทำงานร่วมกัน และมีระบบอาจารย์ที่ปรึกษา ให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา และอาจารย์สามารถถ่ายทอดและมีเทคนิคการสอนให้เข้าใจได้โดยง่าย

- มีความต้องการอาจารย์ที่ปรึกษา ให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา และมี ความพร้อมของหน่วยงานดูแลและให้บริการจัดตั้งชุมนุมต่างๆ และมี ความเหมาะสมของหน่วยงานบริการให้คำแนะนำช่วยเหลือปัญหาต่างๆ และมีระบบจำนวนอาจารย์ที่ปรึกษา สอดคล้องกับจำนวนนักศึกษา

- มีความต้องการระบบความเหมาะสมการวัดผลก่อนและหลังเรียน การติดตามผล และการสื่อสารกับนักศึกษา และมีความพร้อมของห้องสมุดอิเล็กทรอนิกส์ ในการสืบค้นหนังสือ เอกสารวิชาการต่างๆ และมีระบบอาจารย์ที่ปรึกษา ให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา และอาจารย์ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น

- มีความต้องการพื้นที่และวัสดุในการทำกิจกรรมต่างๆ ให้กับนักศึกษาอย่างเหมาะสม และหลักสูตร เป็นความรู้ที่นักศึกษาสามารถนำไปใช้ในการประกอบอาชีพและดำรงชีวิตได้ และมี ความพร้อมทางด้านเทคโนโลยีสำหรับเชื่อมต่อเข้าระบบเครือข่ายของมหาวิทยาลัย และมีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ

- มีความต้องการหลักสูตรที่อาจารย์นำความรู้และประสบการณ์ในงานวิจัย หรือระบบงานธุรกิจมาประยุกต์สอน และมีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ และหลักสูตรเปิดโอกาสให้นักศึกษาลงทะเบียนวิชาต่างๆ ได้ตามความเหมาะสม และมีวัสดุในการศึกษา เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ ในการเรียนเพียงพอ

จากข้อมูลกลุ่มที่ 1 และ 2 ที่มีความพึงพอใจน้อยและปานกลางในด้านต่างๆ ดังนั้นมหาวิทยาลัยจะต้องปรับปรุงในหัวข้อต่างๆ (ที่สรุปไว้แล้วในบทที่ 4) และนำการวิเคราะห์ความสัมพันธ์ที่ได้มาปรับปรุงการให้บริการนักศึกษา เพื่อสร้างความพึงพอใจให้เกิดประโยชน์สูงสุด

**2. กลุ่มนักศึกษาชั้นปี 2** มีค่าความพึงพอใจเฉลี่ย( $\bar{x}$ ) เท่ากับ 3.645 โดยแบ่งออกเป็น 6 กลุ่ม ดังนี้

**กลุ่มที่ 1** จำนวนนักศึกษา เท่ากับ 26.84% ที่ค่าระดับความพอใจเฉลี่ยเท่ากับ 3.67 ควรปรับปรุงในหัวข้อต่างๆ เกี่ยวต่อการเรียนในระบบสองปริญญาหรือมากกว่าพร้อมกัน

**กลุ่มที่ 2** จำนวนนักศึกษา เท่ากับ 0.15% ไม่สามารถนำมาวิเคราะห์และแปลผลได้ ดังนั้นผู้วิจัยจึงไม่กล่าวถึงข้อมูลในกลุ่มนี้

**กลุ่มที่ 3** จำนวนนักศึกษา เท่ากับ 5.96% ที่มี ควรปรับปรุงในหัวข้อต่างๆ ในทุกหัวข้อของแบบสอบถาม (แบบสอบถามในภาคผนวก ก ในด้านเนื้อหาหลักสูตร ตั้งแต่ข้อ 1-12 ด้านคุณสมบัติอาจารย์ ตั้งแต่ข้อ 1-11 และด้านสิ่งอำนวยความสะดวก หลักสูตร ตั้งแต่ข้อ 1-12) ยกเว้นด้านคุณสมบัติของอาจารย์ ในเรื่องอาจารย์มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอน (lec1)

**กลุ่มที่ 4** จำนวนนักศึกษา เท่ากับ 25.43% ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก หรือมากที่สุด

**กลุ่มที่ 5** จำนวนนักศึกษา เท่ากับ 17.89% ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก หรือมากที่สุด

**กลุ่มที่ 6** จำนวนนักศึกษา เท่ากับ 23.70% ควรปรับปรุงในหัวข้อต่างๆ ตามที่ได้กล่าวมาแล้วในบทที่ 4

ดังนั้น นักศึกษาชั้นปีที่ 2 43.32% มีความพึงใจในด้านต่างๆ ทุกข้ออยู่ในระดับมากและมากที่สุด แต่นักศึกษา 56.68% มีความพึงพอใจในด้านต่างๆ น้อย โดยจะต้องปรับปรุงแก้ไข ตามที่กล่าวไว้แล้วในบทที่ 4

โดยงานวิจัยนี้ให้ความสำคัญในการปรับปรุงประสิทธิภาพความพึงพอใจให้กับนักศึกษา ดังนั้นจึงวิเคราะห์ข้อมูลในกลุ่มที่มีความพึงใจน้อย จึงได้ถกความสัมพันธ์ต่างๆ ดังนี้



- มีความต้องการระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม และระยะเวลา (จำนวนปี) ที่กำหนดไว้ในแผนการเรียนเหมาะสม และ อุปกรณ์ ห้องเรียนและห้องปฏิบัติการมีจำนวนเพียงพอและทันสมัย

- มีความต้องการความพร้อมของห้องสมุดอิเล็กทรอนิกส์ในการสืบค้นหนังสือเอกสารวิชาการต่างๆ และอาจารย์สอนสอดคล้องคุณธรรม จริยธรรมและจรรยาบรรณในวิชาชีพ และความเหมาะสมของจำนวนหน่วยกิตที่สถาบันกำหนดในแต่ละภาคเรียน และเรียงลำดับการจัดการเรียนการสอนตามความยากง่ายและเนื้อหา

**3. กลุ่มนักศึกษาชั้นปี 3** มีค่าความพึงพอใจเฉลี่ย( $\bar{x}$ ) เท่ากับ 3.67 โดยแบ่งออกเป็น 5 กลุ่ม ดังนี้คือ

**กลุ่มที่ 1** จำนวนนักศึกษา เท่ากับ 23.98% ที่ค่าระดับความพอใจทุกข้ออยู่ในระดับมาก ไม่ต้องปรับปรุงในหัวข้อต่างๆ

**กลุ่มที่ 2** จำนวนนักศึกษา เท่ากับ 24.51% ควรปรับปรุงในหัวข้อต่างๆ ตามที่ได้กล่าวมาแล้วในบทที่ 4

**กลุ่มที่ 3** จำนวนนักศึกษา เท่ากับ 22.57% ควรปรับปรุงในหัวข้อต่างๆ ตามที่ได้กล่าวมาแล้วในบทที่ 4

**กลุ่มที่ 4** จำนวนนักศึกษา เท่ากับ 18.87% ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก และมากที่สุดไม่ต้องปรับปรุงในหัวข้อต่างๆ

**กลุ่มที่ 5** จำนวนนักศึกษา เท่ากับ 10.05% ควรปรับปรุงในหัวข้อต่างๆ ในทุกหัวข้อของแบบสอบถาม (แบบสอบถามในภาคผนวก ก ในด้านของเนื้อหาหลักสูตร ตั้งแต่ข้อ 1-12 ด้านคุณสมบัติอาจารย์ หลักสูตร ตั้งแต่ข้อ 1-11 และด้านสิ่งอำนวยความสะดวก หลักสูตร ตั้งแต่ข้อ 1-12)

ดังนั้น นักศึกษาชั้นปีที่ 3 จำนวน 42.85% มีความพึงพอใจในด้านต่างๆ ทุกข้ออยู่ในระดับมาก และมากที่สุด แต่จำนวนนักศึกษา 57.15% มีความพึงพอใจในด้านต่างๆ น้อย โดยจะต้องปรับปรุงแก้ไข ตามที่กล่าวไว้แล้วในบทที่ 4

โดยงานวิจัยนี้ให้ความสำคัญในการปรับปรุงประสิทธิภาพความพึงพอใจให้กับนักศึกษา ดังนั้นจึงวิเคราะห์ข้อมูลในกลุ่มที่มีความพึงพอใจน้อย จึงได้ศึกษาความสัมพันธ์ต่างๆ ดังนี้

- ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น และมีวัสดุในการศึกษา (เช่น กระดาษ หมีกิมพ์ และวัสดุอื่นๆ) ในการเรียนเพียงพอ และเป็นความรู้ที่นักศึกษาสามารถนำไปใช้ในการประกอบอาชีพและดำรงชีวิตได้ และระยะเวลา (จำนวนปี) ที่กำหนดไว้ในแผนการเรียนเหมาะสม และมีระบบอิเล็กทรอนิกส์สนับสนุนการลงทะเบียนออนไลน์

- ส่งเสริมให้นักศึกษาได้ศึกษาค้นคว้าและการทำงานร่วมกัน และมีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุดเพียงพอและทันสมัย และมีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม

**4. กลุ่มนักศึกษาชั้นปี 4** มีค่าความพึงพอใจเฉลี่ย( $\bar{X}$ ) เท่ากับ 3.77 โดยแบ่งออกเป็น 3 กลุ่ม ดังนี้คือ

**กลุ่มที่ 1** จำนวนนักศึกษา เท่ากับ 27.72% ควรปรับปรุงในหัวข้อต่างๆ ควรปรับปรุงในหัวข้อต่างๆ ตามที่ได้กล่าวมาแล้วในบทที่ 4

**กลุ่มที่ 2** จำนวนนักศึกษา เท่ากับ 27.13% ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมากที่สุด ไม่ต้องปรับปรุงในหัวข้อต่างๆ

**กลุ่มที่ 3** จำนวนนักศึกษา เท่ากับ 45.13% ซึ่งนักศึกษากลุ่มนี้มีความพึงพอใจในทุกหัวข้ออยู่ในระดับมาก

ดังนั้น นักศึกษาชั้นปีที่ 4 จำนวน 72.26% มีความพึงพอใจในด้านต่างๆ ทุกข้ออยู่ในระดับมากและมากที่สุด แต่จำนวนนักศึกษา 27.74% มีความพึงพอใจในด้านต่างๆ น้อย โดยจะต้องปรับปรุงแก้ไข ตามที่กล่าวไว้แล้วในบทที่ 4

โดยงานวิจัยนี้ให้ความสำคัญในการปรับปรุงประสิทธิภาพความพึงพอใจให้กับนักศึกษา ดังนั้นจึงวิเคราะห์ข้อมูลในกลุ่มที่มีความพียงใจน้อย จึงได้ทูลความสัมพันธ์ต่างๆ ดังนี้

- มีความต้องการส่งเสริมให้คนมีความรู้ ความเข้าใจ ในหลักสูตร เช่น การประชาสัมพันธ์ ผ่านทางการแจกแผ่นพับ หรือทางอินเทอร์เน็ต และ การแจกคู่มือหลักสูตร และให้อาจารย์มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอน และให้ อาจารย์นำความรู้และประสบการณ์ในงานวิจัย หรือระบบงานธุรกิจมาประยุกต์สอน

### 5.3 ปัญหาที่เกิดขึ้นระหว่างการดำเนินงานวิจัย

1. การจัดเก็บข้อมูลเป็นจำนวนมาก จากมหาวิทยาลัยต่างๆ ใช้ระยะเวลานานและความยุ่งยากในการติดต่อประสานในการเก็บข้อมูลในแต่ละมหาวิทยาลัย มีความยุ่งยาก ซับซ้อน แตกต่างกันไป
2. ระยะเวลาในการทำงานวิจัย ใช้เวลานานมาก ในการวิเคราะห์แบบสอบถามและทำความเข้าใจเกี่ยวกับเทคนิคและหลักการของเหมืองข้อมูล
3. การติดต่อประสานงานไปยังมหาวิทยาลัยต่างๆ ในการเก็บข้อมูล และนำแบบสอบถามเก็บมาวิเคราะห์

### 5.4 ข้อเสนอแนะ

1. การพัฒนาแบบจำลองเป็นระบบสนับสนุนการตัดสินใจ
2. สามารถนำข้อมูลจากแบบสอบถามที่จัดเก็บมาจากมหาวิทยาลัยทั้ง 7 แห่ง นำไปพัฒนาต่อ โดยเทคนิคเหมืองข้อมูลแบบอื่นๆ เช่น Association Rule เพื่อหาความสัมพันธ์ระหว่างตัวแปรต่างๆ
3. เพิ่มการวิเคราะห์เชิงลึก ในมุมมองการแบ่งข้อมูลในด้านอื่นๆ เช่น แบ่งตามชั้นปีและเกรดเฉลี่ย แบ่งตามชั้นปีและเพศ หรือแบ่งตามประเภทของมหาวิทยาลัยรัฐและเอกชนและแยกแบ่งย่อยตามชั้นปี เป็นต้น

## บรรณานุกรม

- Jacka, Mike J. and Keller, Paulette J.2002. **Business Process Mapping: Improving Customer Satisfaction** .New York: John Wiley & Sons,Inc.
- Jame, Craig and Dawn Jutla.2001. **E-Business: A customer- focused framework**. Boston : Addison-Wesley.
- Jill Dyche. 2001.**The CRM Handbook: A Business Guide to Customer Relationship Management**. Addison-Wesley.
- Mohamed, H. Peeru and Sagadevan. 2002. **A Customer Relationship Management: Step-by-Step**. New Delhi: Vikas Publishing House.
- Stanley A. Brown.2002.**Performance Driven In CRM**. Ontario: John Wiley and Sons.
- Storbacka, Kaj and Lehtinen, Jarmo R. 2001.**Customer Relationship Management**. Singapore: McGraw – Hill.
- Alan Paul. 2008. **Landscape Study of Student Lifecycle Relationship Management**. Birmingham : Delyth Chamber Ltd.
- Clifton Lemon. 2004. **Challenges in Marketing and Student Relationship Management in Higher Education**. np.
- M.Kantardzic.2003. **Data Mining Concepts, Models, Methods, and Algorithms**. John Wiley & Sons, Inc.
- Berson A. and Smith J.S.1997. **Data Warehousing, Data Mining, and OLAP**. McGraw-Hill,Inc.
- Han J. and Kamber M.2001. **Data Mining Concepts and Techniques**. Morgan Kaufmann Publishers.
- Fayyad U.M ,Piatetsky-Shaprio, Smyth P., and Uthurusamy R. 1996. **Advances in Knowledge Discovery and Data Mining**. AAAI/MIT Press.
- Intelliworks. 2004.**Challenges in Marketing and Student Relationship Management in Higher Education**.np

## บรรณานุกรม (ต่อ)

Najmuddin Shaik .2005. “**Service Center: A SRM Strategy to Promote Student Retention**”. Madison,20<sup>th</sup> Annual conference on Distance Teaching and Learning.

Alan Paul. 2008 . “**A generic student lifecycle relationship management system**” . Available from : [http://srmblog.alanpaul.co.uk/blog/\\_archives/2008/7/16/3795408.html](http://srmblog.alanpaul.co.uk/blog/_archives/2008/7/16/3795408.html)

ชุติมา อุดมะมุณีชัย.2553.”การตัวแบบระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์ สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา”. วิทยานิพนธ์ปริญญา มหาบัณฑิต,สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม.

ปรีชา ยามันสะบีดิน.2549.” การบริหารความสัมพันธ์กับนักศึกษาในสถาบัน ระดับอุดมศึกษาโดยการประยุกต์การทำเหมืองข้อมูล”.วิทยานิพนธ์ปริญญา มหาบัณฑิต,สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม.

รศ.ดร.กฤษณะ ไวยมัย และกัลยาภัทร อุดมสิน. 2550. “คลังข้อมูลอัจฉริยะด้านการศึกษา (INTELLIGENCE DATA WAREHOUSE FOR EDUCATION)”.วิทยานิพนธ์ปริญญา มหาบัณฑิต, คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์.

Hengqing Tong and Xiaochuan Lu .2009. “**Consumption Psychoanalysis and Customer Relationship Management Based on Association Rules Mining**”. World Congress on Computer Science and Information Engineering.

Vasile Paul Bresfelean .2007. “**Analysis and Predictions on Students' Behavior**” . Cavtat, Croatia :Proceedings of the ITI 2007 29th Int. Conf. on Information Technology Interfaces.

ภาคผนวก ก  
การหาข้อมูลทางสถิติเบื้องต้น

ข้อสรุปเบื้องต้นทางสถิติ ที่ได้จากการใส่สูตรในโปรแกรมไมโครซอฟต์เอ็กเซล โดยคำนวณเป็นเปอร์เซ็นต์ ตามข้อมูลดิบจากกลุ่มตัวอย่างทั้งหมด 2,727 ตัวอย่าง

1. กลุ่มตัวอย่างได้แบ่งสัดส่วนเป็นผู้หญิงมากกว่าผู้ชาย โดยมีเพศชาย 35.21 เปอร์เซ็นต์ ในขณะที่มีเพศหญิงถึง 62.07 เปอร์เซ็นต์

ชาย	35.21
หญิง	62.07

2. คณะที่ตอบคำถามไม่ได้บ่งบอกถึงความนิยมในการเข้าเรียน แต่เป็นสัดส่วนที่การลงพื้นที่ของเจ้าหน้าที่ ที่ลงไปสำรวจข้อมูล ได้เก็บตัวอย่างในรูปแบบใดบ้างก็เก็บจากกลุ่มตัวอย่างในห้องเรียน ซึ่งอาจเป็นผลให้จำนวนเปอร์เซ็นต์มีจำนวนสูง หากเป็นการเก็บข้อมูลจากนักศึกษาที่ผ่านไปมา ทำให้คณะที่สังกัดเรียนอยู่มีเปอร์เซ็นต์จำนวน ที่มีตัวเลขไม่สูงนัก และก็เป็นตัวเลขที่ไม่ใช่มีปัจจัยของคณะที่นิยมเข้าศึกษา

บัญชี	4.18
ศิลปศาสตร์	1.11
บริหารธุรกิจ	17.57
เทคโนโลยีสารสนเทศ	3.25
นิเทศศาสตร์	5.36
สถาปัตยกรรมศาสตร์	0.07
นิติศาสตร์	1.79
วิศวกรรมศาสตร์	5.04
เศรษฐศาสตร์	1.46
ดิจิทัลมีเดีย	0.29
วิทยาศาสตร์และเทคโนโลยี	12.11
มนุษยศาสตร์	3.93
สาธารณสุขศาสตร์และสิ่งแวดล้อม	8.82

พยาบาลศาสตร์	3.71
ภาษาและวัฒนธรรมจีน	2.96
เกษตรศาสตร์	1.46
เทคนิคการแพทย์	1.18
ครุศาสตร์	0.25
มนุษยศาสตร์และสังคมศาสตร์	0.04
วิทยาการจัดการ	0.25
เทคโนโลยีอุตสาหกรรม	0.11
วิทยาการสารสนเทศ	0.21
วิทยาศาสตร์	1.07
ศึกษาศาสตร์	0.07
วิทยาศาสตร์และศิลปศาสตร์	0.75
การจัดการและการท่องเที่ยว	0.75
ศิลปกรรมศาสตร์	0.07
รัฐศาสตร์และนิติศาสตร์	0.54

3. เนื่องจากกลุ่มตัวอย่าง สํารวจจากมหาวิทยาลัยที่อยู่ในกรุงเทพฯ และภาคตะวันออก ดังนั้นตัวเลขเปอร์เซ็นต์จึงเห็นชัดเจนว่า กลุ่มตัวอย่างมีฐานที่อยู่อยู่ในภาคกลาง และภาคตะวันออกเป็นหลัก

ภาคกลาง	49.43
ภาคเหนือ	5.32
ภาคใต้	9.86
ภาคตะวันออก	21.43
ภาคตะวันตก	1.71
ภาคตะวันออกเฉียงเหนือ	11.57

4. จากตัวเลขทำให้เห็นว่า กลุ่มตัวอย่างส่วนใหญ่มีคอมพิวเตอร์ส่วนตัวใช้ในการศึกษาเล่าเรียน ทำให้เห็นถึงว่าปัจจุบัน เครื่องคอมพิวเตอร์ถือเป็นปัจจัยหนึ่ง และเป็นเครื่องมือชิ้นหนึ่ง ที่ถูกนำมาประกอบเพื่อใช้งานในการศึกษาเล่าเรียน

มี	87.18
ไม่มี	12.54

5. การเข้าการศึกษาของกลุ่มตัวอย่าง เปรอร์เซ็นต์ตัวเลขที่สูงไปตกอยู่กับ การเข้าศึกษาแบบ  
รับสมัครตรง สูงกว่าถึง 3 เท่าของคนที่เข้ารับการศึกษาดังวิธี Entrance

Entrance/O-Net/A-Net	23.64
รับสมัครตรง	66.89
โควตาพิเศษ	6.82
ทุนการศึกษา	1.86

6. จากตารางที่แสดงเกรดเฉลี่ยก่อนเข้ารับการศึกษ และเกรดเฉลี่ยระหว่างการศึกษา จะเห็น  
ได้ว่าแนวโน้มของเกรดเฉลี่ยก่อนเข้าการศึกษา กลุ่มตัวอย่างมีคะแนนค่อนข้างสูง ในขณะที่เกรด  
เฉลี่ยระหว่างการศึกษา มีแนวโน้มในระดับกลางเป็นส่วนใหญ่ ซึ่งก็ยังไม่สามารถบอกไม่ได้ว่าเข้าศึกษา  
ระดับอุดมศึกษาแล้วเกรดลดลง เนื่องจากการหาค่าของกลุ่มตัวอย่างทั้งหมด ไม่ได้เป็นรายคน

เกรดเฉลี่ยก่อนเข้ารับการศึกษ

เกรดเฉลี่ย 1.00-1.50	0.29
เกรดเฉลี่ย 1.51-2.00	4.32
เกรดเฉลี่ย 2.01-2.50	18.00
เกรดเฉลี่ย 2.51-3.00	28.82
เกรดเฉลี่ย 3.01-3.51	32.25
เกรดเฉลี่ย 3.51-4.00	15.86

เกรดเฉลี่ยระหว่างการศึกษา

เกรดเฉลี่ย 1.00-1.50	1.43
เกรดเฉลี่ย 1.51-2.00	10.71
เกรดเฉลี่ย 2.01-2.50	29.54
เกรดเฉลี่ย 2.51-3.00	30.64
เกรดเฉลี่ย 3.01-3.51	20.25
เกรดเฉลี่ย 3.51-4.00	6.93



7. จากข้อมูลสรุปปัจจัยหลักที่กลุ่มตัวอย่าง เลือกศึกษาในการเข้าเรียน มีแนวโน้มสูงด้วยความต้องการการได้งานทำเป็นหลัก โดยมีเปอร์เซ็นต์เกือบครึ่งหนึ่งของทั้งหมด ในขณะที่กลุ่มตัวอย่างบางส่วนก็ให้ความสำคัญกับหลักสูตรของสถาบันการศึกษา

มีโอกาสได้งานสูง	41.14
หลักสูตรน่าเรียน	29.07
เรียนตามเพื่อน	1.93
ชื่อเสียงสถาบัน	5.14
เรียนตามเพื่อน	2.00
บิดามารดาชี้แนะ	7.75
เทคโนโลยีทันสมัย	4.96
อื่นๆ	7.54

#### 8. ตารางสรุปค่าเฉลี่ย

จากตารางที่คำนวณออกเป็นเปอร์เซ็นต์ จากการตอบแบบสอบถามจากกลุ่มตัวอย่าง โดยไล่ระดับคะแนนจากมากไปหาน้อย ตามแบบสอบถาม จะสังเกตเห็นคะแนนไปตกอยู่ในช่องความพึงพอใจ ระดับ มาก (ระดับที่ 4) เป็นส่วนใหญ่ สังเกตจากช่องที่ระบายสีพื้นเอาไว้ ไม่ว่าจะเป็นการตอบในหัวข้อเนื้อหาของหลักสูตร คุณสมบัติของอาจารย์ และการจัดการวัสดุการศึกษา ตำรา อาคารสถานที่ และระบบการให้บริการ

ภาพรวมของเนื้อหาของหลักสูตร	Mean				
ส่งเสริมให้นัก.มีความรู้ความเข้าใจในหลักสูตร	20.26	49.96	26.81	2.29	0.68
เป็นความรู้ที่นัก.สามารถนำไปใช้ในการประกอบอาชีพได้	28.32	51.34	18.87	1.18	0.29
เป็นพื้นฐานในการศึกษาชั้นสูงหรือศึกษาหาความรู้ใหม่	24.23	51.32	22.44	1.75	0.25
เรียงลำดับการจัดการเรียนการสอนตามความยากง่าย	16.21	50.25	30.31	2.76	0.47
ได้ฝึกทักษะการปฏิบัติงานในด้านต่างๆ	19.72	42.38	31.80	4.98	1.11
จำนวนอาจารย์ที่ปรึกษาสอดคล้องกับจำนวนนัก.	17.05	41.30	33.67	6.59	1.40
อาจารย์ที่ปรึกษาให้เวลาเพียงพอในการให้คำปรึกษา	15.62	39.15	34.10	8.85	2.29
ระยะเวลาที่กำหนดไว้ในแผนการเรียนเหมาะสม	24.60	49.25	23.13	2.40	0.61
เปิดโอกาสให้นัก.ลงทะเบียนวิชาต่างๆ	20.92	39.30	29.63	7.49	2.65
ความเหมาะสมของจำนวนหน่วยกิตที่สถาบันกำหนด	16.80	47.28	30.84	4.08	1.00
การจัดตารางเรียนสอดคล้องกับความต้องการเหมาะสม	13.10	38.51	36.01	9.56	2.83
เอื้อต่อการเรียนในระบบสองปริญญาหรือมากกว่า	11.53	34.09	39.92	10.85	3.62

ภาพรวมของคุณสมบัติอาจารย์ทุกคน					
อาจารย์มีความรู้และประสบการณ์ตรงตามรายวิชาที่สอน	35.44	49.19	14.04	1.08	0.25
สามารถถ่ายทอดและมีเทคนิคการสอนให้เข้าใจได้โดยง่าย	21.18	47.89	27.63	2.76	0.54
ส่งเสริมให้นัก.คิดวิเคราะห์อย่างเป็นระบบ	19.43	47.49	29.67	3.12	0.29
รับฟังและแลกเปลี่ยนความคิดเห็นกับนักศึกษา	18.88	49.57	27.14	3.91	0.50
ส่งเสริมให้นัก.ได้ศึกษาค้นคว้าและการทำงานร่วมกัน	20.68	49.34	26.54	2.91	0.54
นำความรู้และประสบการณ์ในงานวิจัยมาประยุกต์สอน	20.32	46.32	29.55	3.45	0.36
ใช้แหล่งการเรียนรู้การเรียนการสอนรูปแบบอิเล็กทรอนิกส์	25.99	42.38	27.07	3.67	0.90
ตำราหรือเอกสารการสอนที่อาจารย์พัฒนาขึ้นเอง	22.27	49.23	24.79	3.06	0.65
ความเหมาะสมการวัดผลก่อนและหลังเรียน	17.93	49.80	28.67	2.98	0.61
ส่งเสริมให้นัก.ให้ร่วมกิจกรรมเสริมหลักสูตรต่างๆ	18.62	47.47	29.82	3.41	0.68
สอนสอดแทรกคุณธรรมจริยธรรมและจรรยาบรรณในวิชาชีพ	22.39	45.49	27.83	3.39	0.90

ภาพรวมของการจัดวัสดุการศึกษาตำราอาคารสถานที่และระบบการให้บริการ					
มีวัสดุในการศึกษาในการเรียนเพียงพอ	22.30	44.56	27.65	4.60	0.90
อุปกรณ์ห้องเรียนและห้องปฏิบัติการมีจำนวนเพียงพอและทันสมัย	20.62	43.38	29.08	5.41	1.51
มีพื้นที่และวัสดุในการทำกิจกรรมต่างๆให้กับนศ.อย่างเหมาะสม	16.86	42.91	31.86	6.35	2.01
ความพร้อมทางด้านเทคโนโลยีสำหรับเชื่อมต่อเข้าระบบเครือข่าย	22.51	43.93	27.60	4.70	1.26
แหล่งการเรียนรู้และการสืบค้นผ่านเครือข่ายมีจำนวนเพียงพอ	19.53	43.77	29.23	6.14	1.33
มีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุด	19.07	42.44	30.45	6.64	1.40
ความพร้อมของห้องสมุดอิเล็กทรอนิกส์ในการสืบค้นหนังสือ	21.06	41.70	30.30	5.76	1.22
ความพร้อมของหน่วยงานดูแลและให้การบริการจัดตั้งชุมนุมต่างๆ	15.40	39.96	37.06	6.21	1.36
ความเหมาะสมของหน่วยงานให้คำแนะนำช่วยเหลือปัญหาต่างๆ	13.28	42.43	36.07	6.64	1.58
มีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม	16.61	41.64	33.86	6.52	1.37
มีระบบสืบค้นและจัดเก็บประวัติการเรียนของนักศึกษา	20.93	44.23	29.43	4.73	0.68
มีระบบอิเล็กทรอนิกส์สนับสนุนการลงทะเบียนออนไลน์	27.14	44.24	23.97	3.74	0.90

ภาคผนวก ข

แบบสอบถามงานวิจัย

เรื่อง การสร้างโมเดลการจัดการระบบนักศึกษาสัมพันธ์โดยใช้เทคนิคเหมืองข้อมูล

**The Development of Student Relationship Management Model  
using Data Mining Technique**

คำแนะนำในการตอบแบบสอบถาม

ตอนที่ 1 โปรดทำเครื่องหมาย / ลงใน ( ) และกรอกข้อความในช่องตามความเป็นจริง

ตอนที่ 2 โปรดประเมิน โดยเขียนเครื่องหมาย / ลงในช่องระดับความคิดเห็นตามเกณฑ์

ต่อไปนี้

- |           |                        |
|-----------|------------------------|
| 5 หมายถึง | ท่านเห็นด้วยมากที่สุด  |
| 4 หมายถึง | ท่านเห็นด้วยมาก        |
| 3 หมายถึง | ท่านเห็นด้วยปานกลาง    |
| 2 หมายถึง | ท่านเห็นด้วยน้อย       |
| 1 หมายถึง | ท่านเห็นด้วยน้อยที่สุด |

ตอนที่ 3 โปรดเขียนข้อเสนอแนะและความคิดเห็นในด้านต่างๆ

ตอนที่ 1 ข้อมูลเกี่ยวกับสถานภาพของผู้ตอบแบบสอบถาม

1. ชื่อสถาบัน..... วิทยาเขต.....

คณะ..... สาขาวิชา.....

2. เพศ [1] ชาย [2] หญิง

3. ภูมิลำเนาอยู่

[1] ภาคกลาง

[2] ภาคเหนือ

[3] ภาคใต้

[4] ภาคตะวันออก

[5] ภาคตะวันตก

[6] ภาคตะวันออกเฉียงเหนือ

4. ปัจจุบันเป็น นักศึกษา ระดับชั้นปีที่ .....

5. ท่านกำลังศึกษาใน

[1] ภาคปกติ .....ปี

[2] ภาคปกติต่อเนื่อง .....ปี

[3] ภาคค่ำ

หลักสูตร ....ปี [4] ภาคเสาร์-อาทิตย์หลักสูตร.....ปี [5] อื่น ๆ

(ระบุ).....

6. ปัจจุบันนักศึกษามีเครื่องคอมพิวเตอร์ส่วนตัวหรือไม่  
 [1] มี [2] ไม่มี
7. ลักษณะสถานศึกษาก่อนเข้าศึกษาระดับปริญญาตรีจากที่ใด  
 [1] รัฐบาล [2] เอกชน
8. วุฒิการศึกษาจากสถานศึกษาก่อนเข้าศึกษาระดับปริญญาตรี  
 [1] มัธยมศึกษาตอนปลาย/เทียบเท่า สาขาวิชา.....  
 [2] ปวช สาขาวิชา.....  
 [3] ปวส สาขาวิชา.....  
 [4] อื่นๆ ระบุสาขาวิชา.....
9. เกรดเฉลี่ย สะสม (GPA) ที่ท่านสำเร็จก่อนเข้าศึกษาระดับปริญญาตรี  
 [1] เกรดเฉลี่ย 1.00-1.50 [2] เกรดเฉลี่ย 1.51-2.00 [3] เกรดเฉลี่ย 2.01-2.50  
 [4] เกรดเฉลี่ย 2.51-3.00 [5] เกรดเฉลี่ย 3.01-3.50 [6] เกรดเฉลี่ย 3.51-4.00
10. ท่านเข้ามาศึกษาระดับปริญญาตรีโดยวิธี  
 [1] Entrance/O-Net/A-Net [2] รับสมัครตรง  
 [3] โควต้าพิเศษ [4] ทุนการศึกษา
11. เกรดเฉลี่ยสะสมในขณะนี้  
 [1] เกรดเฉลี่ย 1.00-1.50 [2] เกรดเฉลี่ย 1.51-2.00 [3] เกรดเฉลี่ย 2.01-2.50  
 [4] เกรดเฉลี่ย 2.51-3.00 [5] เกรดเฉลี่ย 3.01-3.50 [6] เกรดเฉลี่ย 3.51-4.00
12. ท่านคิดว่าเข้ามาศึกษาในสาขาวิชานี้เพื่ออะไร  
 [1] มีโอกาสได้งานสูง [2] หลักสูตรน่าเรียน [3] เรียนตามเพื่อน  
 [4] ชื่อเสียงสถาบัน [5] เรียนตามเพื่อน [6] บิดามารดาชี้แนะ  
 [7] เทคโนโลยีทันสมัย [8] อื่นๆ.....
13. ท่านเคยถูกพักการเรียนหรือไม่  
 [1] เคย [2] ไม่เคย
14. ท่านกู้เงินกองทุนเรียนหรือไม่  
 [1] กู้ [2] ไม่กู้

ตอนที่ 2 การประเมินความเหมาะสมของหลักสูตร อาจารย์ และการให้บริการต่างๆ ของมหาวิทยาลัย

2.1 ท่านเห็นด้วยกับภาพรวมของเนื้อหาของหลักสูตรดังต่อไปนี้ในระดับใด

คำถาม	ระดับความคิดเห็น				
	5	4	3	2	1
1. ส่งเสริมให้นักศึกษามีความรู้ ความเข้าใจ ในหลักสูตร เช่น การประชาสัมพันธ์ ผ่านทางการแจกแผ่นพับ หรือ ทางอินเทอร์เน็ต และการแจกคู่มือหลักสูตร	...	...	...	...	...
2. เป็นความรู้ที่นักศึกษาสามารถนำไปใช้ในการประกอบอาชีพและดำรงชีวิตได้	...	...	...	...	...
3. เป็นพื้นฐานในการศึกษาชั้นสูงหรือศึกษาหาความรู้ และวิทยาการใหม่ๆ	...	...	...	...	...
4. เรียงลำดับการจัดการเรียนการสอนตามความยากง่ายและเนื้อหา	...	...	...	...	...
5. ได้ฝึกทักษะการปฏิบัติงานในด้านต่างๆ เช่น ฝึกงาน ,สหกิจศึกษา	...	...	...	...	...
6. จำนวนอาจารย์ที่ปรึกษา สอดคล้องกับจำนวนนักศึกษา	...	...	...	...	...
7. อาจารย์ที่ปรึกษาให้เวลาเพียงพอในการให้คำปรึกษานักศึกษา	...	...	...	...	...
8. ระยะเวลา (จำนวนปี) ที่กำหนดไว้ในแผนการเรียนเหมาะสม	...	...	...	...	...
9. เปิดโอกาสให้นักศึกษาลงทะเบียนวิชาต่างๆ ได้ตามความต้องการ	...	...	...	...	...
10. ความเหมาะสมของจำนวนหน่วยกิตที่สถาบันกำหนดในแต่ละภาคเรียน	...	...	...	...	...
11. การจัดตารางเรียน ตารางสอบของแต่ละวิชาสอดคล้องกับความต้องการเหมาะสม	...	...	...	...	...
12. เนื้อต่อการเรียนรู้ในระบบสองปริญญาหรือมากกว่าพร้อมกัน	...	...	...	...	...

2.2 ท่านเห็นด้วยกับภาพรวมของคุณสมบัติอาจารย์ทุกคนดังต่อไปนี้ในระดับใด

คำถาม	ระดับความคิดเห็น				
	5	4	3	2	1
1. อาจารย์มีความรู้ และมีประสบการณ์ตรงตามรายวิชาที่สอน	...	...	...	...	...
2. อาจารย์สามารถถ่ายทอดและมีเทคนิคการสอนให้เข้าใจได้โดยง่าย	...	...	...	...	...
3. ส่งเสริมให้นักศึกษาคิดวิเคราะห์อย่างเป็นระบบและกล้าแสดงความคิดเห็น	...	...	...	...	...
4. รับฟังและแลกเปลี่ยนความคิดเห็นกับนักศึกษา ระหว่างการสอน	...	...	...	...	...

คำถาม	ระดับความคิดเห็น				
	5	4	3	2	1
5. ส่งเสริมให้นักศึกษาได้ศึกษาค้นคว้าและการทำงานร่วมกัน	...	...	...	...	...
6. อาจารย์นำความรู้และประสบการณ์ในงานวิจัย หรือระบบงานธุรกิจมา ประยุกต์สอน	...	...	...	...	...
7. ใช้แหล่งการเรียนรู้การเรียนการสอนรูปแบบอิเล็กทรอนิกส์ ที่หลากหลาย รูปแบบ เช่น โต้ตอบ Online, บทเรียนผ่านทางเว็บ, Power point	...	...	...	...	...
8. ตำราหรือเอกสารการสอนที่อาจารย์พัฒนาขึ้นเอง สอดคล้องตาม เนื้อหาวิชาและการเปลี่ยนแปลงทางวิชาการ	...	...	...	...	...
9. ความเหมาะสมการวัดผลก่อนและหลังเรียน การติดตามผล และการ สื่อสารกับนักศึกษา	...	...	...	...	...
10. ส่งเสริมให้นักศึกษาให้ร่วมกิจกรรมเสริมหลักสูตรต่างๆ เพื่อพัฒนา ตนเองและสังคม	...	...	...	...	...
11. อาจารย์สอนสอดแทรกคุณธรรม จริยธรรมและจรรยาบรรณในวิชาชีพ	...	...	...	...	...

### 2.3 ท่านเห็นด้วยกับภาพรวมของการจัดวัสดุการศึกษา ตำรา อาคารสถานที่ และระบบการ ให้บริการต่างๆ ดังต่อไปนี้ในระดับใด

คำถาม	ระดับความคิดเห็น				
	5	4	3	2	1
1. มีวัสดุในการศึกษา (เช่น กระดาษ หมึกพิมพ์ และวัสดุอื่นๆ) ในการเรียน เพียงพอ	...	...	...	...	...
2. อุปกรณ์ ห้องเรียนและห้องปฏิบัติการมีจำนวนเพียงพอและทันสมัย	...	...	...	...	...
3. มีพื้นที่และวัสดุในการทำกิจกรรมต่างๆ ให้กับนักศึกษาอย่างเหมาะสม	...	...	...	...	...
4. ความพร้อมทางด้านเทคโนโลยีสำหรับเชื่อมต่อเข้าระบบเครือข่ายของ มหาวิทยาลัย	...	...	...	...	...
5. แหล่งการเรียนรู้และการสืบค้นผ่านเครือข่ายมีจำนวนเพียงพอต่อความ ต้องการ	...	...	...	...	...
6. มีตำราและเอกสารเฉพาะสาขาวิชาในห้องสมุดเพียงพอและทันสมัย	...	...	...	...	...

คำถาม	ระดับความคิดเห็น				
	5	4	3	2	1
7. ความพร้อมของห้องสมุดอิเล็กทรอนิกส์ในการสืบค้นหนังสือ เอกสาร วิชาการต่างๆ	...	...	...	...	...
8. ความพร้อมของหน่วยงานดูแลและให้บริการจัดตั้งชุมนุมต่างๆ	...	...	...	...	...
9. ความเหมาะสมของหน่วยงานบริการให้คำแนะนำช่วยเหลือปัญหาต่างๆ	...	...	...	...	...
10. มีระบบการสนับสนุนให้ทุนการศึกษาเพียงพอและเหมาะสม	...	...	...	...	...
11. มีระบบสืบค้นและจัดเก็บประวัติการเรียนของนักศึกษา (E Student) ซึ่ง สามารถตรวจสอบได้ด้วยตัวเอง	...	...	...	...	...
12. มีระบบอิเล็กทรอนิกส์สนับสนุนการลงทะเบียนออนไลน์	...	...	...	...	...

ตอนที่ 3 โปรดเขียนข้อเสนอแนะและความคิดเห็นต่างๆ เกี่ยวกับมหาวิทยาลัยของท่าน

3.1 ความคิดเห็นทั่วไปเกี่ยวกับเนื้อหาของหลักสูตร

.....

.....

3.2 ความคิดเห็นทั่วไปเกี่ยวกับคุณลักษณะอาจารย์

.....

.....

3.3 ความคิดเห็นทั่วไปเกี่ยวกับการจัดวัสดุ ตำรา อาคารสถานที่ และระบบการให้บริการต่างๆ

.....

.....