

# รายงานการวิจัย

## เรื่อง

การใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผล  
ทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์

Utilizing Data Mining Techniques in Knowledge Based Creation for Education  
Achievement Prediction of Ratchaphruek College Students

โดย

นางสาวเรวดี ศักดิ์คุณธรรม

การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากวิทยาลัยราชพฤกษ์

ปีการศึกษา 2552

# รายงานการวิจัย

## เรื่อง

การใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผล  
ทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์

Utilizing Data Mining Techniques in Knowledge Based Creation for Education  
Achievement Prediction of Ratchaphruek College Students

## โดย

นางสาวเรวดี ศักดิ์คุลยธรรม

การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากวิทยาลัยราชพฤกษ์

ปีการศึกษา 2552

ปีที่ทำการวิจัยแล้วเสร็จ 2553

ชื่อโครงการวิจัย	การใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์
ชื่อผู้วิจัย	นางสาวเรวดี ศักดิ์คุลยธรรม Miss Raywadee Sakdulyatham
ปีที่ทำการวิจัย	2552

## บทคัดย่อ

งานวิจัยนี้ เป็นการประยุกต์ใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ เพื่อนำไปใช้ประกอบการตัดสินใจให้คำปรึกษาทางการเรียนของอาจารย์ที่ปรึกษา ซึ่งเป็นการทำนายเพื่อหาความถนัดของนักศึกษาว่าควรเรียนในสาขาวิชาใด โดยการนำข้อมูลส่วนตัว และข้อมูลผลการลงทะเบียนของนักศึกษา จากคณะบริหารธุรกิจ ทั้ง 4 สาขา คือ สาขาวิชาคอมพิวเตอร์ธุรกิจ, สาขาวิชาการตลาด, สาขาวิชาการจัดการ และสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว มาทำการสร้างโมเดล 2 โมเดล คือ โมเดลสำหรับวิเคราะห์หาพฤติกรรมทางการเรียนของนักศึกษาในแต่ละสาขาวิชา ซึ่งโมเดลบอกได้ว่า ผลการเรียนกลุ่มวิชาแกนการเงิน มีผลต่อผลการเรียนในกลุ่มวิชาเอกบังคับของสาขาวิชาคอมพิวเตอร์ธุรกิจ และสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยวมากที่สุด และผลการเรียนกลุ่มวิชาแกนธุรกิจมีผลต่อผลการเรียนในกลุ่มวิชาเอกบังคับของสาขาวิชาการตลาดและการจัดการมากที่สุด นอกจากนี้ยังมีการสร้างโมเดลที่ใช้ในการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา โดยโมเดลทำนายผลการเรียนของสาขาวิชาคอมพิวเตอร์ มีความถูกต้องคิดเป็น 73.49%, โมเดลทำนายผลการเรียนของสาขาวิชาการตลาด มีความถูกต้องคิดเป็น 83.58%, โมเดลทำนายผลการเรียนของสาขาวิชาการจัดการ มีความถูกต้องคิดเป็น 78.12% และโมเดลทำนายผลการเรียนของสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว มีความถูกต้องคิดเป็น 86.67% อย่างไรก็ตาม ผลการทำนายที่ได้ เป็นเพียงข้อมูลที่ช่วยประกอบการตัดสินใจในการให้คำปรึกษาของอาจารย์ที่ปรึกษาเท่านั้น ในความเป็นจริงควรจะอาศัยปัจจัยอีกหลายอย่างประกอบการตัดสินใจด้วย เช่น การเอาใจใส่ในการเรียน, การมาเรียน, การส่งงาน เป็นต้น

คำสำคัญ : ดาต้าไมน์นิ่ง, ฐานความรู้, อาจารย์ที่ปรึกษา

Research Title : Utilizing Data Mining Techniques in Knowledge Based Creation  
for Education Achievement Prediction of Ratchaphruek College  
Students

Researcher : Miss Raywadee Sakdulyatham

Year : 2009

### **Abstract**

This research applies Data Mining techniques in knowledge based creation for education achievement predict of Ratchaphruek College students. The research purpose is assist in decision making to give education advice of advisor about prediction for ability searching which program that they should study. By using students profile and registration information in Business faculty consist of 4 programs are Business computer, Marketing, Management and Hotel and Tourism Management for 2 models building are the first model for analysis about study profile of students each program, the result is core subject in finance grade to affect grade of major in Business computer and Hotel and Tourism Management program mostly and core subject in business grade to affect grade of major in Marketing and Management program mostly, and the second model is model for grade prediction of students each program, the result is analysis accuracy of Business computer program prediction model is 73.49%, analysis accuracy of Marketing program prediction model is 83.58%, analysis accuracy of Management program prediction model is 78.12% and analysis accuracy of Hotel and Tourism Management program prediction model is 86.67%. However this prediction results are the information for assist in decision making for advisors give education advice only, the truth should apply many factors in decision making to advice such as to pay attention in study and to attend class etc.

Key words : Data Mining, Knowledge based, Advisor

## กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณวิทยาลัยราชพฤกษ์ ที่ได้จัดสรรงบประมาณเพื่อสนับสนุนการวิจัยในครั้งนี้ งานวิจัยนี้สำเร็จได้อย่างดีด้วยความอนุเคราะห์ไฟล์ข้อมูลนักศึกษา คณะบริหารธุรกิจ ปีการศึกษา 2549-2551 จากฐานข้อมูลนักศึกษา วิทยาลัยราชพฤกษ์ โดยได้รับการประสานงานจากหน่วยงาน ศูนย์เทคโนโลยีสารสนเทศ วิทยาลัยราชพฤกษ์และคำแนะนำในเรื่องของการดำเนินการวิจัยที่ดี จากหน่วยงานวิจัย วิทยาลัยราชพฤกษ์

ขอขอบพระคุณผู้เชี่ยวชาญในการวิเคราะห์ข้อมูลในครั้งนี้ คือ รศ.ดร.วรพจน์ กรีสุระเดช และดร.อำนวยการพล แจ่มเจริญ ที่ให้คำแนะนำ ข้อคิดเห็น และแนวทางแก้ไขปัญหาในการวิจัย รวมถึง นายธีรภพ กลองใหญ่ ที่ให้ความร่วมมือในการจัดเตรียมข้อมูล และการใช้โปรแกรมในการสร้าง โมเดล สำหรับการวิเคราะห์ข้อมูล รวมไปถึง มีส่วนร่วมในการทดสอบความถูกต้องของโมเดลเพื่อใช้ในการทำนายผลการเรียนของนักศึกษา

เรวดี ศักดิ์คุณุชธรรม

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญตาราง	ฉ
สารบัญภาพ	ช
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหาวิจัย	1
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตของการวิจัย	3
1.4 นิยามศัพท์เฉพาะหรือนิยามศัพท์เชิงปฏิบัติการ	3
1.5 ประโยชน์ที่คาดว่าจะได้รับการวิจัย	6
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	7
2.1 ความรู้ (Knowledge)	7
2.2 ฐานความรู้ (Knowledge-based)	10
2.3 ดาต้าไมน์นิง หรือการทำเหมืองข้อมูล (Data Mining)	11
2.4 อาจารย์ที่ปรึกษา (Advisor)	29
2.5 ผลสัมฤทธิ์ทางการเรียน (Education Achievement)	31
2.6 งานวิจัยที่เกี่ยวข้อง	31
บทที่ 3 วิธีดำเนินการวิจัย	38
3.1 การศึกษาข้อมูลที่เกี่ยวข้อง	38
3.2 รวบรวม คัดเลือก และจัดเตรียมข้อมูล	39
3.3 การสร้างโมเดลในการจัดกลุ่มข้อมูล และการพยากรณ์ข้อมูล	50
3.4 การทดสอบความถูกต้อง และความน่าเชื่อถือของโมเดล โดยการใช้ข้อมูลทดสอบ	54
3.5 แปลความหมาย และนำความรู้ที่ได้มาใช้	56
บทที่ 4 ผลการวิเคราะห์ข้อมูล	57
4.1 ผลการจัดกลุ่มข้อมูล (Classification)	57
4.2 ผลการทำนาย / พยากรณ์ผลการเรียน (Predictive)	68

## สารบัญ (ต่อ)

	หน้า
บทที่ 5 สรุป อภิปรายผล และข้อเสนอแนะ	79
5.1 สรุปผลการวิจัย	79
5.2 อภิปรายผลการวิจัย	81
5.3 ข้อเสนอแนะ	82
บรรณานุกรม	83
ภาคผนวก ตัวอย่างข้อมูลนักศึกษา คณะบริหารธุรกิจ วิทยาลัยราชพฤกษ์ ปีการศึกษา 2549-2551 จำนวน 250 คนจาก 1,154 คน	85

## สารบัญตาราง

ตารางที่		หน้า
2-1	การเปรียบเทียบความถูกต้องและความผิดพลาด	37
3-1	จำนวนนักศึกษาแยกตามสาขาวิชา	40
3-2	ตัวอย่างข้อมูลส่วนตัวของนักศึกษา	41
3-3	ตัวอย่างข้อมูลการลงทะเบียน และผลการเรียนของนักศึกษา	42
3-4	โครงสร้างหลักสูตรของหลักสูตร 2 ปีต่อเนื่อง และหลักสูตร 4 ปี	42
3-5	รายวิชาแกนของแต่ละหลักสูตร	43
3-6	รายวิชาเอกบังคับของแต่ละสาขาวิชา	44
3-7	การจัดกลุ่มรายวิชาแกน	47
3-8	การเฉลี่ยผลการเรียนของกลุ่มวิชาแกนและวิชาเอกบังคับ	48
3-9	ตัวอย่างข้อมูลที่ผ่านการเตรียมข้อมูล (Data Preparation)	50
3-10	คำอธิบายของข้อมูลที่นำมาใช้ในการจัดกลุ่ม (Classification)	50
3-11	คำอธิบายของข้อมูลที่นำมาใช้ในการทำนาย (Predictive)	52



## สารบัญภาพ

ภาพที่	หน้า	
2-1	พีรามิดแสดงลำดับขั้นของความรู้	8
2-2	กระบวนการทำงานของ Data Mining	13
2-3	กระบวนการทำงานของ Classification	16
2-4	ตัวอย่างของ Decision Tree เพื่อวิเคราะห์โอกาสที่ลูกค้าบ้านเช่าจะซื้อบ้าน	18
2-5	Decision Tree เพื่อวิเคราะห์โอกาสที่ลูกค้าจะเล่นสกี	21
2-6	โมเดลของ Neurons ในระบบคอมพิวเตอร์	22
2-7	ลักษณะโครงสร้างของ Neural Network	23
2-8	รูปแบบ Back-propagation neural network	24
2-9	ตัวอย่างการทำ Clustering	26
3-1	ตัวอย่างโมเดลการทำนายผลการเรียน ด้วยเทคนิคของ Neural Network	53
4-1	รายละเอียดการนำเข้าข้อมูล (สาขาวิชาคอมพิวเตอร์ธุรกิจ)	57
4-2	แสดงความสำคัญของแอททริบิวต์ของนักศึกษาสาขาคอมพิวเตอร์ธุรกิจ	58
4-3	การจัดกลุ่มข้อมูลของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ	59
4-4	รายละเอียดการนำเข้าข้อมูล (สาขาวิชาการตลาด)	60
4-5	แสดงความสำคัญของแอททริบิวต์ของนักศึกษาสาขาการตลาด	60
4-6	การจัดกลุ่มข้อมูลของนักศึกษาสาขาวิชาการตลาด	61
4-7	รายละเอียดการนำเข้าข้อมูล (สาขาวิชาการจัดการ)	62
4-8	แสดงความสำคัญของแอททริบิวต์ของนักศึกษาสาขาการจัดการ	63
4-9	การจัดกลุ่มข้อมูลของนักศึกษาสาขาวิชาการจัดการ	64
4-10	รายละเอียดการนำเข้าข้อมูล (สาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว)	65
4-11	แสดงความสำคัญของแอททริบิวต์ของนักศึกษาสาขาวิชาการจัดการการโรงแรมฯ	66
4-12	การจัดกลุ่มข้อมูลของนักศึกษาสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว	67
4-13	รายละเอียดการนำเข้าข้อมูลของแต่ละสาขาวิชา เพื่อสร้าง โมเดลการทำนายผลการเรียน	68
4-14	ผลการวิเคราะห์ข้อมูลการทำนายด้วยข้อมูลเรียนรู้ สาขาวิชาคอมพิวเตอร์ธุรกิจ	69
4-15	การทำนายผลการเรียนด้วยข้อมูลทดสอบ สาขาวิชาคอมพิวเตอร์ธุรกิจ	70
4-16	สรุปผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียน สาขาวิชาคอมพิวเตอร์ธุรกิจ	71

## สารบัญภาพ (ต่อ)

ภาพที่	หน้า
4-17 ผลการวิเคราะห์ข้อมูลการทำนายด้วยข้อมูลเรียนรู้ สาขาวิชาการตลาด	72
4-18 ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ สาขาวิชาการตลาด	72
4-19 สรุปผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียน สาขาวิชาการตลาด	73
4-20 ผลการวิเคราะห์ข้อมูลการทำนายด้วยข้อมูลเรียนรู้ สาขาวิชาการจัดการ	74
4-21 ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ สาขาวิชาการจัดการ	75
4-22 สรุปผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียน สาขาวิชาการจัดการ	76
4-23 ผลการวิเคราะห์ข้อมูลการทำนายด้วยข้อมูลเรียนรู้สาขาวิชาการจัดการการโรงแรมฯ	76
4-24 ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ สาขาวิชาการจัดการการโรงแรมฯ	77
4-25 สรุปผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียน สาขาวิชาการจัดการการโรงแรมฯ	78

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหาวิจัย

วิทยาลัยราชพฤกษ์ เป็นสถาบันอุดมศึกษาที่มุ่งเน้นการพัฒนาคุณภาพการศึกษา ให้เป็นสถาบันแห่งการเรียนรู้ยุคใหม่ พร้อมผลิตบัณฑิตให้เป็นคนเก่ง มีความรู้ ทักษะ ความชำนาญในแต่ละสาขาวิชา ซึ่งบัณฑิตต้องเป็นบุคคลที่มุ่งมั่นจะเรียนรู้และพัฒนาตนเองอยู่ตลอดเวลา เพื่อนำความรู้ความสามารถ เป็นกำลังสำคัญในการพัฒนาประเทศชาติ ทางวิทยาลัยได้ให้ความสำคัญต่อการดูแลและให้คำแนะนำในเรื่องการเรียนของนักศึกษา ซึ่งถือเป็นปัจจัยสำคัญที่มีส่วนช่วยพัฒนาคุณภาพการศึกษาของวิทยาลัย โดยจัดให้มีอาจารย์ที่ปรึกษาคอยให้คำแนะนำแนวทางในการเรียนให้แก่ นักศึกษา

ในปัจจุบัน การตัดสินใจให้คำแนะนำของอาจารย์ที่ปรึกษา เป็นการพิจารณาจากผลการเรียนที่ผ่านมา โดยดูจากรายวิชาในแต่ละสาขาวิชาของนักศึกษาเป็นหลัก ดังนั้น หากมีเครื่องมือที่ช่วยในการวิเคราะห์ความถนัด และจุดอ่อนทางการศึกษาของนักศึกษาเป็นรายบุคคล โดยสามารถคาดการณ์ผลการศึกษาในสาขาวิชาอื่น นอกจากสาขาวิชาที่กำลังศึกษาอยู่ได้ ส่งผลให้กระบวนการตัดสินใจ ให้คำแนะนำของอาจารย์ที่ปรึกษา รวมทั้งกระบวนการพัฒนาคุณภาพการศึกษามี ประสิทธิภาพมากยิ่งขึ้น

การสกัดความรู้ที่เป็นประโยชน์ และน่าสนใจจากฐานข้อมูลขนาดใหญ่ (Knowledge Discovery from very large Database: KDD) หรือที่เรียกกันว่า คาด้าไมน์นิง (Data Mining) ซึ่งเป็นสาขาหนึ่งในวิทยาศาสตร์คอมพิวเตอร์ที่กำลังได้รับความสนใจอย่างสูงในปัจจุบัน เกิดจากในหน่วยงาน หรือองค์กรต่างๆ มีการนำเอาระบบเทคโนโลยีสารสนเทศมาใช้ ทำให้มีการจัดเก็บข้อมูลที่ง่าย และมีข้อมูลเป็นจำนวนมาก แต่ยังไม่มีการนำข้อมูลดังกล่าวไปใช้ประโยชน์สูงสุด เมื่อใช้เทคนิคคาด้าไมน์นิง ข้อมูลขนาดใหญ่จะถูกวิเคราะห์ และสกัดความรู้หรือสิ่งที่สำคัญออกมา จากนั้นจะรวบรวมความรู้ที่ได้ให้อยู่ในรูปแบบของฐานความรู้ (Knowledge Base) เพื่อนำไปใช้ประโยชน์ต่อไป โดยในปัจจุบันได้มีการนำเทคนิคคาด้าไมน์นิงไปประยุกต์ใช้ในงานด้านต่างๆ มากขึ้น ทั้งในด้านการส่งเสริมการขายสินค้าในห้างสรรพสินค้า, ด้านการวิเคราะห์เครดิตลูกค้าในธนาคาร และในด้านอื่นๆ อีกมาก ปัจจุบันเริ่มมีการนำเทคนิคคาด้าไมน์นิงมาประยุกต์กับงานด้านการศึกษา ทั้งที่ในปัจจุบันตามสถาบันการศึกษาส่วนใหญ่ มีการจัดเก็บข้อมูลของนักศึกษา แต่ยังไม่มีการนำเอาข้อมูลดังกล่าวมาวิเคราะห์ และนำไปประยุกต์ใช้ให้เกิดประโยชน์เท่าที่ควร (กฤษณะ ไวยมัย, ชิดชนก ส่องศิริ และชนาวินท์ รักธรรมานนท์, 2544: 134-142)

ดังนั้นในการวิจัยครั้งนี้ จึงสังเกตเห็นถึงความสำคัญในการนำเทคนิคต่างๆ ของดาต้าไมน์นิ่งมาประยุกต์ใช้ในการวิเคราะห์เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ โดยนำเทคนิคดาต้าไมน์นิ่งมาใช้ในการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา และนำผลที่ได้ไปประกอบการตัดสินใจ เพื่อหาแนวทางการศึกษาที่เหมาะสมสำหรับนักศึกษาแต่ละท่าน โดยข้อมูลที่น่ามาวิจัยนี้เป็นข้อมูลของนักศึกษาคณะบริหารธุรกิจ โดยแบ่งข้อมูลเป็น 2 ส่วน คือ ส่วนแรก เป็นฐานข้อมูลการลงทะเบียนเรียนและผลการเรียนในแต่ละวิชา และส่วนที่ 2 เป็นฐานข้อมูลส่วนตัวของนักศึกษา เช่น อายุ เพศ ที่อยู่ ประวัติการศึกษา เกรดเฉลี่ยสะสม เป็นต้น

## 1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 เพื่อทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ โดยใช้เทคนิคดาต้าไมน์นิ่ง

1.2.2 เพื่อนำผลที่ได้ไปวิเคราะห์ หรือหาแนวโน้มการพัฒนาการศึกษาของนักศึกษา สำหรับอาจารย์ที่ปรึกษา

## 1.3 ขอบเขตของการวิจัย

การวิจัยครั้งนี้ ผู้วิจัยมุ่งเน้นการทำนายผลการเรียน เพื่อประกอบการตัดสินใจในการให้คำแนะนำของอาจารย์ที่ปรึกษา โดยใช้เทคนิคดาต้าไมน์นิ่ง โดยมีตัวแปรที่ศึกษา ดังนี้

1.3.1 แหล่งข้อมูลที่ใช้เป็นตัวแปรต้น คือ ข้อมูลของนักศึกษาคณะบริหารธุรกิจ โดยแบ่งเป็น 2 ส่วน คือ ส่วนแรก เป็นฐานข้อมูลการลงทะเบียนเรียนและผลการเรียนในแต่ละวิชา และส่วนที่ 2 เป็นฐานข้อมูลส่วนตัวของนักศึกษา เช่น อายุ เพศ ประวัติการศึกษา เกรดเฉลี่ยสะสม เป็นต้น

1.3.2 ตัวแปรตาม คือ ผลการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา

## 1.4 นิยามศัพท์เฉพาะ หรือนิยามศัพท์เชิงปฏิบัติการ

ในการวิจัยครั้งนี้ เป็นการใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษาวิทยาลัยราชพฤกษ์ โดยมีการนำผลที่ได้ไปวิเคราะห์ หรือหาแนวโน้มการพัฒนาการศึกษาของนักศึกษา สำหรับอาจารย์ที่ปรึกษา ซึ่งมีคำศัพท์เฉพาะ หรือคำศัพท์เชิงปฏิบัติการที่เกี่ยวข้องกับงานวิจัย คือ ความรู้ (Knowledge) ดาต้าไมน์นิ่ง (Data Mining) อาจารย์ที่ปรึกษา และสัมฤทธิ์ผลทางการเรียน โดยมีการให้คำนิยามศัพท์เฉพาะดังกล่าวได้ดังนี้

#### 1.4.1 ความรู้ (Knowledge)

ความรู้ คือ สารสนเทศที่นำไปสู่การปฏิบัติ เป็นเนื้อหาข้อมูล ซึ่งประกอบด้วยข้อเท็จจริง ความคิดเห็น ทฤษฎี หลักการ รูปแบบ กรอบความคิด หรือข้อมูลอื่นๆ ที่มีความจำเป็น และเป็นกรอบของการผสมผสานระหว่างประสบการณ์ ค่านิยม ความรู้ในบริบท สำหรับการประเมินค่า และการนำเอาประสบการณ์กับสารสนเทศใหม่ๆ มาผสมรวมเข้าด้วยกัน (จดหมายข่าวจากทีมงาน KMUTI\_KM: 2549)

มีผู้เชี่ยวชาญชาวญี่ปุ่น ชื่อ Hideo Yamazaki ได้ให้คำจำกัดความของ “ความรู้” ในรูปของปิรามิด ซึ่งแสดงถึงลำดับขั้นของความรู้ ดังนี้ (การนิคมอุตสาหกรรมแห่งประเทศไทย, 2551)

1) ข้อมูล (Data) เป็นข้อเท็จจริง ข้อมูลดิบ หรือตัวเลขต่างๆ ที่ยังไม่ได้ผ่านกระบวนการประมวลผล หรือเป็นกลุ่มของข้อมูลดิบที่เกิดขึ้นจากการปฏิบัติงาน

2) สารสนเทศ (Information) เป็นข้อมูลที่ผ่านกระบวนการประมวลผล โดยรวบรวมและสังเคราะห์ วิเคราะห์ เอาเฉพาะข้อมูลที่มีความหมาย และเป็นประโยชน์ต่องานที่ทำ เพื่อนำมาใช้ประโยชน์ในการบริหารจัดการ และตัดสินใจ โดยมักจะอยู่ในรูปของข้อมูลที่วัดได้ หรือจับต้องได้ สารสนเทศอาจมีข้อจำกัดในเรื่องช่วงเวลาที่ใช้ และขอบข่ายของงานที่จะนำมาใช้

3) ความรู้ (Knowledge) เป็นผลจากการจัดเกลา และเลือกใช้สารสนเทศ โดยมีการจัดระบบความคิดเสียใหม่ ให้เป็น “ความรู้และความเชี่ยวชาญเฉพาะเรื่อง” นั่นก็คือ สารสนเทศที่ผ่านกระบวนการคิดเปรียบเทียบเชื่อมโยงกับความรู้อื่นๆ จนเกิดความเข้าใจ และนำไปใช้ประโยชน์ในการสรุป และตัดสินใจในสถานการณ์ต่างๆ ได้โดยไม่จำกัดช่วงเวลา หรือเป็นสารสนเทศที่ก่อให้เกิดประโยชน์กับเราในการนำไปใช้งาน

4) ความเฉลียวฉลาด (Wisdom) เป็นการนำเอาความรู้ต่างๆ มาบูรณาการเข้าด้วยกัน เพื่อใช้ให้เกิดประโยชน์ต่อการทำงานในสาขาวิชาชีพต่างๆ

#### 1.4.2 ดาต้าไมนิ่ง (Data Mining)

เนื่องจากในปัจจุบันองค์กรธุรกิจส่วนใหญ่ มีการนำเอาเทคโนโลยีการจัดเก็บข้อมูลมาใช้เป็นจำนวนมาก จึงมักเผชิญกับปัญหาของข้อมูลดิบจำนวนมาก แต่ข้อมูลที่ประยุกต์ใช้ได้นั้นมีน้อย ดาต้าไมนิ่งจึงเป็นสาขาหนึ่งที่น่ามาใช้ประยุกต์ใช้อย่างแพร่หลาย เนื่องจากดาต้าไมนิ่ง สามารถดึงความรู้ออกมาจากข้อมูลจำนวนมากที่ถูกเก็บสะสม และซ่อนไว้

ดาต้าไมนิ่ง คือกระบวนการในการดึงความรู้ (Knowledge) ออกจากฐานข้อมูลขนาดใหญ่ เพื่อนำไปใช้ประโยชน์ในการตัดสินใจ เป็นการวิเคราะห์ค้นหาสารสนเทศหรือความรู้

(Information/Knowledge Discovery) ซึ่งอาจจะไม่ได้ปรากฏอยู่จริงในฐานข้อมูลก็ได้ โดยใช้เทคโนโลยีทางคอมพิวเตอร์ด้านต่างๆ เช่น ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) เข้ามาช่วยในการค้นหาสารสนเทศหรือความรู้ คำค้นหาใดหนึ่งมีหลายโมเดล แต่ละโมเดลจะมีสูตรทางธุรกิจ (Business Formula) เข้ามาเกี่ยวข้องและให้ผลลัพธ์ในรูปแบบที่แตกต่างกัน เช่น เป็นกฎ If-then หรือเป็นแผนภูมิการตัดสินใจ (Decision Trees) เป็นต้น (วิจัย กองศาสนะ, 2546)

1.4.2.1 เทคนิคการทำงานของคำค้นหา (Data Mining) โดยทั่วไป มีดังนี้ (เอกสารประกอบรายวิชาการเปรียบเทียบวิธีวิจัยทางเทคโนโลยีสารสนเทศ, ม.ขอนแก่น)

- 1) การจำแนกหมวดหมู่ หรือการวิเคราะห์แบบรวมกลุ่ม (Cluster Analysis)
- 2) การวิเคราะห์ความแปลกแยก ความผิดปกติไปจากข้อมูลทั่วไปในฐานเดียวกัน (Outlier Analysis)
- 3) การจัดกลุ่ม และพยากรณ์ (Classification and Prediction)
- 4) การหาความสัมพันธ์ของข้อมูล (Association / Correlation and Causality)
- 5) การวิเคราะห์แนวโน้ม และวิวัฒนาการ (Trend and Evolution Analysis)

1.4.2.2 ขั้นตอนการทำคำค้นหา (เอกสารประกอบรายวิชาการเปรียบเทียบวิธีวิจัยทางเทคโนโลยีสารสนเทศ, ม.ขอนแก่น)

- 1) รวบรวมข้อมูล (Data integration) โดยรวมข้อมูลจากหลากหลายแหล่งเก็บข้อมูล
- 2) คัดเลือกข้อมูล (Data selection) โดยเลือกเฉพาะข้อมูลที่ต้องการนำมาวิเคราะห์
- 3) กลั่นกรองข้อมูล (Data cleaning) โดยนำข้อมูลที่ไม่มีค่า ข้อมูลที่ขาดหาย ข้อมูลขยะ และข้อมูลที่ไม่แน่นอนออกไป
- 4) แปลรูปข้อมูล (Transformation) โดยรวมกลุ่มข้อมูลและแปรข้อมูลให้อยู่ในรูปแบบที่พร้อมจะนำไปใช้ในการวิเคราะห์ หรือการทำคำค้นหา
- 5) การทำคำค้นหา (Data mining) โดยการประยุกต์เทคนิคการทำงานต่างๆ ของกระบวนการคำค้นหา เพื่อให้ได้ความรู้หรือสิ่งที่สนใจออกจากข้อมูลที่ได้แปรรูปไว้แล้ว เช่น ร้านค้าสะดวกซื้อแห่งหนึ่ง เก็บข้อมูลการขายรายวัน และนำมาวิเคราะห์โดยใช้เทคนิคคำค้นหาในการหาความสัมพันธ์ของข้อมูล (Association) พบว่า เมื่อลูกค้าซื้อกาแฟ ก็มักจะซื้อขนมปังควบคู่ด้วย ดังนั้น ทางบริษัทสามารถนำกลยุทธ์นี้ไปปรับวิธีการขาย เพื่อดึงความสนใจของลูกค้ามากขึ้น
- 6) ประเมินรูปแบบ (Pattern evaluation) ซึ่งรูปแบบที่น่าสนใจควรเป็นสิ่งที่ผู้ใช้เชื่อถือได้ หรือยืนยันสมมติฐานที่มีเหตุผล

7) นำเสนอความรู้สู่ผู้ใช้ (Knowledge presentation) เป็นการนำเอาความรู้ที่ได้จากการทำคิต้าไม้นั้นไปประยุกต์ใช้ในการตัดสินใจ เพื่อนำไปวางแผน และปรับกลยุทธ์ทางธุรกิจให้มีประสิทธิภาพมากขึ้น

### 1.4.3 อาจารย์ที่ปรึกษา

จากคู่มืออาจารย์ที่ปรึกษา วิทยาลัยราชพฤกษ์ กล่าวว่า ระบบอาจารย์ที่ปรึกษาถือเป็นสิ่งจำเป็นอย่างยิ่งในการช่วยเหลือนักศึกษาแก้ไขปัญหาต่างๆ โดยอาจารย์ที่ปรึกษาจะให้คำแนะนำ และคำปรึกษาเพื่อให้นักศึกษามีความพร้อมทั้งร่างกายและจิตใจในการเรียน การร่วมกิจกรรมต่างๆ และการใช้ชีวิตในสังคมได้อย่างมีความสุข

#### 1.4.3.1 วัตถุประสงค์ของอาจารย์ที่ปรึกษา

- 1) เพื่อให้คำปรึกษาแนะนำทางด้านวิชาการเกี่ยวกับหลักสูตร การเลือกวิชาเรียน การลงทะเบียนเรียน วิธีการเรียน และการวัดผล เพื่อให้นักศึกษาเลือกเรียน ได้ตรงกับความสามารถ
- 2) เพื่อให้คำแนะนำเกี่ยวกับระเบียบข้อบังคับต่างๆ ของวิทยาลัย รวมถึงการให้คำแนะนำเกี่ยวกับการบริการ และสวัสดิการต่างๆ ที่ทางวิทยาลัยจัดบริการให้กับนักศึกษา
- 3) เพื่อเป็นที่พึ่งของนักศึกษา ในการรับฟังปัญหา และให้คำปรึกษา ข้อแนะนำแก่นักศึกษาในการแก้ไขปัญหาต่างๆ ได้
- 4) เพื่อช่วยพัฒนาบุคลิกภาพของนักศึกษาทางด้านวิชาการ, วิชาชีพ, การเข้าสังคม รวมทั้งด้านร่างกายและจิตใจ เพื่อเตรียมความพร้อมให้นักศึกษาก้าวออกไปสู่สังคมได้เป็นอย่างดี
- 5) เพื่อสร้างความสัมพันธ์ ความเข้าใจอันดี ระหว่างนักศึกษา คณาจารย์ และวิทยาลัย

#### 1.4.3.2 บทบาทหน้าที่ของอาจารย์ที่ปรึกษา

- 1) ด้านวิชาการ มีการพิจารณาคำร้องของนักศึกษา และดำเนินการตามระเบียบของวิทยาลัย มีการให้คำปรึกษาแนะนำนักศึกษาเกี่ยวกับหลักสูตรและการเลือกวิชาเรียน ดูแลเรื่องการลงทะเบียนของนักศึกษา รวมไปถึงให้คำปรึกษาแนะนำหรือให้การช่วยเหลือ เมื่อผลการเรียนของนักศึกษาต่ำลง และให้คำปรึกษาแนะนำเกี่ยวกับการศึกษาต่อในระดับที่สูงขึ้น
- 2) ด้านบริการและพัฒนานักศึกษา มีการให้คำปรึกษาที่เกี่ยวกับปัญหาส่วนตัว เช่น สุขภาพอนามัยทั้งสุขภาพกายและสุขภาพจิต รวมไปถึงปัญหาสังคม เช่น การปรับตัว การคบเพื่อน การพัฒนาบุคลิกภาพ ความประพฤติ และจริยธรรม ปัญหาด้านอาชีพ เป็นต้น
- 3) ด้านอื่นๆ เช่น ประสานงานกับอาจารย์ผู้สอน และฝ่ายต่างๆ ที่เกี่ยวข้อง รวมไปถึงกำหนดเวลาให้นักศึกษาเข้าพบ เพื่อขอคำปรึกษา แนะนำอย่างสม่ำเสมอ และเก็บข้อมูลรายละเอียดของนักศึกษาที่อยู่ในความรับผิดชอบ เพื่อใช้เป็นข้อมูลพื้นฐานสำหรับให้คำปรึกษา พร้อมให้ความร่วมมือ และความเข้าใจอันดีระหว่างนักศึกษา คณาจารย์ สาขาวิชา คณะฯ และวิทยาลัย

#### 1.4.4 สัมฤทธิ์ผลทางการเรียน

สัมฤทธิ์ผลทางการเรียน (Education achievement) หมายถึง ผลที่เกิดจากการจัดการศึกษา และเป็นปัจจัยหนึ่งที่สามารถบอกถึงคุณภาพการศึกษา ดังที่ กู๊ด (Good 1973 :7 ,อ้างถึงใน ชุติพร ฤทธิเดช 2547 : 56) กล่าวถึงสัมฤทธิ์ผลทางการเรียน สรุปได้ว่า หมายถึงความรู้หรือทักษะอันเกิดจากการเรียนรู้ที่ได้เรียนมาแล้วที่ได้จากผลการสอนของอาจารย์ผู้สอน ซึ่งอาจพิจารณาจากคะแนนสอบที่กำหนด ให้คะแนนที่ได้จากงานที่ครูมอบหมายให้หรือทั้งสองอย่าง จัดว่าเป็นการตรวจสอบระดับความสามารถของนักศึกษาว่าเรียนแล้วมีความรู้เท่าใด สามารถวัดได้โดยการใช้แบบทดสอบต่าง ๆ เช่น ใช้ข้อสอบวัดผลสัมฤทธิ์ ข้อสอบวัดภาคปฏิบัติ เป็นต้น

จากความหมายข้างต้นสรุปได้ว่า สัมฤทธิ์ผลทางการเรียน หมายถึง ผลที่ได้จากการวัดความรู้ที่ได้จากการเรียนรู้ในเนื้อหาสาระที่ได้จากผลการสอนของอาจารย์ผู้สอน เพื่อตรวจสอบว่ามีความเข้าใจและเกิดการเรียนรู้ในเนื้อหานั้นๆ มากขึ้นเท่าใด มีความสามารถด้านไหน ซึ่งสามารถทำการวัดผลได้จากแบบทดสอบวัดสัมฤทธิ์ผลทางการเรียนในลักษณะต่างๆ ทั้งทางทฤษฎี และเชิงปฏิบัติ เพื่อบอกถึงคุณภาพการศึกษา

#### 1.5 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1.5.1 นำผลการวิจัยที่ได้ไปประกอบการตัดสินใจ ให้คำแนะนำของอาจารย์ที่ปรึกษา เพื่อหาแนวทางในการพัฒนาการศึกษาของนักศึกษาให้มีคุณภาพ

1.5.2 ขยายผลการนำเทคนิคดาต้าไมน์นิ่งมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลด้านต่างๆ เพื่อเป็นแนวทางในการพัฒนาคุณภาพการศึกษาของนักศึกษาต่อไป



## บทที่ 2

### เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยเรื่อง การใช้เทคนิคดาต้าไมนนิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ ผู้วิจัยได้ทำการศึกษาถึงแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง เพื่อใช้เป็นแนวทางสำหรับการดำเนินงานวิจัย ประกอบไปด้วย

- 2.1 ความรู้ (Knowledge)
- 2.2 ฐานความรู้ (Knowledge-based)
- 2.3 ดาต้าไมนนิ่ง หรือการทำเหมืองข้อมูล (Data Mining)
- 2.4 อาจารย์ที่ปรึกษา (Advisor)
- 2.5 สัมฤทธิ์ผลทางการเรียน (Education achievement)
- 2.6 งานวิจัยที่เกี่ยวข้อง

#### 2.1 ความรู้ (Knowledge)

เนื่องจากมีผู้ให้คำนิยามและความหมายของ “ความรู้ (Knowledge)” ไว้เป็นจำนวนมาก ในการวิจัยครั้งนี้ จึงขอกล่าวถึงความหมายและคำนิยามของ “ความรู้ (Knowledge)” จากแหล่งข้อมูลต่างๆ ดังนี้

ความรู้ คือ สารสนเทศที่นำไปสู่การปฏิบัติ เป็นเนื้อหาข้อมูล ซึ่งประกอบด้วยข้อเท็จจริง ความคิดเห็น ทฤษฎี หลักการ รูปแบบ กรอบความคิด หรือข้อมูลอื่นๆ ที่มีความจำเป็น และเป็นกรอบของการผสมผสานระหว่างประสบการณ์ คำนิยาม ความรู้ในบริบท เพื่อใช้สำหรับการประเมินค่า และการนำเอาประสบการณ์กับสารสนเทศใหม่ๆ มาผสมรวมเข้าด้วยกัน (จดหมายข่าวจากทีมงาน KMUTI\_KM: 2549)

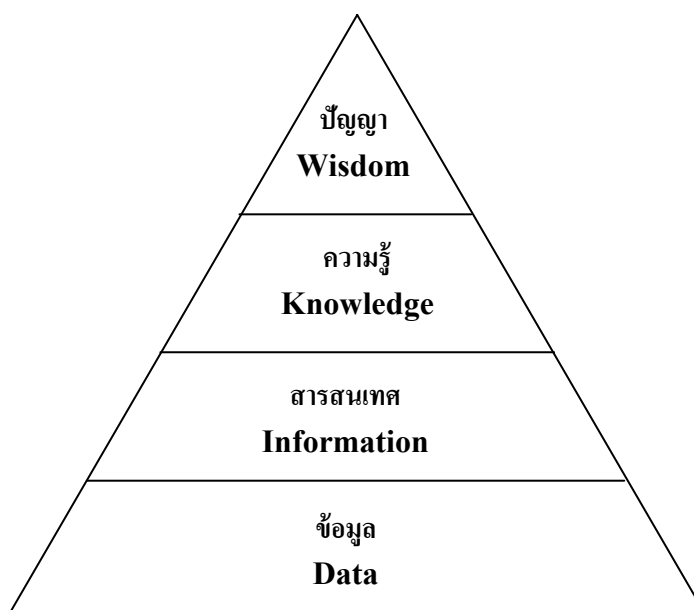
ทางด้านผู้เชี่ยวชาญชาวญี่ปุ่น ชื่อ Hideo Yamazaki ได้ให้คำจำกัดความของ “ความรู้” ในรูปของ พีระมิด ซึ่งแสดงถึงลำดับขั้นของความรู้ ดังนี้ (การนิคมอุตสาหกรรมแห่งประเทศไทย, 2551)

- 1) ข้อมูล (Data) เป็นข้อเท็จจริง ข้อมูลดิบ หรือตัวเลขต่างๆ ที่ยังไม่ได้ผ่านกระบวนการประมวลผล หรือเป็นกลุ่มของข้อมูลดิบที่เกิดขึ้นจากการปฏิบัติงาน
- 2) สารสนเทศ (Information) เป็นข้อมูลที่ผ่านกระบวนการประมวลผล โดยรวบรวมและสังเคราะห์ วิเคราะห์ เอาเฉพาะข้อมูลที่มีความหมาย และเป็นประโยชน์ต่องานที่ทำ เพื่อนำมาใช้

ประโยชน์ในการบริหารจัดการ และตัดสินใจ โดยมักจะอยู่ในรูปของข้อมูลที่วัดได้ หรือจับต้องได้ สารสนเทศอาจมีข้อจำกัดในเรื่องช่วงเวลาที่ใช้ และขอบข่ายของงานที่จะนำมาใช้

3) ความรู้ (Knowledge) เป็นผลจากการจัดเกลา และเลือกใช้สารสนเทศ โดยมีการจัดระบบความคิดเสียใหม่ ให้เป็น “ความรู้และความเชี่ยวชาญเฉพาะเรื่อง” นั่นก็คือ สารสนเทศที่ผ่านกระบวนการคิดเปรียบเทียบเชื่อมโยงกับความรู้อื่นๆ จนเกิดความเข้าใจ และนำไปใช้ประโยชน์ในการสรุป และตัดสินใจในสถานการณ์ต่างๆ ได้โดยไม่จำกัดช่วงเวลา หรือเป็นสารสนเทศที่ก่อให้เกิดประโยชน์กับเราในการนำไปใช้งาน

4) ความเฉลียวฉลาด (Wisdom) เป็นการนำเอาความรู้ต่างๆ มาบูรณาการเข้าด้วยกัน เพื่อใช้ให้เกิดประโยชน์ต่อการทำงานในสาขาวิชาชีพต่างๆ



ภาพที่ 2-1 พีระมิดแสดงลำดับขั้นของความรู้ (การนิคมอุตสาหกรรมแห่งประเทศไทย, 2551)

คำว่า ความรู้ (Knowledge) ในทัศนะของฮอสเปอร์ (อ้างถึงในมาโนช เวชพันธ์ 2532, 15-16) ถือว่าเป็นขั้นแรกของพฤติกรรมที่เกี่ยวข้องกับความสามารถในการจดจำ ซึ่งอาจจะทำได้โดยการนึกได้ การมองเห็น ได้ยิน หรือได้ฟังความรู้นี้ ซึ่งเป็นหนึ่งในขั้นตอนของการเรียนรู้ โดยประกอบไปด้วยคำจำกัดความหรือความหมาย ข้อเท็จจริง ทฤษฎี กฎ โครงสร้าง วิธีการแก้ไขปัญหา และมาตรฐาน เป็นต้น ซึ่งอาจกล่าวได้ว่า ความรู้เป็นเรื่องของการจำอะไรได้ ระลึกได้ โดยไม่จำเป็นต้องใช้ความคิดที่ซับซ้อนหรือใช้ความสามารถของสมองมากนัก ด้วยเหตุดังกล่าวการจำได้จึงถือว่าเป็นกระบวนการที่สำคัญในทางจิตวิทยา และเป็นขั้นตอนที่นำไปสู่พฤติกรรมที่ก่อให้เกิดความเข้าใจ

การนำความรู้ไปใช้ในการวิเคราะห์ การสังเคราะห์ การประเมินผล ซึ่งเป็นขั้นตอนที่ได้ใช้ความคิดและความสามารถทางสมองมากขึ้นเป็นลำดับ ส่วนความเข้าใจ (Comprehension) นั้น ฮอสเปอร์ชี้ให้เห็นว่า เป็นขั้นตอนต่อมาจากความรู้ โดยเป็นขั้นตอนที่จะต้องใช้ความสามารถของสมองและทักษะในขั้นที่สูงขึ้น จนถึงระดับของการสื่อความหมาย ซึ่งอาจจะเป็นไปได้โดยการใช้ปากเปล่า ข้อเขียน ภาษา หรือการใช้สัญลักษณ์ โดยมักเกิดขึ้นหลังจากที่บุคคลได้รับข่าวสารต่าง ๆ แล้ว อาจจะใช้การฟัง การเห็น การได้ยิน หรือเขียน แล้วแสดงออกมาในรูปของการใช้ทักษะหรือการแปลความหมายต่าง ๆ เช่น การบรรยายข่าวสารที่ได้ยินมาโดยคำพูดของตนเอง หรือการแปลความหมายจากภาษาหนึ่งไปเป็นอีกภาษาหนึ่ง โดยคงความหมายเดิมเอาไว้ หรืออาจเป็นการแสดงความคิดเห็นหรือให้ข้อสรุปหรือการคาดคะเนก็ได้

เบนจามิน บลูม และคณะ (Benjamin S. Bloom อ้างถึงใน อักษร สวัสดิ์ 2542, 26-28) ได้เสนอแนวคิดเกี่ยวกับการรับรู้หรือพุทธิพิสัย (cognitive domain) ของคน ว่าประกอบด้วยความรู้ตามระดับต่าง ๆ รวม 6 ระดับ ซึ่งอาจพิจารณาจากระดับความรู้ในขั้นต่ำไปสู่ระดับของความรู้ในระดับที่สูงขึ้นไป โดยบลูมและคณะ ได้แจกแจงรายละเอียดของแต่ละระดับไว้ดังนี้

1) ความรู้ (Knowledge) หมายถึง การเรียนรู้ที่เน้นถึงการจำ และการระลึกได้ถึงความคิด วัตถุ และปรากฏการณ์ต่าง ๆ ซึ่งเป็นความจำที่เริ่มจากสิ่งง่าย ๆ ที่เป็นอิสระแก่กัน ไปจนถึงความจำในสิ่งที่ยู่ยากซับซ้อนและมีความสัมพันธ์ระหว่างกัน

2) ความเข้าใจหรือความคิดรวบยอด (Comprehension) เป็นความสามารถทางสติปัญญาในการขยายความรู้ ความจำ ให้กว้างออกไปจากเดิมอย่างสมเหตุสมผล การแสดงพฤติกรรมเมื่อเผชิญกับสื่อความหมาย และความสามารถในการแปลความหมาย การสรุปหรือการขยายความสิ่งใดสิ่งหนึ่ง

3) การนำไปปรับใช้ (Application) เป็นความสามารถในการนำความรู้ (knowledge) ความเข้าใจหรือความคิดรวบยอด (comprehension) ในเรื่องใด ๆ ที่มีอยู่เดิม ไปแก้ไขปัญหาที่แปลกใหม่ของเรื่องนั้น โดยการใช้ความรู้ต่าง ๆ โดยเฉพาะอย่างยิ่งวิธีการกับความคิดรวบยอดมาผสมผสานกับความสามารถในการแปลความหมาย การสรุปหรือการขยายความสิ่งนั้น

4) การวิเคราะห์ (Analysis) เป็นความสามารถ และทักษะที่สูงกว่าความเข้าใจ และการนำไปปรับใช้ โดยมีลักษณะเป็นการแยกแยะสิ่งที่จะพิจารณาออกเป็นส่วนย่อย ๆ ที่มีความสัมพันธ์กัน รวมทั้งการสืบค้นความสัมพันธ์ของส่วนต่าง ๆ เพื่อดูว่าส่วนประกอบปลีกย่อยนั้นสามารถเข้ากันได้หรือไม่ อันจะช่วยให้เกิดความเข้าใจต่อสิ่งหนึ่งสิ่งใดอย่างแท้จริง

5) การสังเคราะห์ (Synthesis) เป็นความสามารถในการรวบรวมส่วนประกอบย่อย ๆ หรือส่วนใหญ่ ๆ เข้าด้วยกันเพื่อให้เป็นเรื่องราวอันหนึ่งอันเดียวกัน การสังเคราะห์จะมีลักษณะของการ

เป็นกระบวนการรวบรวมเนื้อหาสาระของเรื่องต่างๆ เข้าไว้ด้วยกันเพื่อสร้างรูปแบบหรือโครงสร้างที่ยังไม่ชัดเจนขึ้นมาก่อน ถือว่าเป็นกระบวนการที่ต้องอาศัยความคิดสร้างสรรค์ภายในขอบเขตของสิ่งที่กำหนดให้

6) การประเมินผล (Evaluation) เป็นความสามารถในการตัดสินใจเกี่ยวกับความคิด ค่านิยม ผลงาน คำตอบ รวมถึงวิธีการและเนื้อหาสาระ เพื่อวัตถุประสงค์บางอย่าง โดยมีการกำหนดเกณฑ์ (criteria) เป็นฐานในการพิจารณาตัดสินใจ ซึ่งการประเมินผลนั้น จัดได้ว่าเป็นขั้นตอนที่สูงที่สุดของพุทธิลักษณะ (characteristics of cognitive domain) ที่ต้องใช้ความรู้ความเข้าใจ การนำไปปรับใช้ การวิเคราะห์และการสังเคราะห์เข้ามาพิจารณาประกอบกันเพื่อทำการประเมินผลสิ่งหนึ่งสิ่งใด

จากคำนิยามและความหมายของ “ความรู้ (Knowledge)” ที่ได้อ้างถึงมานั้น สามารถสรุปได้ว่าความรู้ คือสิ่งที่ได้จากการเรียนรู้ ด้วยวิธีต่างๆ ทั้งการจดจำเนื้อหาสาระ ความเข้าใจในเนื้อหาเหล่านั้น รวมไปถึงสิ่งที่ได้จากการวิเคราะห์ สังเคราะห์ และการประเมินผลจากสิ่งที่ได้เรียนรู้มา หรือกล่าวได้ว่าเป็นองค์ความรู้ใหม่ ที่สามารถนำไปปรับใช้ให้เกิดประโยชน์ เช่นองค์ความรู้ที่ได้ไปใช้ในการประกอบการตัดสินใจ หรือแก้ไขปัญหาต่างๆ ต่อไป

## 2.2 ฐานความรู้ (Knowledge-based)

มีผู้เชี่ยวชาญได้ให้ความหมายของ “ฐานความรู้ (Knowledge-based)” ไว้เป็นจำนวนมาก ในการวิจัยครั้งนี้ จึงขอกล่าวถึงความหมายของ “ฐานความรู้ (Knowledge-based)” จากแหล่งข้อมูลต่างๆ ดังนี้

ฐานความรู้ หมายถึง ส่วนของโปรแกรมในระบบผู้เชี่ยวชาญที่รวมความรู้ของผู้เชี่ยวชาญในแต่ละสาขาวิชาไว้ ซึ่งปกติแล้วจะอยู่ในกฎของ IF/THEN เช่น ถ้า (IF) ถังน้ำมีแรงอัดมากเกินกว่า 600 ปอนด์ต่อตารางนิ้ว แล้ว (THEN) จะมีเสียงเตือนขึ้นมาให้ทราบ (กิดานันท์ มลิทอง, 2539)

ฐานความรู้ อีกความหมายหนึ่ง คือ เป็นการนำเสนอความรู้ในรูปของโปรแกรมคอมพิวเตอร์ในขอบเขตความรู้ใดความรู้หนึ่ง โดยจัดเป็นขั้นเป็นตอนให้สามารถค้นคว้าหาความรู้นั้นได้อย่างมีการเชื่อมโยงและมีปฏิสัมพันธ์ ผู้ใช้สามารถศึกษาหาความรู้ได้ด้วยตนเองเพื่อให้ได้สิ่งที่สนใจไปรู้ (นงคราญ ใจปัญญา, 2542: 30)

ฐานความรู้จัดว่าเป็นหัวใจของระบบฐานความรู้ เป็นส่วนที่เก็บกฎและความสัมพันธ์ต่างๆ ที่เกี่ยวข้องกับปัญหา ซึ่งอาจเก็บอยู่ในรูปแบบต่างๆ อย่างเช่น if X then Y โดยมีกฎเกณฑ์จำนวนมาก อย่างไรก็ตาม การแทนความรู้เป็นเรื่องที่ซับซ้อน เนื่องจากรูปแบบของความรู้ไม่มีรูปแบบที่แน่นอนตายตัว (สมชาย นำประเสริฐชัย, 2544)

ฐานความรู้เป็นส่วนหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence) ซึ่งหมายถึง ที่จัดเก็บความรู้ที่หลากหลายซึ่งถูกรวบรวมจากแหล่งต่างๆมาไว้ เพื่อให้ระบบคอมพิวเตอร์ สามารถค้นคืนสำหรับนำไปประกอบการวินิจฉัยหรือใช้เป็นเหตุผลประกอบการทำงาน หรือจัดการปัญหาที่เกิดขึ้นได้ (สำราญ กมลายุทธ, 2545)

ฐานความรู้ หมายถึง แหล่งความรู้ที่ถูกรวบรวมเอาไว้เป็นหมวดหมู่ เพื่อให้สามารถค้นหาได้อย่างมีระบบ โดยมีขอบเขตเฉพาะเจาะจงและลึกในเรื่องใดเรื่องหนึ่งเพื่อความสะดวกในการศึกษาค้นคว้า มีการเชื่อมโยงปฏิสัมพันธ์ภายในฐานความรู้นั้นๆ เพื่ออำนวยความสะดวกในการศึกษาหาความรู้ด้วยตนเองของผู้ใช้ (สุรัสวดี อัจฉนนท์ลา, 2549)

จากความหมายของ “ฐานความรู้ (Knowledge-based)” ที่ได้รวบรวมมานั้น สามารถสรุปได้ว่า ฐานความรู้ เป็นแหล่งที่เก็บรวบรวมความรู้ต่างๆ ที่ผ่านการวิเคราะห์มาแล้ว โดยจะจัดเก็บความรู้ที่มีความเฉพาะเจาะจงในเรื่องใดเรื่องหนึ่ง ซึ่งจะจัดเก็บในรูปแบบของกฎความสัมพันธ์ของข้อมูลในเนื้อหาสาระที่ต้องการจะศึกษา ซึ่งสามารถใช้ฐานความรู้ที่ได้ไปประกอบการตัดสินใจในเรื่องนั้นๆ ได้อย่างมีประสิทธิภาพ

## 2.3 ดาต้าไมนิ่ง หรือการทำเหมืองข้อมูล (Data Mining)

### 2.3.1 ความหมายของการทำเหมืองข้อมูล (Data Mining)

มีผู้ให้คำจำกัดความของดาต้าไมนิ่ง หรือการทำเหมืองข้อมูล (Data Mining) ไว้เป็นจำนวนมาก ในการวิจัยครั้งนี้จึงได้ทำการรวบรวมคำจำกัดความจากแหล่งข้อมูลต่างๆ ไว้ดังนี้

การทำเหมืองข้อมูล หมายถึง ขบวนการทำงาน(Process) ที่สกัดข้อมูล (Extract data) ออกจากฐานข้อมูลขนาดใหญ่ (Large Information) เพื่อให้ได้สารสนเทศ (Useful Information) บางมุมที่เรายังไม่รู้ (Unknown data) โดยเป็นสารสนเทศที่มีเหตุผล (Valid) และสามารถนำไปใช้ได้ (Actionable) ซึ่งเป็นสิ่งสำคัญในการที่จะช่วยการตัดสินใจในการทำธุรกิจ (Data Mining & Data Exploration Lab จากคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2547)

Berry , Michael J.A. and Linn off (อ้างถึงใน ฉวีรุณี เจริญเกียรติบวร, 2549 : 4) กล่าวว่า ดาต้าไมนิ่ง คือ การค้นหาความสัมพันธ์และรูปแบบทั้งหมด ที่มีอยู่จริงในฐานข้อมูล แต่ถูกซ่อนไว้ภายในข้อมูลจำนวนมาก โดยดาต้าไมนิ่งจะเหมาะกับการแก้ปัญหาบางชนิดเท่านั้น อย่างเช่น ปัญหาที่ต้องใช้เหตุผลในการแก้ หรือ ปัญหาที่เกี่ยวข้องกับเศรษฐศาสตร์ และ การเงินการศึกษา เป็นต้น มีเทคนิคต่าง ๆ ของดาต้าไมนิ่งที่ใช้ในการแก้ปัญหาอยู่หลายเทคนิค ซึ่งจะไม่มีเทคนิคใด

เลยที่สามารถแก้ปัญหของดาต้าไมน์นิ่งได้ทุกปัญหา ดังนั้น ความหลากหลายของเทคนิคเป็นสิ่งที่จะต้องนำมาไปสู่วิธีการแก้ปัญหที่ดีที่สุดของดาต้าไมน์นิ่ง

ดาต้าไมน์นิ่ง เป็นกระบวนการของการกลั่นกรองสารสนเทศ (Information) ที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่ เพื่อใช้ในทำนายแนวโน้ม และพฤติกรรม โดยอาศัยข้อมูลในอดีต และเพื่อใช้สารสนเทศเหล่านี้ในการสนับสนุนการตัดสินใจทางธุรกิจ (คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศิลปากร, 2551)

ดาต้าไมน์นิ่ง คือ ชุดซอฟต์แวร์ สำหรับวิเคราะห์ข้อมูลที่ได้ถูกออกแบบมา เพื่อระบบสนับสนุนการตัดสินใจของผู้ใช้ ซึ่งเป็นซอฟต์แวร์ที่สมบูรณ์ทั้งเรื่องการค้นหา การทำรายงาน และโปรแกรมในการจัดการ ที่เรียกว่า Executive Information System (EIS) หรือระบบข้อมูลสำหรับการตัดสินใจในการบริหาร ซึ่งเป็นเครื่องมือชิ้นใหม่ที่สามารถค้นหาข้อมูลในฐานข้อมูลขนาดใหญ่ หรือข้อมูลที่เป็นประโยชน์ในการบริหาร ซึ่งเป็นการเพิ่มคุณค่าให้กับฐานข้อมูลที่มีอยู่ (เอกสารประกอบรายวิชา การจัดการฐานข้อมูล มหาวิทยาลัยขอนแก่น, 2552)

จากคำจำกัดความของการทำเหมืองข้อมูล (Data Mining) ที่ได้รวบรวมมานั้น สามารถสรุปได้ว่า การทำเหมืองข้อมูล (Data Mining) คือกระบวนการในการสกัดสารสนเทศ เพื่อให้ได้ความรู้ หรือข้อมูลบางมุมที่ซ่อนเร้นอยู่ในสารสนเทศนั้นๆ หรือฐานข้อมูลขนาดใหญ่ โดยอาศัยข้อมูลในอดีต เพื่อนำความรู้ที่ได้ไปช่วยประกอบการตัดสินใจของผู้บริหารในด้านต่างๆ

### 2.3.2 ทำไมจึงต้องมี Data Mining

1) ข้อมูลที่ถูกเก็บไว้ในฐานข้อมูล หากเก็บไว้เฉย ๆ ก็จะไม่เกิดประโยชน์ ดังนั้นจึงต้องมีการสกัดสารสนเทศไปใช้ การสกัดสารสนเทศ หมายถึง การคัดเลือกข้อมูลออกมาใช้งานในส่วนที่เราต้องการ

2) ในอดีตเราใช้คนเป็นผู้สืบค้นข้อมูลต่าง ๆ ในฐานข้อมูลซึ่งผู้สืบค้นจะทำการสร้างเงื่อนไขขึ้นมาตามภูมิปัญญาของผู้สืบค้น

3) ในปัจจุบันการวิเคราะห์ข้อมูลจากฐานข้อมูลเดียวอาจไม่ให้ความรู้เพียงพอ และลึกซึ้งสำหรับการดำเนินงานภายใต้ภาวะที่มีการแข่งขันสูง มีการเปลี่ยนแปลงที่รวดเร็ว จึงจำเป็นที่จะต้องรวบรวมฐานข้อมูลหลาย ๆ ฐานข้อมูลเข้าด้วยกัน เรียกว่า “คลังข้อมูล” (Data Warehouse)

ดังนั้นเราจึงจำเป็นต้องใช้ดาต้าไมน์นิ่ง (Data Mining) ในการดึงข้อมูลจากฐานข้อมูลที่มีขนาดใหญ่ เพื่อที่จะนำข้อมูลนั้นมาใช้งานให้เกิดประโยชน์สูงสุด (ฝ่ายพัฒนาทรัพยากรบุคคล บจก. กสท. โทรคมนาคม, 2550)

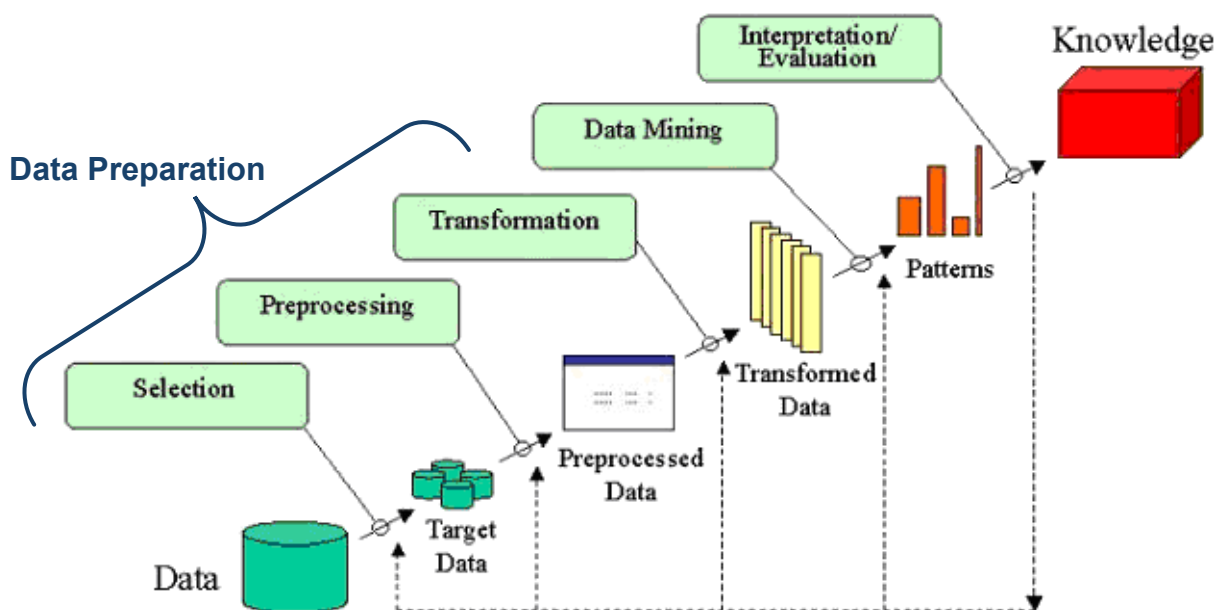
### 2.3.3 กระบวนการทำงานของ Knowledge Discovery in Database (KDD) หรือ Data Mining

ในมุมมองเชิงวิชาการนั้น นักวิจัยจะอ้างถึงกระบวนการทั้งหมดในการทำ Data Mining ว่า “Knowledge discovery in database (KDD)” และใช้คำว่า “ดาต้าไมน์นิง (Data Mining)” แทนขั้นตอนขั้นหนึ่งของกระบวนการที่เกี่ยวข้องกับการค้นหารูปแบบ ความสัมพันธ์ของข้อมูลเท่านั้น อย่างไรก็ตาม ในแง่มุมมองธุรกิจจะใช้คำว่า “ดาต้าไมน์นิง (Data Mining)” แทนความหมายของขั้นตอนทั้งหมด

กระบวนการ KDD หมายถึง กระบวนการในการค้นหาความรู้หรือสารสนเทศของข้อมูลที่อยู่ในกลุ่มข้อมูลจำนวนมาก ซึ่งมีขั้นตอนของการทำดาต้าไมน์นิงเป็นกระบวนการที่สำคัญในการค้นหาลักษณะที่น่าสนใจของข้อมูลเหล่านี้ เช่น รูปแบบ ความสัมพันธ์ การเปลี่ยนแปลง โครงสร้างที่เด่นชัด หรือลักษณะที่ผิดปกติของข้อมูลจากข้อมูลจำนวนมากที่เก็บอยู่ในฐานข้อมูล ซึ่งวิธีการต่างๆ ที่นำมาใช้ในการทำดาต้าไมน์นิงนี้มีวัตถุประสงค์ต่างกัน ขึ้นอยู่กับผลลัพธ์ของกระบวนการโดยรวมที่ต้องการ ดังนั้น จึงควรมีการนำเสนอวิธีการที่หลากหลายสำหรับเป้าหมายที่แตกต่างกัน เพื่อให้ได้ผลลัพธ์ที่เหมาะสมตามที่ต้องการหลังจากนำไปใช้งานแล้ว

(มหาวิทยาลัยศิลปากร, วิไลวรรณ เชมยสุทธิ, 2551)

จากที่ได้กล่าวมาแล้วว่าดาต้าไมน์นิงเป็นขั้นตอนหนึ่งที่สำคัญในกระบวนการค้นหาหรือสกัดความรู้และสารสนเทศที่มีประโยชน์ออกจากฐานข้อมูลขนาดใหญ่ (Knowledge Discovery in Database : KDD) ซึ่งโดยทั่วไปกระบวนการของ KDD นั้นประกอบด้วยขั้นตอนต่างๆ ดังนี้



ภาพที่ 2-2 กระบวนการทำงานของ Data Mining

(ฝ่ายพัฒนาทรัพยากรบุคคล บจก. กสท. ไตรคมนาคม, 2550)

จากภาพที่ 2-2 แสดงกระบวนการทำงานของดาต้าไมน์นิ่งประกอบไปด้วยขั้นตอนการทำงานหลักอยู่ 5 ขั้นตอน สามารถอธิบายขั้นตอนของกระบวนการต่างๆ ได้ ดังนี้

1) Business Object ขั้นตอนแรกจะต้องมีการกำหนด ขอบเขต เป้าหมายในการทำ KDD หรือดาต้าไมน์นิ่งก่อนว่าจะทำดาต้าไมน์นิ่ง เพื่อใช้ในการตัดสินใจเรื่องใด ซึ่งขั้นตอนนี้จะมีผลต่อทุกๆ ขั้นตอนในการทำ KDD โดยนักวิเคราะห์ธุรกิจ (Business Analyst) จะต้องมีการกำหนดปัญหาที่เกิดขึ้นหรือความต้องการทางธุรกิจให้ครอบคลุม ชัดเจน และกำหนดวัตถุประสงค์ทางธุรกิจด้วย

2) Data Preparation ขั้นตอนนี้ มีหน้าที่คือ จัดการข้อมูลให้สามารถนำเข้าสู่กระบวนการของดาต้าไมน์นิ่งได้ โดยในขั้นตอนนี้มีกระบวนการย่อยออกเป็น 3 ส่วน ดังนี้

2.1) Data Selection เป็นการคัดเลือกข้อมูลที่เกี่ยวข้องกับเป้าหมาย หรือวัตถุประสงค์ที่ได้ตั้งไว้ เพื่อให้ผลการวิเคราะห์ข้อมูลหรือผลของการทำกระบวนการดาต้าไมน์นิ่ง สามารถนำไปใช้ในการประกอบการตัดสินใจได้ตามความต้องการ โดยสามารถเลือกจากแหล่งข้อมูลต่างๆ ทั้งที่อยู่ภายใน และภายนอกองค์กร ซึ่งแหล่งข้อมูลหลักของการทำดาต้าไมน์นิ่ง คือ คลังข้อมูล (Data Warehouse) หรือฐานข้อมูล (Database)

2.2) Data Preprocessing เป็นขั้นตอนการปรับข้อมูลที่ได้คัดเลือกมานั้น ให้มีคุณภาพมากยิ่งขึ้น เนื่องจากการคัดเลือกข้อมูลจากหลากหลายแหล่งข้อมูล อาจจะทำให้มีรูปแบบ (Format) ในการเก็บข้อมูลไม่เหมือนกัน ส่งผลให้เกิดการผิดพลาดในการทำกระบวนการดาต้าไมน์นิ่งได้ ซึ่งการปรับข้อมูลในขั้นตอน Data Preprocessing นี้ ประกอบไปด้วยกระบวนการต่างๆ ดังนี้

- Data Cleaning เป็นกระบวนการทำความสะอาดข้อมูล โดยการเติมข้อมูลที่ขาดหายไปในบาง Field รวมทั้งเลือกข้อมูลที่จำเป็นต่อการทำกระบวนการดาต้าไมน์นิ่ง เช่น ถ้าต้องการที่จะวิเคราะห์ความต้องการของลูกค้า ควรจะเลือกข้อมูลที่สำคัญต่อการวิเคราะห์ ซึ่งอาจจะตัดข้อมูลในส่วน of รายละเอียดที่ไม่จำเป็น เช่น เลขที่บัตรประชาชน หรือเบอร์โทรศัพท์ เป็นต้น นอกจากนี้ยังจัดการกับความผิดปกติของข้อมูล และจัดการกับข้อมูลที่ไม่สอดคล้องกัน เช่น อายุ ไม่ตรงกับปีที่เกิด, วันที่มีเดือนที่ 13, เพศหญิง แต่มีคำนำหน้าว่า นาย เป็นต้น

- Data Integration เป็นกระบวนการรวมข้อมูล โดยต้องมีการตรวจสอบก่อนว่า การรวมข้อมูลจากแหล่งข้อมูลต่างๆ นั้น มีรูปแบบในการจัดเก็บที่เหมือนกันหรือต่างกัน เช่น ข้อมูลที่มีรูปแบบการจัดเก็บต่างกัน หรืออาจจะมีหน่วยวัดที่ต่างกัน สัญลักษณ์ที่ใช้ไม่เหมือนกัน รวมไปถึงลดความซ้ำซ้อนของข้อมูลด้วย

- Data Reduction เป็นกระบวนการในการลดขนาดของข้อมูล ซึ่งสามารถทำได้ 2 แนวทางคือ Data size reduction ใช้แถวเป็นหลักในการลดข้อมูล และ Dimensionality reduction ใช้ Attribute เป็นหลักในการลดข้อมูล เนื่องจากในบางเทคนิคของกระบวนการของดาต้า



ไมน์นิง (Data Mining) จะต้องอาศัยข้อมูลในเรื่องต่างๆ เป็นจำนวนที่เท่าๆ กัน เพื่อความเที่ยงตรง และความถูกต้องของผลการทำดาต้าไมน์นิง

2.3) Data Transformation เป็นขั้นตอนในการแปลงข้อมูล เพื่อให้ใช้กับโมเดลต่างๆ ของกระบวนการทำดาต้าไมน์นิงได้ ซึ่งมีวิธีการแปลงข้อมูลในรูปแบบต่างๆ ดังนี้

- Normalization คือการแปลงข้อมูลให้อยู่ในช่วงค่าที่กำหนด จัดได้ว่าเป็นการลดระยะห่างของข้อมูล เนื่องจากข้อมูลชนิดเดียวกัน แต่มีระยะห่างมากเกินไป ซึ่งทำให้มีผลต่อการวิเคราะห์ข้อมูลในบางโมเดลของดาต้าไมน์นิง เช่น ข้อมูลที่มีค่าช่วง 0-1,000,000 แปลงให้เป็นช่วง 0-20 เป็นต้น

- Discretization คือกระบวนการแปลงข้อมูลที่เป็นจำนวนจริงต่อเนื่อง ให้เป็นช่วงข้อมูล จัดได้ว่าเป็นกำหนดกลุ่มให้ข้อมูล ลดการกระจายของข้อมูล เช่น กำหนดกลุ่มให้คะแนนของนักศึกษา อายุของลูกค้า แปลงให้เป็นช่วงข้อมูล เช่น ช่วงอายุ 20-25 ปี, ช่วง 26-30 ปี, ช่วง 31-35 ปี ตามลำดับ

- 1 Of N Coding คือกระบวนการแปลงข้อมูลจากตัวอักษรให้เป็นตัวเลข ที่ไม่ได้บอกถึงลำดับความมากน้อย เช่นการกำหนดรหัสให้กับข้อมูล เป็นต้น

- Generalization คือกระบวนการแปลงข้อมูลที่มีรายละเอียดมากเกินไป ให้มีความกระชับมากยิ่งขึ้น เช่น ที่อยู่ของลูกค้า อาจจะทำการแปลงข้อมูลให้กระชับขึ้น เป็นอยู่ในกทม. หรืออยู่ต่างจังหวัด เป็นต้น

3) Data Mining เป็นขั้นตอนการทำเหมืองข้อมูล โดยมีเทคนิค หรือ Operation ในการทำดาต้าไมน์นิง หลายรูปแบบ เช่น Database Segmentation (Clustering), Data Classification และ Link Analysis (Association) เป็นต้น ซึ่งแต่ละเทคนิคนั้น จะมีอัลกอริทึมให้เลือกใช้ ซึ่งจะกล่าวถึงในหัวข้อถัดไป เรื่อง เทคนิคการทำดาต้าไมน์นิง

4) Interpretation / Evaluation หรือ Analysis of Result and Knowledge Presentation เป็นขั้นตอนสุดท้ายสำหรับนักวิเคราะห์ข้อมูล ที่จะต้องเก็บผลลัพธ์ของกระบวนการดาต้าไมน์นิงมาทำการสรุปความหมาย หรือตีความของผลลัพธ์ที่ได้ ออกมาเป็นข้อมูลความรู้ใหม่ (Knowledge) ที่สามารถนำไปเป็นสารสนเทศที่ช่วยในการตัดสินใจของผู้บริหารต่อไปได้

อาจจะกล่าวได้ว่ากระบวนการดาต้าไมน์นิง (Data Mining) เป็นเพียงเครื่องมือที่ช่วยสกัดเอาความรู้ที่ซ่อนเร้นอยู่ในข้อมูลเป็นจำนวนมาก แต่ข้อมูลความรู้ที่ได้มานั้น จะนำไปใช้ประโยชน์ได้มากน้อยเพียงใด ขึ้นอยู่กับการตีความหมาย หรือการสรุปความหมายของนักวิเคราะห์

ในกระบวนการของ KDD หรือดาต้าไมน์นิง (Data Mining) นั้น จะใช้เวลาในการเตรียมข้อมูล สำหรับการทำเหมืองข้อมูลมากถึง 60 % เนื่องจากว่าผลการวิเคราะห์เพื่อให้ได้ความรู้ใหม่

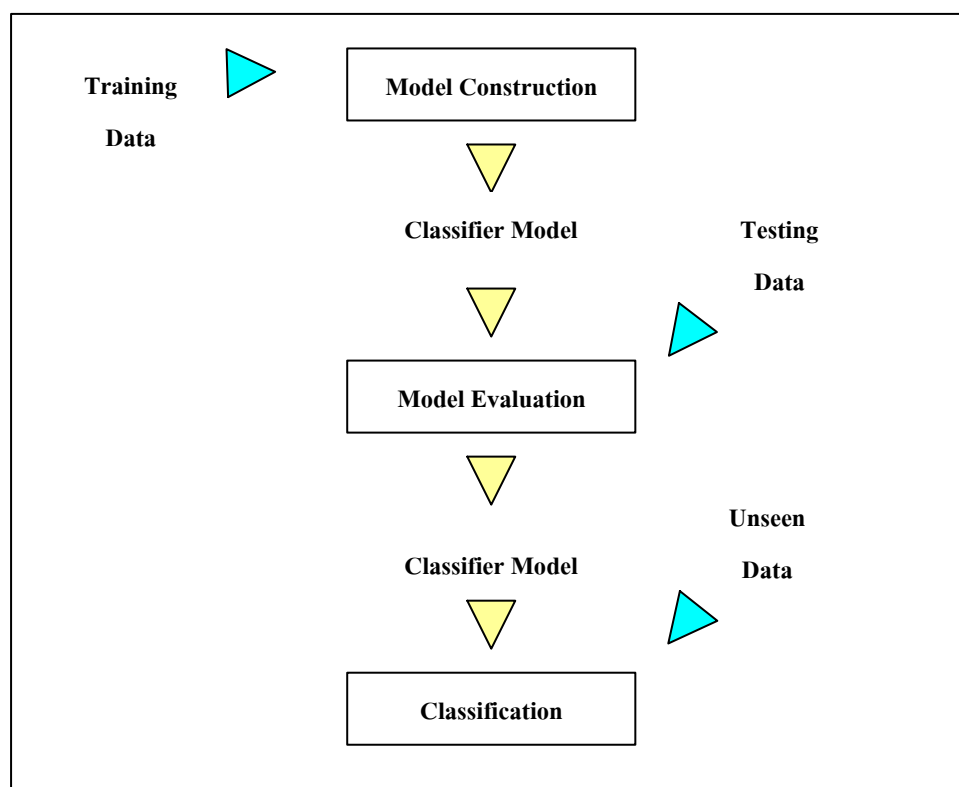
หรือสารสนเทศที่ซ่อนเร้นอยู่ในข้อมูลเป็นจำนวนมากนั้น จะมีความถูกต้องมากน้อยแค่ไหน ก็ขึ้นอยู่กับว่าข้อมูลที่นำไปวิเคราะห์นั้นมีความสอดคล้อง และมีความผิดพลาดเท่าใด ดังนั้นจึงต้องให้ความสำคัญกับการเตรียมข้อมูล ส่วนการทำดาต้าไมน์นิ่งจริงๆ นั้น ใช้เวลาเพียง 10% เท่านั้น

### 2.3.4 เทคนิคการทำ Data Mining

เทคนิคการทำดาต้าไมน์นิ่ง (Data Mining) ซึ่งในปัจจุบันมีหลายรูปแบบ โดยผู้วิจัยได้ทำการศึกษา ค้นคว้า เกี่ยวกับเทคนิคการทำดาต้าไมน์นิ่งจากแหล่งข้อมูลต่างๆ ซึ่งสามารถบอกถึงเทคนิคต่างๆ ของ Data Mining ที่ได้รับความนิยม ดังนี้

#### 1) Classification หรือ Predictive

เป็นเทคนิคที่ใช้ในการจำแนกกลุ่มของข้อมูลด้วยคุณลักษณะต่างๆ ที่ได้มีการกำหนดไว้แล้ว ซึ่งกระบวนการ Classification นั้นจะมีการแบ่งข้อมูลออกเป็น 2 ส่วน คือ กลุ่มข้อมูลที่ใช้ในการเรียนรู้ (Training data) และกลุ่มข้อมูลที่ใช้ในการทดสอบ (Testing data) โดยกระบวนการจะใช้ Model Construction (Learning) ซึ่งเป็นการสร้างโมเดลที่มีการเรียนรู้จากข้อมูลเดิมที่มีการกำหนดคลาสไว้เรียบร้อยแล้ว (Training data) ซึ่งข้อมูลนั้นสามารถกระทำกับข้อมูลเชิงตัวเลขและข้อมูลที่มีลักษณะเป็นกลุ่มได้ เพื่อให้โมเดลในการจำแนกข้อมูลเกิดการเรียนรู้พฤติกรรมของข้อมูลเดิม ซึ่งยังมีข้อมูล Training data มากแค่ไหน ก็ส่งผลให้โมเดลสามารถจำแนกกลุ่มได้อย่างถูกต้องมากขึ้น



ภาพที่ 2-3 กระบวนการทำงานของ Classification (มหาวิทยาลัยขอนแก่น, รัตติกาล ทาโบราณ: 7)

เทคนิคที่ใช้ทำค้ำไม้ในแบบ Classification หรือ Predictive ได้แก่ Decision tree และ Neural Network ดังนี้

1.1) Decision tree เป็นเทคนิคที่ให้ผลลัพธ์ในลักษณะของโครงสร้างต้นไม้ ซึ่งเมื่อมีข้อมูลที่ต้องการจะจัดกลุ่ม ก็จะนำ attribute ต่างๆ ของข้อมูลนั้นไปเทียบกับ Decision tree ตามเส้นทางใน tree จนกระทั่งคลาสปลาซทางซึ่งก็คือ กลุ่มของข้อมูลที่เหมือนกัน โดยปกติมักประกอบด้วยกฎในรูปแบบ “ถ้า เงื่อนไข แล้ว ผลลัพธ์” เช่น

“If Income = High and Married = No THEN Risk = Poor”

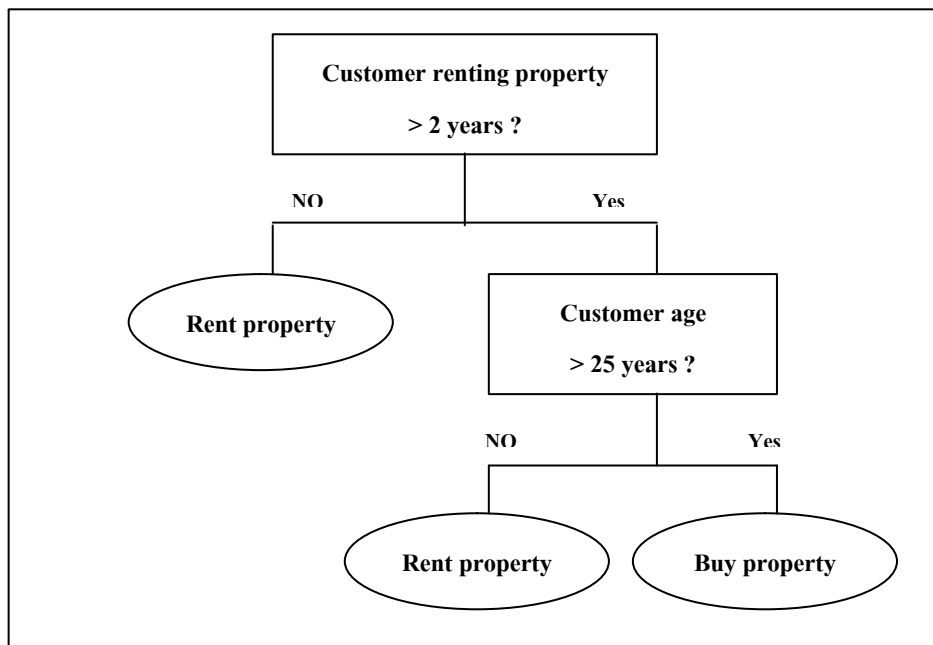
“If Income = High and Married = Yes THEN Risk = Good”

เนื่องจากผู้ใช้สามารถทำความเข้าใจผลลัพธ์ได้ง่าย เทคนิค Decision tree จะจำกัดข้อมูลที่เป็นตัวแปรตาม (dependent variable) 1 ตัวต่อ 1 แบบจำลอง ถ้าต้องการทำนายตัวแปรตามหลายๆ ตัว ก็จะต้องสร้างแบบจำลองสำหรับตัวแปรตามแต่ละตัว ซึ่ง algorithm ของเทคนิคแบบ Decision tree นี้ ส่วนใหญ่ไม่รองรับข้อมูลแบบต่อเนื่อง (continuous data) จะต้องมีการแบ่งให้เป็นข้อมูลแบบไม่ต่อเนื่องหรือแบบช่วงข้อมูล (discrete data) เสียก่อน ซึ่ง algorithm ที่ใช้ในการดำเนินการของ Decision tree นั้น ได้แก่ Chi-squared Automatic Interaction Detection (CHAID) , Classification and Regression Trees (CART) , C4.5 และ C5.0 โดย algorithm เหล่านี้ส่วนมากมักเหมาะกับปัญหาแบบ classification (มหาวิทยาลัยศิลปากร, วิไลวรรณ เสรษญศิริ, 2551)

ตัวอย่าง สมมติว่าบริษัทขนาดใหญ่แห่งหนึ่ง ทำธุรกิจอสังหาริมทรัพย์มีสำนักงานสาขาอยู่ประมาณ 50 แห่ง แต่ละสาขามีพนักงานประจำ เป็นผู้จัดการและพนักงานขาย พนักงานเหล่านี้แต่ละคนจะดูแลอาคารต่าง ๆ หลายแห่งรวมทั้งลูกค้าจำนวนมาก บริษัทจำเป็นต้องใช้ระบบฐานข้อมูลที่กำหนดความสัมพันธ์ระหว่างองค์ประกอบเหล่านี้ เมื่อรวบรวมข้อมูลแบ่งเป็นตารางพื้นฐานต่างๆ เช่น ข้อมูลสำนักงานสาขา (Branch) ข้อมูลพนักงาน (Staff) ข้อมูลทรัพย์สิน (Property) และข้อมูลลูกค้า (Client) พร้อมทั้งกำหนดความสัมพันธ์ (Relationship) ของข้อมูลเหล่านี้ เช่น ประวัติการเช่าบ้านของลูกค้า (Customer\_rental) รายการให้เช่า (Rentals) รายการขายสินทรัพย์ (Sales) เป็นต้น ส่วนหนึ่งของรายงานจากฐานข้อมูลสรุปว่า

“40 % ของลูกค้าที่เช่าบ้านนานกว่าสองปี และมีอายุเกิน 25 ปี จะซื้อบ้านเป็นของตนเอง โดยกรณีเช่นนี้เกิดขึ้น 35 % ของลูกค้าผู้เช่าบ้านของบริษัท”

ดังภาพที่ 2-4 แสดงให้เห็นถึง Decision Tree สำหรับการวิเคราะห์หว่าลูกค้าบ้านเช่า จะมีความสนใจที่จะซื้อบ้านเป็นของตนเองหรือไม่ โดยปัจจัยที่ใช้ในการวิเคราะห์ คือระยะเวลาที่ลูกค้าได้เช่าบ้านมา และอายุของลูกค้า (มหาวิทยาลัยขอนแก่น, รัตติกาล ทาโบราณ: 7)



ภาพที่ 2-4 ตัวอย่างของ Decision Tree เพื่อวิเคราะห์โอกาสที่ลูกค้าบ้านเช่าจะซื้อบ้าน  
(มหาวิทยาลัยขอนแก่น, รัตติกาล ทาโบราณ: 7)

### ทฤษฎี Information Gain

จากการศึกษาการสร้าง Decision tree ในการทำ Classification มีทฤษฎีที่ใช้ในการสร้าง Decision tree คือ ทฤษฎี Information gain เป็นทฤษฎีที่นำมาใช้ในการเลือก attribute ในแต่ละ node ของ tree ซึ่ง attribute ตัวใดที่มีค่า Information gain สูงสุด หรือว่ามีค่า Entropy น้อย ก็จะถูกเลือกเป็น attribute ของ node นั้น (ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเชียงใหม่, 2548)

ทฤษฎีนี้มีสมการการคำนวณ Information gain และค่า Entropy ดังนี้

$$GAIN_{split} = Entropy(p) - \left( \sum_{i=1}^k \frac{n_i}{n} Entropy(i) \right)$$

$$Entropy(t) = - \sum_j p(j|t) \log_2 p(j|t)$$

จากสมการดังกล่าว สามารถยกตัวอย่างการคำนวณหาค่า Information gain และค่า Entropy ได้ โดยมีตัวอย่างการจัดกลุ่มข้อมูลการเล่นสกีที่ใช้ในการคำนวณหา Information gain และค่า Entropy ดังนี้ (สถาบันบัณฑิตพัฒนบริหารศาสตร์, 2552)

Gender	Marital Status	Play Ski
1	1	1
0	2	0
1	3	1
0	2	0
1	1	1
0	2	0
1	3	1
0	1	0
1	1	1
0	2	1

Decision Tree

Sol <sup>n</sup> #1			
Female		Male	
Play	Not Play	Play	Not Play
5	0	1	4

Sol <sup>n</sup> #2					
Single		Married		Divorced	
Play	Not Play	Play	Not Play	Play	Not Play
3	1	1	3	2	0

Gender 0 = ผู้หญิง, 1 = ผู้ชาย    Marital Status 1 = Single, 2 = Married, 3 = Divorced    Play Ski 0 = Not Play, 1 = Play

จากข้อมูลตัวอย่างดังกล่าวข้างต้น สามารถแสดงการคำนวณหาค่า Entropy ของ attribute Gender และ attribute Marital ได้ดังนี้

Sol <sup>n</sup> #1			
Female		Male	
Play	Not Play	Play	Not Play
5	0	1	4

- Entropy (Gender)
 
$$= 0.5 * (-1 \log_2 1 - 0 \log_2 0)$$

$$+ 0.5 * (-0.2 \log_2 0.2 - 0.8 \log_2 0.8)$$

$$\approx 0.36$$

จากสูตร  $Entropy(t) = -\sum_j p(j|t) \log_2 p(j|t)$  สามารถอธิบายได้ ดังนี้

- ค่า P คือ จำนวนเพศหญิง และชาย หาคด้วยจำนวนข้อมูลทั้งหมด ซึ่ง ณ ที่นี้มีเพศหญิง 5 คน และเพศชาย 5 คน จากทั้งหมด 10 คน ดังนั้นจะได้ค่า P ของเพศชายมีค่าเท่ากับ 0.5 และค่า P ของเพศหญิง มีค่าเท่ากับ 0.5

- ค่า J ของเพศหญิง อธิบายได้ ดังนี้

จำนวนของเพศหญิงที่เล่นสกี หาคด้วยจำนวนเพศหญิงทั้งหมด ดังนั้นค่า J ของเพศหญิงที่เล่นสกี จะมีค่าเท่ากับ  $5/5 = 1$

จำนวนของเพศหญิงที่ไม่เล่นสกี หาคด้วยจำนวนเพศหญิงทั้งหมด ดังนั้นค่า J ของเพศหญิงที่ไม่เล่นสกี จะมีค่าเท่ากับ  $0/5 = 0$

- ค่า  $J$  ของเพศชาย อธิบายได้ ดังนี้

จำนวนของเพศชายที่เล่นสกี หารด้วยจำนวนเพศชายทั้งหมด ดังนั้นค่า  $J$  ของเพศชายที่เล่นสกี จะมีค่าเท่ากับ  $1/5 = 0.2$

จำนวนของเพศชายที่ไม่เล่นสกี หารด้วยจำนวนเพศชายทั้งหมด ดังนั้นค่า  $J$  ของเพศชายที่ไม่เล่น จะมีค่าเท่ากับ  $4/5 = 0.8$

Sol <sup>n</sup> #2					
Single		Married		Divorced	
Play	Not Play	Play	Not Play	Play	Not Play
3	1	1	3	2	0

• Entropy (Marital Status)

$$\begin{aligned}
 &= 0.4 * (-0.75 \log_2 0.75 - 0.25 \log_2 0.25) \\
 &\quad + 0.4 * (-0.25 \log_2 0.25 - 0.75 \log_2 0.75) \\
 &\quad + 0.2 * (-1 \log_2 1 - 0 \log_2 0) \\
 &\approx 0.65
 \end{aligned}$$

จากสูตร  $Entropy(t) = -\sum_j p(j|t) \log_2 p(j|t)$  สามารถอธิบายได้ ดังนี้

- ค่า  $P$  คือ จำนวนผู้ที่มีสถานภาพโสด แต่งงาน และหย่าร้าง หารด้วยจำนวนข้อมูลทั้งหมด ซึ่ง ณ ที่นี้ผู้ที่มีสถานภาพโสดมี 4 คน แต่งงานมี 4 คน และหย่าร้างมี 2 คน จากทั้งหมด 10 คน ดังนั้นจะได้ค่า  $P$  ของผู้ที่มีสถานภาพโสดมีค่าเท่ากับ 0.4 และค่า  $P$  ของผู้ที่แต่งงาน มีค่าเท่ากับ 0.4 และค่า  $P$  ของผู้ที่ย่าร้างมีค่าเท่ากับ 0.2

- ค่า  $J$  ของผู้ที่มีสถานภาพโสด อธิบายได้ ดังนี้

จำนวนของผู้ที่มีสถานภาพโสดที่เล่นสกี หารด้วยจำนวนผู้ที่มีสถานภาพโสดทั้งหมด ดังนั้นค่า  $J$  ของผู้ที่มีสถานภาพโสดที่เล่นสกี จะมีค่าเท่ากับ  $3/4 = 0.75$

จำนวนของผู้ที่มีสถานภาพโสดที่ไม่เล่นสกี หารด้วยจำนวนผู้ที่มีสถานภาพโสดทั้งหมด ดังนั้นค่า  $J$  ของผู้ที่มีสถานภาพโสดที่ไม่เล่นสกี จะมีค่าเท่ากับ  $1/4 = 0.25$

- ค่า  $J$  ของผู้ที่แต่งงานแล้ว อธิบายได้ ดังนี้

จำนวนของผู้ที่แต่งงานแล้วที่เล่นสกี หารด้วยจำนวนผู้ที่แต่งงานแล้วทั้งหมด ดังนั้นค่า  $J$  ของผู้ที่แต่งงานแล้วที่เล่นสกี จะมีค่าเท่ากับ  $1/4 = 0.25$

จำนวนของผู้ที่แต่งงานแล้วที่ไม่เล่นสกี หารด้วยจำนวนผู้ที่แต่งงานแล้วทั้งหมด ดังนั้นค่า  $J$  ของผู้ที่แต่งงานแล้วที่ไม่เล่นสกี จะมีค่าเท่ากับ  $3/4 = 0.75$

- ค่า  $J$  ของผู้ที่หย่าร้าง อธิบายได้ ดังนี้

จำนวนของผู้ที่หย่าร้างที่เล่นสกี หาดด้วยจำนวนผู้ที่หย่าร้างทั้งหมด ดังนั้นค่า  $J$  ของผู้ที่หย่าร้างที่เล่นสกี จะมีค่าเท่ากับ  $2/2 = 1$

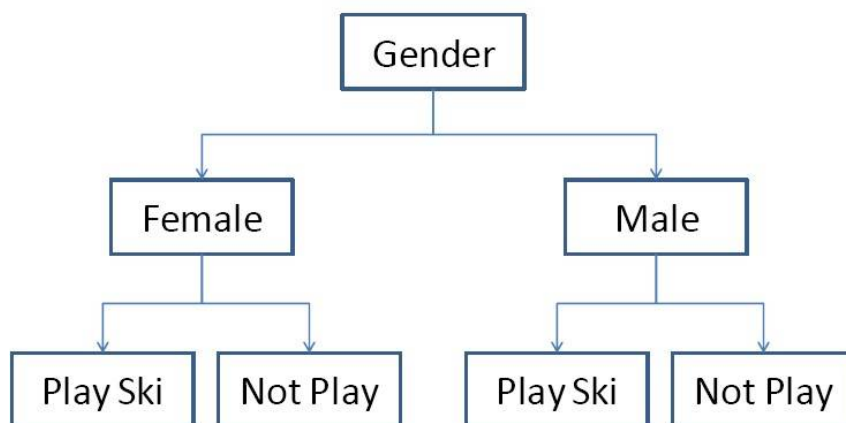
จำนวนของผู้ที่หย่าร้างที่ไม่เล่นสกี หาดด้วยจำนวนผู้ที่หย่าร้างทั้งหมด ดังนั้นค่า  $J$  ของผู้ที่หย่าร้างที่ไม่เล่นสกี จะมีค่าเท่ากับ  $0/2 = 0$

จากการคำนวณดังกล่าว สรุปได้ว่า

Entropy (Gender)  $\approx 0.36$  หมายความว่าถ้าต้องการจะแยกแยะข้อมูลว่าคนแบบไหนจะเล่นสกี ถ้าใช้ Gender ในการแยกข้อมูลแล้วก็ยังต้องการข้อมูลเพิ่มเติมอีก 0.36 บิต ในการตัดสินใจ

Entropy (Marital Status)  $\approx 0.65$  หมายความว่า ถ้าต้องการจะแยกแยะข้อมูลว่าคนแบบไหนจะเล่นสกี ถ้าใช้ Marital Status ในการแยกข้อมูลแล้วก็ยังต้องการข้อมูลเพิ่มเติมอีก 0.65 บิต ในการตัดสินใจ

ดังนั้น จึงสรุปได้ว่า ควรจะใช้ Gender แบ่งแยกข้อมูลก่อน เพราะ Gender แบ่งแยกข้อมูลได้ดีกว่า ดังนี้



ภาพที่ 2-5 Decision Tree เพื่อวิเคราะห์โอกาสที่ลูกค้าจะเล่นสกี

(สถาบันบัณฑิตพัฒนบริหารศาสตร์, 2552)

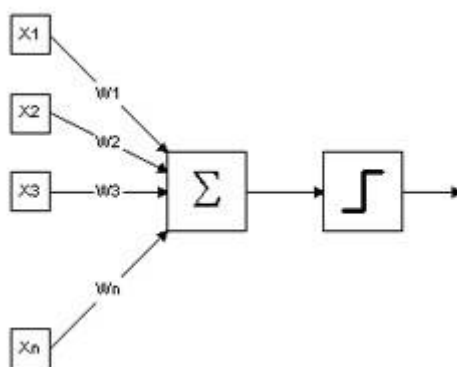
1.2) Neural Network หรือเรียกว่า “โครงข่ายประสาทเทียม” เป็นเทคโนโลยีหนึ่งที่มีที่มาจากงานวิจัยด้านปัญญาประดิษฐ์ Artificial Intelligence:AI เพื่อใช้ในการคำนวณค่าฟังก์ชันจากกลุ่มข้อมูล โดยวิธีการของ นิวรอลเน็ตเวิร์ก (Artificial Neural Networks หรือ ANN ) นั้น เป็นวิธีการที่ให้เครื่องหรือระบบคอมพิวเตอร์มีการเรียนรู้จากตัวอย่างต้นแบบ แล้วทำการฝึก (train) ให้ระบบได้รู้จักที่จะคิดแก้ปัญหาที่กว้างขึ้นได้ โดยอาศัยข้อมูลในการเรียนรู้ (Training data) ซึ่งในโครงสร้างของนิวรอลเน็ตเวิร์กนี้ จะประกอบด้วยโหนด (node) สำหรับ Input – Output และการ

ประมวลผลกระจายอยู่ในโครงสร้างเป็นชั้น ๆ ได้แก่ input layer , output layer และ hidden layers การประมวลผลของนิวรอลเน็ตเวิร์กจะอาศัยการส่งการทำงานผ่านโหนดต่าง ๆ ใน layer เหล่านี้ (มหาวิทยาลัยขอนแก่น, รัตติกาล ทาโบราณ: 7)

ซึ่งสามารถอธิบายถึงโครงสร้างและส่วนประกอบของโครงข่ายประสาทเทียม (Neural Network) ได้ดังนี้

- 1) หน่วยประมวลผล (Processing Element)
- 2) ค่าน้ำหนักและฟังก์ชันการแปลงค่าจากหน่วยความจำไปยังหน่วยประมวลผล (Weight and Transfer Function)
- 3) ชั้นข้อมูลนำเข้า (Input Layer)
- 4) ชั้นซ่อนเร้นของหน่วยประมวลผล (Hidden Layer)
- 5) ความหนาแน่นของจุดที่นิวรอลติดต่อซึ่งกันและกัน (Connections)
- 6) การเรียนรู้ (Training) ซึ่งเกิดจากการปรับค่าน้ำหนักระหว่างหน่วยประมวลผล
- 7) ชั้นแสดงผลลัพธ์ (Output Layer)

สำหรับหลักการทำงานในคอมพิวเตอร์ Neurons ประกอบด้วย input และ output เหมือนกัน โดยจำลองให้ input แต่ละอันมี weight เป็นตัวกำหนดน้ำหนักของ input โดย neuron แต่ละหน่วยจะมีค่า threshold เป็นตัวกำหนดว่าน้ำหนักรวมของ input ต้องมากขนาดไหนจึงสามารถส่ง output ไปยัง neurons ตัวอื่นได้ เมื่อนำ neuron แต่ละหน่วยมาต่อกันให้ทำงานร่วมกัน



ภาพที่ 2-6 โมเดลของ Neurons ในระบบคอมพิวเตอร์  
(ภาควิชาวิทยาการคอมพิวเตอร์มหาวิทยาลัยขอนแก่น, 2548)

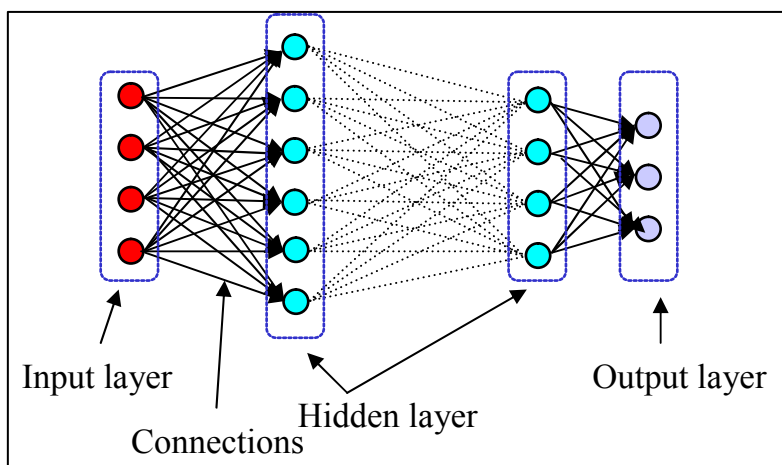
การทำงานของนิวรอลเน็ตเวิร์ก (Neural networks) ก็คือ เมื่อมี input เข้ามายัง network ก็เอา input มาคูณกับ weight ของแต่ละขา ผลที่ได้จาก input ทุก ๆ ขาของ neuron จะเอามารวมกันแล้วก็เอามาเทียบกับ threshold ที่กำหนดไว้ ถ้าผลรวมมีค่ามากกว่า threshold แล้ว neuron ก็



จะส่ง output ออกไป output นี้ก็จะถูกส่งไปยัง input ของ neuron อื่น ๆ ที่เชื่อมกันใน network ถ้าค่าน้อยกว่า threshold ก็จะไม่เกิด output โดยสามารถเขียนอธิบายได้ดังนี้

**if (sum(input \* weight) > threshold) then output**

สิ่งสำคัญคือเราต้องทราบค่า weight และ threshold สำหรับสิ่งที่เราต้องการเพื่อให้คอมพิวเตอร์รู้จัก ซึ่งเป็นค่าที่ไม่แน่นอน แต่สามารถกำหนดให้คอมพิวเตอร์ปรับค่าเหล่านั้นได้โดยการสอนให้มันรู้จัก pattern ของสิ่งที่เราต้องการให้มันรู้จัก เรียกว่า "back-propagation" ซึ่งเป็นกระบวนการย้อนกลับของการรู้จัก ในการฝึก feed-forward neural networks จะมีการใช้อัลกอริทึมแบบ back-propagation เพื่อใช้ในการปรับปรุงน้ำหนักคะแนนของเครือข่าย (network weight) หลังจากใส่รูปแบบข้อมูลสำหรับฝึกให้แก่เครือข่ายในแต่ละครั้งแล้ว ค่าที่ได้รับ (output) มาจากเครือข่ายจะถูกนำไปเปรียบเทียบกับผลที่คาดหวัง แล้วทำการคำนวณหาค่าความผิดพลาด ซึ่งค่าความผิดพลาด จะถูกส่งกลับเข้าสู่เครือข่ายเพื่อใช้แก้ไขค่าน้ำหนักคะแนนต่อไป (ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น, 2548)



ภาพที่ 2-7 ลักษณะ โครงสร้างของ Neural Network

ซึ่ง Output ของแต่ละ Node สามารถเขียนเป็นสมการได้ดังนี้

$$y_i = f(w_i^1 x_1 + w_i^2 x_2 + w_i^3 x_3 + \dots + w_i^m x_m)$$

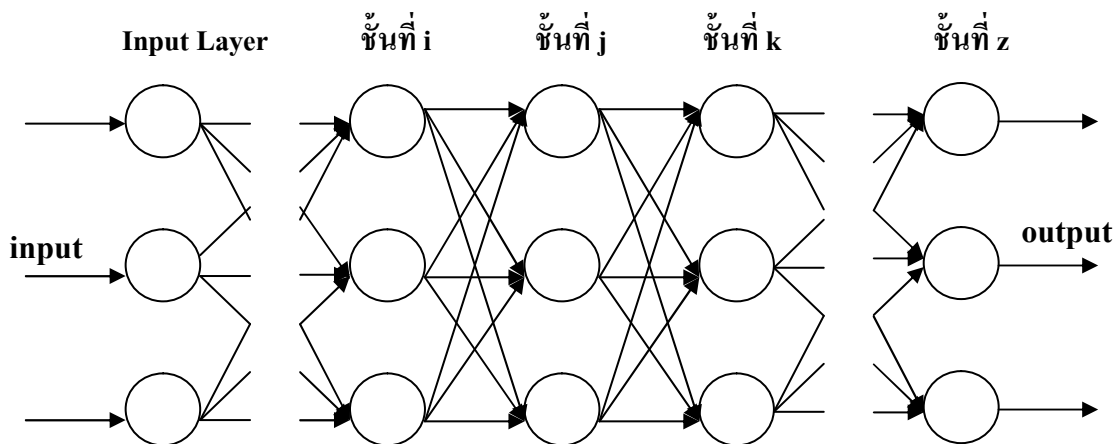
$$= f\left(\sum_j w_i^j x_j\right)$$

เมื่อ  $x_j$  = input จากโหนดอื่นๆ

$w_i^j$  = น้ำหนัก (weight) ของแต่ละแขน (connection)

**Back propagation Algorithm**

Back-propagation เป็นอัลกอริทึมที่ใช้สำหรับเรียนรู้ของเครือข่ายประสาทเทียม (Neural Network) วิธีหนึ่งที่ยิยมใช้ใน multilayer perceptron เพื่อปรับค่าน้ำหนักในเส้นเชื่อมต่อระหว่างโหนดให้เหมาะสม โดยการปรับค่านี้อาจจะขึ้นกับความแตกต่างของค่าเอาต์พุตที่คำนวณได้กับค่าเอาต์พุตที่ต้องการ ดังภาพที่ 2-8 (ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น, 2548)



ภาพที่ 2-8 รูปแบบ Back-propagation neural network

จากภาพที่ 2-8 สามารถเสนอขั้นตอนของการทำ Back-propagation Algorithm ได้ดังนี้ (ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่น, 2548)

1. กำหนดค่าอัตราเร็วในการเรียนรู้ ( $r$  : rate parameter)
2. สำหรับแต่ละตัวอย่างอินพุตให้ทำตามขั้นตอนนี้จนกว่าได้ระดับ performance ที่ต้องการ
  - คำนวณหาค่าเอาต์พุตโดยใช้ค่าน้ำหนักเริ่มต้นซึ่งอาจได้จากการสุ่ม
  - คำนวณหาค่า  $\beta$  : แทนประโยชน์ที่ได้รับในการเปลี่ยนค่าเอาต์พุตของแต่ละโหนด
  - ในชั้นเอาต์พุต (Output Layer) (ชั้นที่ Z)

$$\beta_z = d_z - o_z$$

เมื่อ  $d_z$  = ค่าเอาต์พุตที่ต้องการ  
 $o_z$  = ค่าเอาต์พุตที่คำนวณได้

- ในชั้นซ่อน (Hidden Layer) (ชั้นที่ j และ ชั้นที่ k)

$$\beta_j = \sum_k w_{j \rightarrow k} o_k (1 - o_k) \beta_k$$

เมื่อ  $w_{j \rightarrow k}$  = น้ำหนักของเส้นเชื่อมระหว่างชั้นที่ j กับ k

- คำนวณค่าน้ำหนักที่เปลี่ยนแปลงไปสำหรับในทุกน้ำหนัก ด้วยสมการต่อไปนี้

$$\Delta w_{i \rightarrow j} = r o_i o_j (1 - o_j) \beta_j$$

- เพิ่มค่าน้ำหนักที่เปลี่ยนแปลง สำหรับตัวอย่างอินพุตทั้งหมด และเปลี่ยนค่าน้ำหนักเมื่อผ่านการประมวลผลจนมาถึง output layer เรียบร้อยแล้ว ก็จะเป็นขั้นการประมาณความถูกต้อง โดยอาศัยข้อมูลที่ใช้ทดสอบ (testing data) ซึ่งคลาสที่แท้จริงของข้อมูลที่ใช้ทดสอบนี้จะถูกนำมาเปรียบเทียบกับคลาสที่หามาได้จากโมเดลนิวรอลเน็ตเวิร์ก เพื่อทดสอบความถูกต้อง ซึ่งโดยทั่วไปแล้วจะนำกระบวนการของนิวรอลเน็ตเวิร์กมาประยุกต์ใช้ในการทำนาย (Prediction) ดังตัวอย่างเช่น หายอดขายของเดือนถัดไปจากข้อมูลที่มีอยู่ หรือทำนายโรคจากอาการของคนไข้ในอดีต เป็นต้น

## 2) Database Segmentation หรือ Clustering

เป็นเทคนิคในการจำแนกกลุ่มข้อมูลใหม่ ที่มีลักษณะคล้ายกันไว้กลุ่มเดียวกัน โดยไม่มีการจัดกลุ่มของข้อมูลตัวอย่างไว้ล่วงหน้า ซึ่งแตกต่างจาก Classification ที่มีการกำหนดกลุ่มไว้แล้ว โดยกระบวนการของ Clustering นั้นจะต้องมีการกำหนดจำนวนกลุ่มที่ต้องการว่าต้องการจัดจำแนกกลุ่มออกเป็นกี่กลุ่ม

การศึกษาของ Clustering นั้นเป็นการศึกษาที่เรียกว่า Unsupervised Learning หรือ Segmentation โดยการทำ Clustering เองสามารถทำบนพื้นฐานของข้อมูลในอดีตได้เหมือนกัน แต่ไม่ต้องมีกระบวนการเรียนรู้จาก Training Data

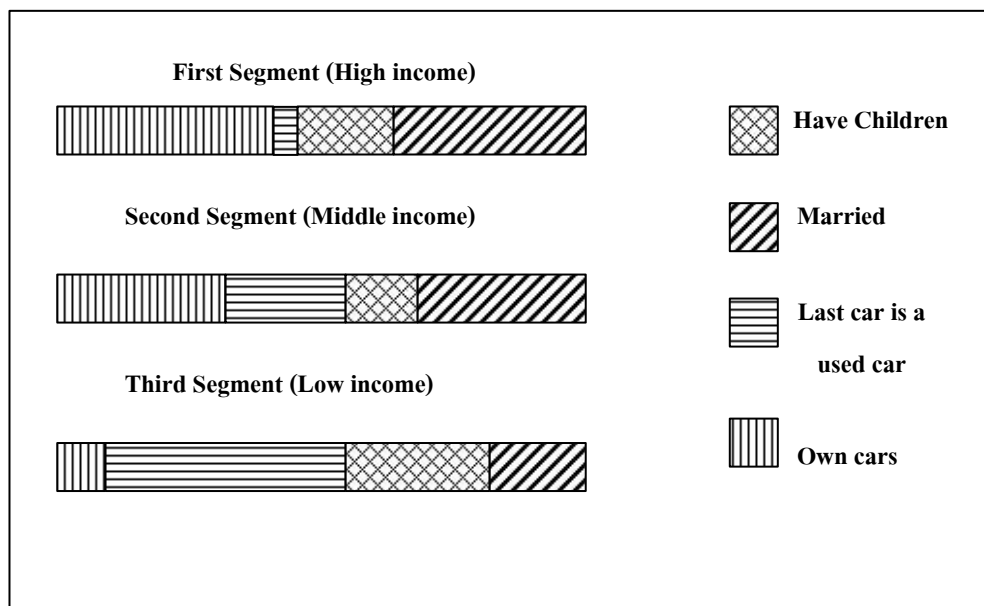
ตัวอย่างของ Clustering เช่น องค์กรต้องการทราบความเหมือนที่มีในกลุ่มลูกค้าของตนเอง เพื่อที่ว่าองค์กรจะสามารถเข้าใจลักษณะเฉพาะของกลุ่มลูกค้าเป้าหมายขององค์กร และสร้างกลุ่มของลูกค้า เพื่อองค์กรจะสามารถขายสินค้าได้ในอนาคต องค์กรจะทำการแยกกลุ่มของลูกค้าออกเป็นกลุ่ม ๆ โดยเทคนิคของ Clustering พยายามมองหาความเหมือนและความแตกต่างภายในกลุ่มของข้อมูลและแบ่งกลุ่มต่าง ๆ ออกเป็นส่วน ๆ (มหาวิทยาลัยศิลปากร, วิไลวรรณ เสรษญศิริ, 2551)

ตัวอย่างเช่น บริษัทจำหน่ายรถยนต์ได้แยกกลุ่มลูกค้าออกเป็น 3 กลุ่ม คือ

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

และภายในแต่ละกลุ่มยังแยกออกได้ ดังนี้

- Have Children
- Married
- Last car is a used car
- Own cars



ภาพที่ 2-9 ตัวอย่างการทำ Clustering (มหาวิทยาลัยขอนแก่น, รัตติกาล ทาโบราณ: 7)

จากข้อมูลข้างต้นทำให้ทางบริษัทรู้ว่าเมื่อมีกลุ่มลูกค้าในลักษณะต่างๆ เข้ามาที่บริษัท ควรจะเสนอขายรถประเภทใด เช่น ถ้าเป็นกลุ่มผู้มีรายได้สูงควรจะเสนอรถใหม่ เป็นรถครอบครัว ขนาดใหญ่พอสมควร แต่ถ้าเป็นผู้มีรายได้ค่อนข้างต่ำควรเสนอรถมือสองขนาดเล็ก เป็นต้น

### 3) Link Analysis หรือ Association

หลักการงานของวิธีนี้คือ การค้นหาความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่ที่มีอยู่ เพื่อนำไปใช้ในการวิเคราะห์หรือทำนายปรากฏการณ์ต่างๆ เช่น การวิเคราะห์การซื้อสินค้าของลูกค้า (Market Basket Analysis) ซึ่งประเมินจากข้อมูลในตารางที่ได้เก็บรวบรวมไว้แล้ว คือ รายการทั้งหมดที่ลูกค้าซื้อต่อครั้งที่ซูเปอร์มาร์เก็ต สามารถใช้ Input Device โดยอาจใช้ Bar Code Scanner ซึ่งมีหลายงานด้วยกัน เช่น ซูเปอร์มาร์เก็ต การเตรียม Inventory การวางแผนการเรียง Shelf การทำ Mailing List สำหรับ Direct Mail และการวางแผนเพื่อจัด Promotion สนับสนุนการขาย เป็นต้น ซึ่งตัวอย่างของ Association เช่น อาจพบว่า 75 % ของผู้ซื้อน้ำอัดลมจะซื้อข้าวโพดคั่ว

ด้วย ซึ่งผลการวิเคราะห์ที่ได้มานั้น จะกลายเป็นคำตอบของปัญหา ถือว่าเป็นการวิเคราะห์แบบ “กฎความสัมพันธ์” (Association Rule) เพื่อหาความสัมพันธ์ของข้อมูล

จากข้อความดังกล่าวข้างต้น สรุปได้ว่า Association เป็นเทคนิคที่ใช้ในการค้นพบองค์ความรู้ใหม่ ด้วยการเชื่อมโยงกลุ่มของข้อมูลที่เกิดขึ้นในเหตุการณ์เดียวกันไว้ด้วยกัน โดยการดูจากความบ่อยหรือจำนวนครั้งของการเกิดเหตุการณ์นั้น ซึ่งโดยทั่วไปผลลัพธ์จะออกมาในรูปแบบของกฎ if.....then.... คือ ถ้าเกิดเหตุการณ์หนึ่ง แล้วจะเกิดอีกเหตุการณ์หนึ่งตามมา เช่นลูกค้าทั่วไปเมื่อซื้อขนมขบเคี้ยวแล้วมักจะซื้อน้ำอัดลมตามมา (หมายถึงซื้อสินค้าที่คู่กันมีความสัมพันธ์กัน) เป็นต้น

ปัจจุบันได้มีการนำเทคนิค Association Rule มาใช้เป็นจำนวนมากมากในการวิเคราะห์ข้อมูลการตลาด (Market Analysis) เช่น การวิเคราะห์หาพฤติกรรมของการซื้อของลูกค้า นั่นคือ การหากฎการซื้อสินค้าประเภทหนึ่ง แล้วจะซื้อสินค้าอีกประเภทหนึ่งด้วย ยกตัวอย่างเช่น ถ้ามีสินค้าอยู่ 3 ประเภทคือ กาแฟ น้ำตาล และครีมเทียม เราอาจหาความสัมพันธ์ที่ได้ดังนี้ (บุญเกียรติ สว่างมณีเจริญ, 2552)

1. ถ้าลูกค้าซื้อ กาแฟ แล้วซื้อน้ำตาลด้วย
2. ถ้าลูกค้าซื้อน้ำตาล แล้ว ซื้อกาแฟด้วย
3. ถ้าลูกค้าซื้อ กาแฟและน้ำตาล แล้วซื้อครีมเทียมด้วย
4. ถ้าลูกค้าซื้อ น้ำตาล และครีมเทียมแล้วซื้อกาแฟด้วย

กฎความสัมพันธ์ร่วมของกลุ่มข้อมูล (Association Rule) ในระบบฐานข้อมูลขนาดใหญ่ การวิเคราะห์ความสัมพันธ์ของกลุ่มข้อมูลในลักษณะที่มี attribute หนึ่งแล้วจะมี attribute หนึ่งด้วย นั้น สามารถแสดงได้ในรูปของกฎความสัมพันธ์ (Association Rule) โดยอาศัยค่าสนับสนุน (Support) และความเชื่อมั่น (Confidence) ในการวัดความเชื่อถือของแต่ละกฎตัวอย่างเช่น ในการค้นหาพฤติกรรมการซื้อของลูกค้าเราสามารถหากฎการซื้อในสถานะ “เมื่อซื้อสินค้าประเภท A แล้วจะซื้อสินค้าประเภท B ด้วย” ความเชื่อถือของกฎดังกล่าว สามารถวัดได้ด้วยจำนวนของลูกค้าที่มีพฤติกรรมซื้อตรงตามกฎเทียบกับจำนวนลูกค้าทั้งหมด และจำนวนลูกค้าซื้อสินค้าทั้งประเภท A และ B เทียบกับจำนวนลูกค้าที่ซื้อสินค้าประเภท A ซึ่งหมายถึงว่าสนับสนุนและค่าความเชื่อมั่นของกฎความสัมพันธ์นั่นเอง (จิรายุทธ ไชยจรรูณิช, 2546)

จากทฤษฎีและเทคนิคของดาต้าไมน์นิ่งที่ได้กล่าวไปแล้วนั้น ทางผู้วิจัยได้เลือกใช้เทคนิคในการจัดกลุ่มข้อมูล และการทำนาย (Classification และ Predictive) โดยใช้วิธีการของ Decision tree เพื่อหาพฤติกรรมการเรียนของนักศึกษา และเลือกใช้วิธีการของ Neural Network ในการสร้างโมเดลเพื่อใช้ในการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์

### 2.3.5 การนำ Data Mining ไปประยุกต์ใช้กับงานด้านต่างๆ

เทคนิคการทำดาต้าไมนิง (Data Mining) นั้น สามารถนำไปวิเคราะห์ข้อมูล เพื่อช่วยประกอบการตัดสินใจ และแก้ไขปัญหาในธุรกิจต่างๆ ได้เป็นจำนวนมาก ดังนี้

1) ทางด้านการตลาด สามารถนำเทคนิคการทำดาต้าไมนิงไปใช้ในการทำนายผลการตอบสนองกับการเปิดตัวสินค้าใหม่, การทำนายยอดขายเมื่อมีการลดราคาสินค้า และการทำนายกลุ่มลูกค้าที่น่าจะใช้สินค้าของเรา เป็นต้น ซึ่งกระบวนการทำนายเหล่านี้ จะต้องอาศัยข้อมูลเกี่ยวกับพฤติกรรมการใช้สินค้าของลูกค้าในอดีต

2) ทางด้านการเงินการธนาคาร สามารถนำเทคนิคในการทำดาต้าไมนิงไปช่วยในการคาดการณ์ถึงโอกาสในการชำระหนี้ของลูกค้า, ค้นหาลูกค้าที่ขาดคุณภาพ เพื่อหลีกเลี่ยงความเสี่ยงในการปล่อยกู้ หรือค้นหาลูกค้าชั้นดี เพื่อเสนอการปล่อยกู้ นอกจากนี้ยังสามารถใช้ในการทำนายแนวโน้มของพฤติกรรมการใช้บัตรเครดิตของลูกค้าได้โดยการเก็บพฤติกรรมการใช้บัตรเครดิตของลูกค้าในอดีต หรือทำการเก็บลักษณะข้อมูลของลูกค้าในแต่ละกลุ่ม ทั้งกลุ่มลูกค้าชั้นดี และกลุ่มลูกค้าที่ขาดคุณภาพ เพื่อนำมาเป็นพื้นฐานในการจัดกลุ่มให้กับข้อมูลที่เข้ามาใหม่ ซึ่งกระบวนการต่างๆ เหล่านี้ส่วนมากจะใช้เทคนิค Classification และเทคนิค Association เป็นต้น

3) ทางด้านสถานีโทรทัศน์หรือวิทยุ สามารถนำเทคนิคการทำดาต้าไมนิงไปใช้ในการค้นหารายการที่ดี และเหมาะสมต่อช่วงเวลามากที่สุด เพื่อทำการวางผังรายการในแต่ละเดือน โดยจะต้องอาศัยข้อมูลพฤติกรรมชมรายการของลูกค้ามาวิเคราะห์ ว่ากลุ่มลูกค้ากลุ่มใดมักจะเลือกดูรายการในช่วงไหนมากที่สุด

4) ทางด้านการค้าขาย สามารถนำเทคนิคในการทำดาต้าไมนิง ไปช่วยในการทำนายผลกำไร เมื่อทำการลงทุนซื้อสินค้ามาเพื่อขาย และสามารถนำมาใช้ในการค้นหาจุดคุ้มทุน เช่น มีการนำรายได้ที่ได้จากการขายในอดีตมาทำนายยอดขายในอนาคต เพื่อทำการวางแผนการผลิตและการขายให้มีความเหมาะสมมากยิ่งขึ้น

5) ทางด้านโรงพยาบาล สามารถนำเทคนิคดาต้าไมนิง มาใช้ในการทำนายอาการของโรคได้อย่างแม่นยำยิ่งขึ้น เพื่อกำหนดวิธีการรักษาได้อย่างเหมาะสมที่สุด เช่น นำข้อมูลอาการของผู้ป่วยในอดีตมาหาความสัมพันธ์โดยใช้เทคนิค Association เพื่อหาว่าส่วนมาก เมื่อผู้ป่วยมีอาการอย่างใดอย่างหนึ่งแล้ว มักจะเกิดอาการใดตามมา ส่งผลให้มีการวางแผนในการรักษาโรคดังกล่าวได้ดียิ่งขึ้น

6) ธุรกิจประกันภัย สามารถนำเทคนิคดาต้าไมนิงมาใช้ในการค้นหาผู้เอาประกันที่มีแนวโน้มจะซื้อประกันเพิ่ม โดยสามารถนำข้อมูลลูกค้า และพฤติกรรมการใช้ประกันของลูกค้ามา

ทำการจัดกลุ่มลูกค้า แล้วนำไปเป็นพื้นฐานในการวิเคราะห์หากกลุ่มลูกค้าชั้นเยี่ยม หรือกลุ่มลูกค้าที่มีแนวโน้มจะซื้อประกันเพิ่ม

7) ธุรกิจทางการศึกษา ซึ่งในปัจจุบันนี้ได้เริ่มนำเทคนิคดาต้าไมน์นิ่งมาใช้ในสถาบันการศึกษามากขึ้น เช่นการนำเทคนิคดาต้าไมน์นิ่งมาช่วยในการวางแผนการประชาสัมพันธ์ หรือหาแนวทางการแนะแนวการศึกษาให้กับสถาบันการศึกษาต่างๆ นอกจากนี้ยังมีการนำข้อมูลผลการเรียนที่ผ่านมา รวมทั้งผลการประเมินการสอนมาวิเคราะห์หาประสิทธิภาพในการศึกษา เพื่อนำไปใช้เป็นแนวทางในการพัฒนาการเรียนการสอนต่อไปในอนาคตได้ เป็นต้น

ในการวิจัยครั้งนี้ มีการนำเทคนิคต่างๆ ของดาต้าไมน์นิ่งมาประยุกต์ใช้ในการวิเคราะห์เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ โดยนำเทคนิคดาต้าไมน์นิ่งมาใช้ในการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา และนำผลที่ได้นั้น ไปประกอบการตัดสินใจ เพื่อหาแนวทางการศึกษาที่เหมาะสมสำหรับนักศึกษา โดยการนำข้อมูลของนักศึกษามาวิเคราะห์ ซึ่งแบ่งเป็น 2 ส่วน คือ ส่วนแรก เป็นฐานข้อมูลการลงทะเบียนเรียนและผลการเรียนในแต่ละวิชา และส่วนที่ 2 เป็นฐานข้อมูลส่วนตัวของนักศึกษา เช่น อายุ เพศ ที่อยู่ ประวัติการศึกษา เกรดเฉลี่ยสะสม เป็นต้น

## 2.4 อาจารย์ที่ปรึกษา (Advisor)

จากคู่มืออาจารย์ที่ปรึกษา วิทยาลัยราชพฤกษ์ กล่าวว่า ระบบอาจารย์ที่ปรึกษานั้นถือว่าเป็นสิ่งจำเป็นอย่างยิ่งในการช่วยเหลือนักศึกษาแก้ไขปัญหาต่างๆ โดยอาจารย์ที่ปรึกษาจะให้คำแนะนำ และคำปรึกษาเพื่อให้นักศึกษามีความพร้อมทั้งร่างกายและจิตใจในการเรียน การร่วมกิจกรรมต่างๆ และการใช้ชีวิตในสังคมได้อย่างมีความสุข

### 2.4.1 วัตถุประสงค์ของอาจารย์ที่ปรึกษา

- 1) เพื่อให้คำปรึกษาแนะนำทางด้านวิชาการที่เกี่ยวกับหลักสูตร การเลือกวิชาเรียน การลงทะเบียนเรียน วิธีการเรียน และการวัดผล เพื่อให้นักศึกษาเลือกเรียนได้ตรงกับความสามารถ
- 2) เพื่อให้คำแนะนำเกี่ยวกับกฎระเบียบข้อบังคับต่างๆ ของวิทยาลัย รวมไปถึงการให้คำแนะนำเกี่ยวกับการบริการ และสวัสดิการต่างๆ ที่ทางวิทยาลัยจัดบริการให้กับนักศึกษา
- 3) เพื่อเป็นที่พึ่งของนักศึกษา ทั้งในการรับฟังปัญหา และการให้คำปรึกษา ข้อเสนอแนะแก่นักศึกษาในการแก้ไขปัญหาต่างๆ ได้
- 4) เพื่อช่วยในการพัฒนาบุคลิกภาพของนักศึกษาทางด้านวิชาการ, วิชาชีพ, การเข้าสังคม รวมทั้งด้านร่างกายและจิตใจ เพื่อเตรียมความพร้อมให้นักศึกษาก้าวออกไปสู่สังคมได้เป็นอย่างดี
- 5) เพื่อสร้างความสัมพันธ์ ความเข้าใจอันดี ระหว่างนักศึกษา คณาจารย์ และวิทยาลัย

## 2.4.2 บทบาทหน้าที่ของอาจารย์ที่ปรึกษา

### 1) ด้านวิชาการ

- 1.1) พิจารณาคำร้องของนักศึกษา และดำเนินการตามระเบียบของวิทยาลัย
- 1.2) ให้คำปรึกษาแนะนำนักศึกษาเกี่ยวกับหลักสูตรและการเลือกวิชาเรียน
- 1.3) ให้คำแนะนำนักศึกษา เกี่ยวกับการลงทะเบียนเรียน รวมถึงดูแลการลงทะเบียนของนักศึกษาให้เป็นไปตามกำหนดการและระเบียบข้อบังคับของวิทยาลัย
- 1.4) ให้คำปรึกษานักศึกษาเพื่อเลือกวิชาเอกเลือกและวิชาเลือกเสรี และการวางแผนการศึกษาให้สอดคล้องกับวัตถุประสงค์ของสาขาวิชา
- 1.5) ให้คำปรึกษานักศึกษาเกี่ยวกับวิธีการเรียน การค้นคว้า และติดตามผลการเรียนของนักศึกษาอย่างสม่ำเสมอ
- 1.6) ทักท้วงการลงทะเบียนของนักศึกษา กรณีการลงทะเบียนที่ไม่เหมาะสม
- 1.7) ให้คำปรึกษาแนะนำหรือตักเตือนเมื่อผลการเรียนของนักศึกษาดำลง
- 1.8) ให้คำปรึกษา ช่วยเหลือนักศึกษา เพื่อแก้ไขปัญหาในการเรียนวิชาต่างๆ
- 1.9) ให้ความรู้ ความเข้าใจเกี่ยวกับระเบียบ ข้อบังคับ ประกาศ และบริการต่างๆ ของวิทยาลัยที่เกี่ยวข้องกับนักศึกษา

1.10) ให้คำปรึกษาแนะนำเกี่ยวกับการคิดค่าระดับคะแนนเฉลี่ยของนักศึกษา

1.11) ให้คำปรึกษาแนะนำเกี่ยวกับการศึกษาต่อในระดับที่สูงขึ้น

### 2) ด้านบริการและพัฒนานักศึกษา

- 2.1) ให้คำปรึกษาเกี่ยวกับปัญหาส่วนตัว อย่างเช่น สุขภาพอนามัยทั้งสุขภาพกายและสุขภาพจิต
- 2.2) ให้คำปรึกษาเกี่ยวกับปัญหาสังคม เช่น การปรับตัว การคบเพื่อนในสังคม
- 2.3) ให้คำปรึกษาเกี่ยวกับการพัฒนาบุคลิกภาพ ความประพฤติ และจริยธรรม
- 2.4) ให้คำปรึกษาเกี่ยวกับปัญหาด้านอาชีพ ได้แก่ การให้ข้อมูลต่างๆ เช่น ลักษณะงาน สภาพแวดล้อมของงาน ตลาดแรงงาน
- 2.5) ให้คำแนะนำเกี่ยวกับการเข้าร่วมกิจกรรมนักศึกษา
- 2.6) ให้คำแนะนำ ควบคุมและสอดส่องนักศึกษาเกี่ยวกับการแต่งกาย ความประพฤติ และมารยาททางสังคม

### 3) ด้านอื่นๆ

3.1) ประสานงานกับอาจารย์ผู้สอน และฝ่ายต่างๆ ที่เกี่ยวข้อง

3.2) กำหนดเวลาให้นักศึกษาเข้าพบ เพื่อขอคำปรึกษา แนะนำอย่างสม่ำเสมอ



3.3) เก็บข้อมูลรายละเอียดของนักศึกษาที่อยู่ในความรับผิดชอบ เพื่อนำไปเป็นข้อมูลพื้นฐานสำหรับให้คำปรึกษา และช่วยเหลือนักศึกษา

3.4) ให้ความร่วมมือ รวมถึงสร้างสัมพันธภาพและความเข้าใจอันดีระหว่างนักศึกษาคณาจารย์ สาขาวิชา คณะฯ และวิทยาลัย

3.5) ให้การรับรองแก่นักศึกษาที่มีความประพฤติเหมาะสมได้ตามสมควร

3.6) ชี้แจงให้นักศึกษาเข้าใจหน้าที่ของอาจารย์ที่ปรึกษา และหน้าที่ของนักศึกษาต่ออาจารย์ที่ปรึกษา

3.7) กรอกแบบรายงานการทำงานที่อาจารย์ที่ปรึกษามายังหัวหน้าสาขาวิชาเกี่ยวกับปัญหาต่างๆ ของนักศึกษา

## 2.5 ผลสัมฤทธิ์ทางการเรียน (Education achievement)

สัมฤทธิ์ผลทางการเรียน (Education achievement) หมายถึง ผลที่เกิดจากการจัดการศึกษา จัดได้ว่า เป็นปัจจัยหนึ่งที่สามารถบอกถึงคุณภาพการศึกษา ดังที่ กู๊ด (Good 1973 :7 ,อ้างถึงใน ชุติพร ฤทธิเดช 2547 : 56) กล่าวถึงสัมฤทธิ์ผลทางการเรียน สรุปได้ว่า หมายถึงความรู้หรือทักษะอันเกิดจากการเรียนรู้ที่ได้เรียนมาแล้วที่ได้จากผลการสอนของอาจารย์ผู้สอน ซึ่งอาจพิจารณาจากคะแนนสอบที่กำหนด ให้คะแนนที่ได้จากงานที่ครูมอบหมายให้หรือทั้งสองอย่าง จัดว่าเป็นการตรวจสอบระดับความสามารถของนักศึกษาว่าเรียนแล้วมีความรู้เท่าใด สามารถวัดได้โดยใช้แบบทดสอบต่าง ๆ เช่น ใช้ข้อสอบวัดผลสัมฤทธิ์ ข้อสอบวัดภาคปฏิบัติ เป็นต้น

จากความหมายข้างต้นสรุปได้ว่า สัมฤทธิ์ผลทางการเรียน หมายถึง ผลที่ได้จากการวัดความรู้ที่ได้จากการเรียนรู้ในเนื้อหาสาระที่ได้จากผลการสอนของอาจารย์ผู้สอน เพื่อตรวจสอบว่ามีความเข้าใจและเกิดการเรียนรู้ในเนื้อหานั้นๆ มากขึ้นเท่าใด มีความสามารถด้านไหน ซึ่งสามารถทำการวัดผลได้จากแบบทดสอบวัดสัมฤทธิ์ผลทางการเรียนในลักษณะต่างๆ ทั้งทางทฤษฎี และเชิงปฏิบัติ เพื่อบอกถึงคุณภาพการศึกษา

## 2.6 งานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการศึกษาค้นคว้าบทความ และงานวิจัยที่เกี่ยวข้อง โดยมีรายละเอียด ดังต่อไปนี้

### 2.6.1 การใช้เทคนิคดาต้าไมน์นิ่ง เพื่อพัฒนาคุณภาพการศึกษาคณะวิศวกรรมศาสตร์

โดย กฤษณะ ไวยมัย, ชิดชนก ส่งศิริ และธนาวินท์ รักรธรรมานนท์ อาจารย์ประจำสาขาวิชาวิศวกรรมคอมพิวเตอร์ และนิติศติปริญญาโทวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

บทความนี้เป็นการศึกษาและวิเคราะห์ระบบฐานข้อมูลนิติต โดยนำความรู้ทางด้านค้ำ  
ไมน์นิ่งมาประยุกต์ใช้กับข้อมูลนิติต คณะวิศวกรรมศาสตร์ เพื่อเป็นแนวทางในการแก้ไขปัญหา  
ต่างๆ เช่น ปัญหาการเลือกสาขาวิชาไม่ตรงกับความสามารถที่แท้จริง ปัญหาผลการเรียนของนิติต  
ตกต่ำจนต้องออกจากสถาบันการศึกษา อันเป็นผลทำให้ไม่ได้มาซึ่งบุคลากรที่มีความสามารถสูงสุด

โดยบทความนี้ได้นำเทคนิคค้ำไมน์นิ่งมาประยุกต์ใช้ในการช่วยนิติตเลือกสาขาวิชาที่  
เหมาะสม สามารถทำได้ 2 แนวทางที่มีผลการทดสอบความถูกต้องค่อนข้างสูง คือ การทำเทคนิค  
การจำแนกประเภทข้อมูล และเทคนิคการพยากรณ์ข้อมูลมาสร้างโมเดล โดยมีรายละเอียดดังนี้

1) โมเดลการจำแนกประเภทข้อมูลสำหรับแต่ละสาขาวิชา โดยพิจารณาว่านิติตเหมาะสม  
กับสาขาวิชานั้นๆ หรือไม่ ในงานวิจัยครั้งนี้ใช้ข้อมูล 2 ส่วน ส่วนแรกคือ ข้อมูลประวัติส่วนตัวของ  
นิติต เช่น ชื่อ ที่อยู่ ภูมิลำเนา อายุ ฯลฯ ข้อมูลอีกส่วนหนึ่งคือ ข้อมูลการลงทะเบียนของนิติต จะมี  
การเลือกสาขาวิชาขึ้นในปี 2 ดังนั้น ในการวิจัยนี้จึงนำเฉพาะวิชาในปีที่ 1 ของการศึกษา (9 วิชา) มา  
เป็นตัวพิจารณา และมีการจัดกลุ่มของเกรด เป็น High, Medium, Low เพื่อลดการกระจายของเกรด

ในการสร้างแต่ละโมเดล มีการแบ่งข้อมูล 70% จากข้อมูลทั้งหมดเป็นข้อมูลสอน  
ระบบ (training data) เพื่อสร้างโมเดล และข้อมูล 30% ที่เหลือเป็นข้อมูลที่ใช้ทดสอบ (testing data)  
เพื่อทดสอบความถูกต้อง โดยในโมเดลการจำแนกประเภทข้อมูลนี้มี 2 คลาส คือ GOOD และ  
BAD โดยมีเกณฑ์ว่า GOOD คือนิติตที่มีเกรดเฉลี่ยสะสมอยู่ในอันดับ 40% แรกของแต่ละสาขาวิชา  
และ BAD คือนิติตที่มีเกรดเฉลี่ยสะสมอยู่ในอันดับ 40% สุดท้ายในแต่ละสาขาวิชา ซึ่งเปอร์เซ็นต์นี้  
สามารถปรับเปลี่ยนให้เหมาะสมได้ ณ ที่นี้ขอยกตัวอย่างการนำข้อมูลนิติตภาคไฟฟ้ามาสร้างโมเดล

โดยได้ทำการทดสอบข้อมูลนิติตคนหนึ่งกับโมเดลของทุกสาขาวิชา และเลือกเฉพาะ  
สาขาวิชาที่คลาสปลาซทางมีสัดส่วนของ GOOD มากกว่า BAD ถ้าผลออกมาว่ามีหลายสาขาวิชาที่  
เหมาะสม สามารถเลือกสาขาวิชาที่ดีที่สุดได้ โดยพิจารณาจากสัดส่วนของ GOOD ในโมเดลที่มี  
มากกว่าเป็นหลัก ได้ผลการทดสอบถูกต้องเฉลี่ย 84.58% ในทุกโมเดล

2) โมเดลการพยากรณ์ข้อมูลในแต่ละสาขาวิชา มีการปรับเปลี่ยนข้อมูลจากการเตรียม  
ข้อมูลนำไปสร้างโมเดลการจำแนกประเภทข้อมูล คือ ใช้ผลการเรียนในวิชาต่างๆ ด้วยตัวเลขจริง  
ไม่ต้องจัดกลุ่มเพื่อลดการกระจายของเกรด

สร้างโมเดลโดยนำผลการเรียนและลักษณะต่างๆ ของนิติตแต่ละสาขาวิชาเป็นตัว  
พิจารณา และสร้างโมเดลเพื่อทำนายเกรดเฉลี่ยตามลักษณะในแต่ละสาขาวิชา หลังจากสร้างโมเดล  
ของทุกสาขาวิชาแล้ว เมื่อมีนิติตคนหนึ่งต้องการทราบว่าตนเหมาะสมกับสาขาวิชาใด จะนำข้อมูล  
ต่างๆ ของนิติตคนนั้นมาเป็นปัจจัยในแต่ละโมเดล เพื่อทำนายค่าเกรดเฉลี่ยที่ต้องการออกมา เมื่อนำ  
ข้อมูลนิติตมาผ่านทุกโมเดลแล้ว จะได้เกรดเฉลี่ยของนิติตคนนั้นตามโมเดลของแต่ละสาขาวิชา

โมเดลนี้สร้างขึ้นเพื่อที่นิตินิติจะได้ทราบว่า ผลการเรียนรู้ของนิตินิติเป็นอย่างไร เมื่อเปรียบเทียบกับนิตินิติส่วนใหญ่ในแต่ละสาขาวิชา โดยกล่าวได้ว่า ถ้านิตินิติเรียนในสาขาวิชาวิศวกรรมคอมพิวเตอร์ แล้วจะมีแนวโน้มผลการเรียนตอนจบการศึกษาสูงกว่าสาขาวิชาอื่นๆ ซึ่งการนำเสนอแบบนี้จะไม่บอกนิตินิติว่า สาขาวิชาที่เหมาะสมที่สุด แต่ระบบจะนำเสนอผลที่ได้ในทุกสาขาวิชาให้กับนิตินิติ เพื่อที่นิตินิติสามารถนำไปพิจารณาประกอบกับความต้องการของนิตินิติได้อย่างเหมาะสม

3) การใช้เทคนิคการค้นหากฎความสัมพันธ์มาช่วยทำนายแนวโน้มเกรด โดยประยุกต์ใช้เทคนิคการสืบค้นกฎความสัมพันธ์กับข้อมูลผลการเรียนนิตินิติ โดยหาความสัมพันธ์ของผลการเรียนในแต่ละวิชาที่ส่งผลต่อกัน ซึ่งจะช่วยให้ทราบว่าวิชาใดบ้างที่มีผลต่อวิชาที่ต้องการจะทำนายเกรด

ซึ่งผลที่ได้จากการใช้เทคนิคการค้นหากฎความสัมพันธ์จะอยู่ในรูปความสัมพันธ์ของวิชาต่างๆ มากมาย โดยมีความสัมพันธ์ที่ได้ในแต่ละสาขามีมากกว่า 200,000 กฎ ซึ่งต้องตัดความสัมพันธ์บางส่วน โดยมีการกำหนดค่าสนับสนุนต่ำสุด (minimum support) และค่าความมั่นใจต่ำสุด (minimum confidence) เพื่อเลือกเฉพาะความสัมพันธ์ที่มีจำนวนนิตินิติมาก และน่าเชื่อถือ นอกจากนี้ต้องกำจัดความสัมพันธ์ที่ลำดับของวิชาไม่ปฏิบัติตามข้อกำหนดหลักสูตรด้วยเช่นกัน

วัตถุประสงค์ของโมเดลนี้ มีการนำเสนอการทำนายเกรดที่มีผลต่อวิชาอื่นเป็นช่วงนั้น เพื่อต้องการชี้แนะแนวทางในการเรียนในวิชานั้นๆ อย่างเหมาะสม ซึ่งการที่นิตินิติทราบว่า ผลการเรียนรู้มีแนวโน้มจะออกมาดี ปานกลาง หรือ ไม่ดีนั้น เป็นการนำเสนอที่มีผลดีต่อการเรียนของนิตินิติ

บทสรุปของบทความนี้ ได้นำเสนอการนำเทคนิคดาต้าไมน์นิ่งมาประยุกต์ใช้เพื่อช่วยพิจารณาหารูปแบบความสัมพันธ์ของข้อมูลในมุมมองต่างๆ เพื่อที่จะได้นำสิ่งที่ได้เป็นประโยชน์ที่ได้จากดาต้าไมน์นิ่งไปเป็นส่วนหนึ่งในการพัฒนาคุณภาพการศึกษาต่อไป ซึ่งผลลัพธ์ที่ได้จากงานวิจัยนี้ค่อนข้างเป็นที่น่าพอใจ โดยมีเปอร์เซ็นต์ความถูกต้องค่อนข้างสูง แต่มีปัญหาบางประการ เช่น จำนวนข้อมูลในบางสาขาวิชาที่มีปริมาณค่อนข้างน้อย ทำให้โมเดลที่ได้ไม่แม่นยำเท่าที่ควร เป็นต้น

## 2.6.2 ตัวแบบพยากรณ์ลักษณะความเหมาะสมของนักศึกษาใหม่สาขาวิชาเทคโนโลยีสารสนเทศ ด้วยกฎการจำแนกประเภทเชิงความสัมพันธ์

โดยจามรรกุล เหล่าเกียรติกุล, จิรารัตน์ สิทธิวรชาติ ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

การวิจัยนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบในการพยากรณ์ลักษณะความเหมาะสมต่อการเข้าศึกษาของนักศึกษาใหม่ในสาขาวิชาเทคโนโลยีสารสนเทศ ด้วยเทคนิคกฎการจำแนกเชิงความสัมพันธ์ โดยใช้ข้อมูลนักศึกษาสาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏจันทรเกษม เป็นกรณีศึกษา เพื่อพยากรณ์คุณลักษณะความเหมาะสมของนักศึกษา โดยใช้ระดับผลการเรียนปัจจุบัน

ซึ่งมีการนำเทคนิคของดาต้าไมน์นิ่งมาประยุกต์ใช้ คือ กฎการจำแนกประเภทเชิงความสัมพันธ์ (CARs) เป็นเทคนิคของกรรมวิธีดาต้าไมน์นิ่ง ที่ผสมผสานระหว่างเทคนิคการวิเคราะห์เพื่อสร้างกฎเชื่อมโยงความสัมพันธ์ (Association-Rules Discovery) เข้ากับวิธีการจำแนกกลุ่มข้อมูล (Data Classification)

### 1) ตัวแบบพยากรณ์ลักษณะความเหมาะสมของนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ

เริ่มจากการจัดเตรียมข้อมูล โดยใช้ข้อมูลนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏจันทรเกษม จากฐานข้อมูล ตามมาตรฐานของสำนักงานคณะกรรมการอุดมศึกษา (สกอ.) จำนวน 1,250 ระเบียบ โดยกำหนดตัวแปรที่จะใช้ในการวิจัยทั้งสิ้น 18 ตัวแปร โดยนำข้อมูลดังกล่าวมา จัดเตรียมข้อมูลสำหรับการเรียนรู้ (Training data) และข้อมูลสำหรับการทดสอบตัวแบบ (Testing data) ให้เหมาะสมกับการวิเคราะห์ด้วยเทคนิคการเชื่อมโยงความสัมพันธ์ โดยปรับข้อมูลบางประเภทให้เป็นช่วงข้อมูลเพื่อง่ายต่อการวิเคราะห์ เช่น อายุ จัดเป็นช่วงอายุ และข้อมูลระดับผลการเรียน จัดเป็นกลุ่มระดับผลการเรียน เป็นต้น

### 2) การสร้างกฎจำแนกประเภทเชิงความสัมพันธ์

นำข้อมูลที่จัดเตรียมสำหรับกระบวนการเรียนรู้ (Training data) เป็นจำนวน 935 ระเบียบ นำเข้าสู่เทคนิคการวิเคราะห์และเชื่อมโยงความสัมพันธ์ เพื่อค้นหาความเชื่อมโยงของปัจจัยต่างๆ กับประเภทของข้อมูลที่ต้องการ ซึ่งในกระบวนการจะกำหนดเป็นค่าคงที่ตามระดับผลการเรียน 4 ระดับ คือ ระดับดีเยี่ยม, ระดับดี, ระดับปานกลาง และระดับต่ำ

ผลการดำเนินการพบว่ามีจำนวนกฎที่ถูกสร้างขึ้นทั้งสิ้น 8,308 กฎ โดยมีกฎที่เป็นไปตามค่าความเชื่อมั่นขั้นต่ำที่กำหนด (Minimum confidence) (0.9) 157 กฎ ซึ่งจำแนกกฎสำหรับจำแนกระดับความเหมาะสมจากระดับผลการเรียน ในระดับดีเยี่ยมจำนวน 13 กฎ ระดับดี 40 กฎ ระดับปานกลาง 61 กฎ และระดับต่ำ 43 กฎ

### 3) การทดสอบกฎจำแนกประเภทเชิงความสัมพันธ์

ในส่วนของการทดสอบข้อมูล ใช้ชุดข้อมูล ซึ่งเตรียมไว้สำหรับการทดสอบ (Testing data) จำนวน 315 ระเบียบ นำมาผ่านกระบวนการของตัวแบบที่ถูกสร้างขึ้น โดยที่สมมติให้ข้อมูลทดสอบเป็นข้อมูลนักศึกษาเข้าใหม่ ที่ต้องการคาดการณ์ว่าจะมีระดับความเหมาะสมอยู่ในระดับใด

เมื่อทำการทดสอบกฎจำแนกประเภทเชิงสัมพันธ์ที่สร้างขึ้นด้วยชุดข้อมูลทดสอบพบว่า กฎสำหรับจำแนกระดับความเหมาะสมของนักศึกษาในระดับดี เป็นกฎที่มีความสามารถในการจำแนกข้อมูลได้ดีที่สุดถึงร้อยละ 98.97 โดยตัวแบบพยากรณ์ฯ ดังกล่าว สามารถจำแนกระดับความเหมาะสมของนักศึกษาโดยรวมได้ร้อยละ 83.81 แม้ว่าจะพบข้อจำกัดในเรื่องของการนำเข้า

ข้อมูลสำหรับเรียนรู้ ที่ระดับความเหมาะสมดีเยี่ยมและระดับต่ำ น้อยเกินไป ทำให้ประสิทธิภาพการพยากรณ์ความเหมาะสมของนักศึกษาประเภทดังกล่าวไม่ดีเท่าที่ควร

ทั้งนี้กฎการจำแนกประเภทเชิงความสัมพันธ์ระหว่างปัจจัยที่ค้นพบ รวมถึงลักษณะความเหมาะสมที่ได้ สามารถนำไปประยุกต์ใช้กับข้อมูลนักศึกษาในสาขาอื่นๆ ได้ และเป็นแนวทางสำหรับการจัดการเรียน กลุ่มเรียน รวมถึงรายวิชาในแผนการเรียนในหลักสูตรให้เหมาะสมกับผู้เรียน ซึ่งเป็นกระบวนการที่จะเกิดขึ้นในงานวิจัยต่อไป

### 2.6.3 การทำนายผลสำเร็จการศึกษาของนักศึกษาระดับอุดมศึกษาด้วยการเรียนรู้แบบเบย์และการทำเหมืองข้อมูล

โดย จิราพร ยิ่งกว่าชาติ, ประสงค์ ปราณีตพลกรัง สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยศรีปทุม และบุญเสริม กิจศิริกุล ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

งานวิจัยนี้ เป็นการนำเสนอทางเลือกเพื่อการทำนายผลสำเร็จการศึกษา โดยอาศัยเทคนิคการทำเหมืองข้อมูล (Data mining) คือ เทคนิคข่างานเบย์ ซึ่งเป็นพื้นฐานของการจัดหมวดหมู่มาใช้เพื่อวิเคราะห์ถึงตัวแปรที่มีผลต่อการทำนายผลสำเร็จการศึกษาของนักศึกษาระดับปริญญาตรี และในการทดสอบแบบจำลองที่ได้จะทำการทดสอบผลบนพื้นฐานวิธี k-fold Cross Validation

#### 1) ทำการรวบรวมและเลือกข้อมูลที่เหมาะสมกับอัลกอริทึมข่างานเบย์

ทำการกำหนดตัวแปรที่จะใช้ในงานวิจัย ซึ่งมีทั้งสิ้น 18 ตัวแปร จากข้อมูลของมหาวิทยาลัยราชภัฏอุบลราชธานีในฐานข้อมูลย้อนหลัง ปี 2534-2545 พบว่ามีจำนวนระเบียบที่มีความเหมาะสมในการสร้างแบบจำลองจำนวนทั้งหมด 5,140 ระเบียบ และข้อมูลของมหาวิทยาลัยศรีปทุม เป็นจำนวน 145 ระเบียบ

#### 2) การนำตัวแปรที่กำหนดมาเข้ากระบวนการสร้างแบบจำลอง

โดยทดลองกำหนดเซตย่อยๆ ที่จะสร้างความเป็นไปได้ในการทำนายผล แล้วสังเกตเซตที่ให้ค่าที่มีความแม่นยำมากที่สุด แล้วจึงนำมาสร้างเป็นแบบจำลอง งานวิจัยนี้ได้เลือกใช้โปรแกรม WEKA เพื่อสร้างแบบจำลอง ซึ่งผลที่ได้ทราบว่า ตัวแปรที่มีผลต่อการสำเร็จการศึกษาคือเกรดเฉลี่ยในชั้นปีแรก, อาชีพของมารดา และรายได้ของครอบครัว โดยแบบจำลองทำนายผลสำเร็จการศึกษาของมหาวิทยาลัยราชภัฏอุบลราชธานี และมหาวิทยาลัยศรีปทุม พบว่ามีลักษณะของตัวแปรที่มีผลต่อการศึกษามีความใกล้เคียงกัน

พบว่า ผลการทดสอบแบบจำลองที่ได้จากข้อมูลของมหาวิทยาลัยราชภัฏอุบลราชธานี มีความแม่นยำที่ได้ คือ 91.26% และผลการทดสอบแบบจำลองทำนายผลสำเร็จการศึกษาที่ได้จากข้อมูลของมหาวิทยาลัยศรีปทุม มีความแม่นยำที่ได้ คือ 90.84%

สรุปผลงานวิจัยเพื่อทำการทำนายผลสำเร็จการศึกษาของนักศึกษานั้น มีวัตถุประสงค์เพื่อให้มีความกระตือรือร้นในการเรียน เพื่อให้มีผลการเรียนดีขึ้น ซึ่งผลที่ได้จากการศึกษานี้ อาจจะช่วยเพิ่มจำนวนผู้สำเร็จการศึกษาให้มีจำนวนมากขึ้น จากผลการวิเคราะห์เบื้องต้น พบว่า การสร้างแบบจำลองด้วยเทคนิคงานเบย์ มีความแม่นยำในการทำนายผลถึง 91.26%

#### 2.6.4 การประยุกต์ใช้ดาต้าไมนิ่งในการบริหารลูกค้าสัมพันธ์สำหรับนักศึกษาระดับอุดมศึกษา

โดย ปรีชา ขามันสะบีดิน, บุญเสริม กิจศิริกุล, ปิยะวัฒน์ จิระพงษ์สุวรรณ และประสงค์ ปราณีตพลกรัง สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยศรีปทุม

งานวิจัยนี้เป็นการศึกษาและนำแนวความคิดของการทำเหมืองข้อมูล (Data Mining) มาออกแบบและพัฒนาใช้สำหรับการบริหารลูกค้าสัมพันธ์ (Customer Relationship Management: CRM) กับนักศึกษา ซึ่งเป็นกลยุทธ์ที่ช่วยวิเคราะห์ถึงพฤติกรรมของนักศึกษา ตลอดจนคาดคะเนถึงแนวโน้มของความเป็นไปได้ในอนาคต เพื่อประกอบการตัดสินใจในการจัดการกับปัญหานักศึกษาที่มีแนวโน้มที่จะถูกคัดชื่อออก เพราะฉะนั้นการบริหารลูกค้าสัมพันธ์ ช่วยในการรักษาสัมพันธ์ภาพที่ดีระหว่างนักศึกษาและมหาวิทยาลัย ซึ่งผลที่ได้คือ ข้อมูลด้านสถิติและรูปแบบของพฤติกรรมนักศึกษา สามารถนำไปใช้กำหนดภารกิจ วัตถุประสงค์ แผนกลยุทธ์ การจัดการของมหาวิทยาลัย

งานวิจัยในครั้งนี้เป็นการประยุกต์ใช้เทคนิคการทำเหมืองข้อมูล เพื่อค้นหาปัจจัยที่ส่งผลกระทบต่อความสำเร็จการศึกษา ของนักศึกษาที่ลงทะเบียนเรียนกับมหาวิทยาลัยสุโขทัยธร มาธิราช

แนวทางในการสร้างโมเดลสำหรับการจำแนกคุณลักษณะของนักศึกษา คือ โมเดลจำแนกประเภทข้อมูลโดยใช้ต้นไม้ตัดสินใจ (Decision Tree) บ่งบอกถึงคุณลักษณะของนักศึกษาและคลาสปลายทางแทนสถานภาพนักศึกษา เพื่อที่จะทำนายนักศึกษแต่ละสถานภาพ โดยมีวิธีการดังนี้

##### 1) การเตรียมข้อมูล

งานวิจัยนี้ใช้ข้อมูลนักศึกษาระดับปริญญาตรีมหาวิทยาลัยสุโขทัยธรมาธิราช จำนวน 94,244 คน โดยข้อมูลที่น่ามาใช้ประกอบด้วยข้อมูล 2 ส่วน ส่วนแรกคือ ข้อมูลประวัติส่วนตัวของนักศึกษา เช่น ชื่อ ที่อยู่ เพศ อายุ ฯลฯ ข้อมูลอีกส่วนคือ ข้อมูลการลงทะเบียนเรียนและเกรดนักศึกษา ประกอบด้วย ปีการศึกษา, ภาคการศึกษา, รหัสวิชา ประเภทการลงทะเบียน และเกรด โดยแบ่งกลุ่มนักศึกษาเป็นประเภทต่างๆ ตามสถานภาพของนักศึกษาออกเป็น 5 ประเภท คือ นักศึกษาที่สำเร็จการศึกษาแล้ว (Graduate), นักศึกษาที่กำลังศึกษาอยู่ (Active), นักศึกษาที่ถูกคัดชื่อออกกลางคัน (Drop-out), นักศึกษาที่ลาออก (Quit) และนักศึกษาที่เรียนครบ 3 เท้าหลักสูตร (Retired)

## 2) การดำเนินการวิเคราะห์ข้อมูลจากโมเดลที่สร้าง

การวิจัยในครั้งนี้ใช้โมเดลจำแนกประเภทข้อมูลโดยใช้ต้นไม้ตัดสินใจ (Decision Tree) โดยมีการแบ่งกลุ่มข้อมูลเป็น 2 ชุด คือ ชุดข้อมูลที่ใช้ในการเรียนรู้ หรือข้อมูลสอน จำนวน 52,362 คน และข้อมูลที่ใช้ในการทดสอบโมเดล จำนวน 41,882 คน ซึ่งมีอัตราส่วนใกล้เคียงกัน

ซึ่งผลการวิเคราะห์ไฟล์ข้อมูลทดสอบ และไฟล์ข้อมูลสอน ได้ความถูกต้องและความผิดพลาด ดังนี้

**ตารางที่ 2-1** การเปรียบเทียบความถูกต้องและความผิดพลาด

ไฟล์	จำนวนข้อมูล	ความถูกต้อง	ร้อยละ	ความผิดพลาด	ร้อยละ
ข้อมูลสอน	52,362	43,355	82.80	9,007	17.20
ข้อมูลทดสอบ	41,882	34,126	81.48	7,756	18.52

สรุปผลการวิจัย สามารถนำผลที่ได้จากการทำนายการสำเร็จการศึกษา มาช่วยในการวางแผนการศึกษาของนักศึกษาให้เหมาะสมกับตัวเอง สำหรับการลงทะเบียนเรียนแต่ละภาคการศึกษา และสามารถช่วยในการตัดสินใจการดำเนินงานของมหาวิทยาลัย ด้านการแนะแนวการศึกษา การจัดสอนเสริม และกิจกรรมการเรียนให้กับนักศึกษา วิธีการดังกล่าว อาจยังไม่ใช่วิธีที่ดีที่สุดในการวิเคราะห์แนวโน้ม ผู้วิจัยจะพัฒนา และนำวิธีการหรืออัลกอริทึมอื่นๆ มาใช้ในการวิจัยต่อไป

## บทที่ 3

### วิธีดำเนินการวิจัย

วิธีการดำเนินการวิจัยของ การใช้เทคนิคค้ำไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ ผู้วิจัยได้ทำการกำหนดขั้นตอนและวิธีการดำเนินการวิจัยออกเป็น 5 ขั้นตอน ดังนี้

- 3.1 การศึกษาข้อมูลที่เกี่ยวข้อง
- 3.2 รวบรวม คัดเลือก และจัดเตรียมข้อมูล
- 3.3 การสร้างโมเดลในการจัดกลุ่มข้อมูล และการพยากรณ์ข้อมูล
- 3.4 การทดสอบความถูกต้อง และความน่าเชื่อถือของโมเดล โดยการใช้ข้อมูลทดสอบ
- 3.5 แปลความหมาย และนำความรู้ที่ได้มาใช้

#### 3.1 การศึกษาข้อมูลที่เกี่ยวข้อง

การใช้เทคนิคค้ำไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ ผู้จัดทำได้มีการศึกษาค้นคว้าหาข้อมูล หรือความรู้ที่สำคัญที่นำมาใช้ในการดำเนินการวิจัย ดังนี้

3.1.1 ศึกษา และทบทวนวรรณกรรม หรือผลงานวิจัยที่เกี่ยวข้อง (Review Literature)

3.1.2 ศึกษาถึงโอกาสทางธุรกิจ หรือปัญหาที่เกิดขึ้นกับธุรกิจ เป็นการระบุดูขอบเขตของข้อมูลที่จะนำมาทำการวิเคราะห์ เพื่อนำมาประกอบการตัดสินใจในกระบวนการที่เกี่ยวข้อง ดังนี้

เนื่องจากวิทยาลัยราชพฤกษ์ เป็นสถาบันอุดมศึกษาที่เน้นการพัฒนาคุณภาพบัณฑิต ให้เป็นคนเก่ง มีความรู้ ทักษะ ความชำนาญในแต่ละสาขาวิชา ซึ่งสิ่งที่เป็นปัจจัยหลักสำคัญในการพัฒนาคุณภาพบัณฑิตนั้น คือ ทางวิทยาลัยได้ให้ความสำคัญต่อการดูแลและให้คำแนะนำในเรื่องการเรียนของ โดยจัดให้มีอาจารย์ที่ปรึกษาคอยให้คำแนะนำแนวทางในการเรียนให้แก่นักศึกษา ซึ่งในปัจจุบัน การตัดสินใจให้คำแนะนำของอาจารย์ที่ปรึกษา เป็นการพิจารณาจากผลการเรียนที่ผ่านมา โดยดูจากรายวิชาในแต่ละสาขาวิชาของนักศึกษาเป็นหลัก

การวิจัยครั้งนี้ จึงได้เห็นถึงโอกาสในการนำเครื่องมือที่เข้ามาช่วยในการวิเคราะห์ความถนัด และจุดอ่อนทางการศึกษาของนักศึกษาเป็นรายบุคคล โดยสามารถคาดการณ์ผลการศึกษาในสาขาวิชาอื่น นอกจากสาขาวิชาที่กำลังศึกษาอยู่ได้ ส่งผลให้กระบวนการตัดสินใจ ให้คำแนะนำของอาจารย์ที่ปรึกษา รวมทั้งกระบวนการพัฒนาคุณภาพการศึกษามีประสิทธิภาพมากยิ่งขึ้น จึงได้มีการนำเทคนิคค้ำไมน์นิ่งมาใช้ในการทำนายผลการเรียนของนักศึกษาในแต่ละสาขา วิชา และนำผลที่



ได้ไปประกอบการตัดสินใจ เพื่อหาแนวทางการศึกษาที่เหมาะสมสำหรับนักศึกษาแต่ละท่าน โดยข้อมูลที่น่ามาวิจัยนี้เป็นข้อมูลของนักศึกษาคณะบริหารธุรกิจ วิทยาลัยราชพฤกษ์ ซึ่งจะประกอบไปด้วย 4 สาขาวิชา คือ สาขาวิชาคอมพิวเตอร์ธุรกิจ, สาขาวิชาการจัดการ, สาขาวิชาการตลาด และสาขาวิชาการจัดการโรงแรมและการท่องเที่ยว โดยแบ่งข้อมูลออกเป็น 2 ส่วน คือ ส่วนแรก เป็นฐานข้อมูลการลงทะเบียนเรียนและผลการเรียนในแต่ละวิชา และส่วนที่ 2 เป็นฐานข้อมูลส่วนตัวของนักศึกษา เช่น อายุ เพศ ประวัติการศึกษา เกรดเฉลี่ยสะสม เป็นต้น

3.1.3 ศึกษาขั้นตอนการทำค้ำไมน์นิ่ง และเทคนิคของค้ำไมน์นิ่งในแบบต่างๆ ซึ่งในการสร้างฐานความรู้ เพื่อทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ เลือกเทคนิคของค้ำไมน์นิ่งที่เหมาะสม ดังนี้

1) การจัดกลุ่มข้อมูล (Classification) ซึ่งเลือกใช้เทคนิคของ Decision Tree เพื่อทำการศึกษาพฤติกรรมการณ์เรียนของนักศึกษา (Profile) โดยนำเสนอข้อมูลในรูปแบบของต้นไม้ (Tree) ซึ่งทำให้ง่ายต่อการเข้าใจ ซึ่งจะเลือกใช้อัลกอริทึม CHAID เพราะเป็นอัลกอริทึมที่สามารถรองรับข้อมูลได้ทั้งที่เป็นตัวเลข (Numerical data) และข้อมูลที่เป็นตัวอักษร (Categorical data)

2) การทำนาย/พยากรณ์ข้อมูล (Predictive) ซึ่งเลือกใช้เทคนิคของ Neural Network เพื่อการทำนายผลการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ โดยอาศัยข้อมูลผลการเรียนของนักศึกษาในอดีต (Training data) เพื่อให้โมเดลทำการเรียนรู้พฤติกรรมการณ์เรียนของนักศึกษาแต่ละสาขาวิชาแล้วใช้ข้อมูลทดสอบ (Testing data) ในการทดสอบความถูกต้อง และความน่าเชื่อถือในการทำนายผลการเรียนของแต่ละโมเดล

3.1.4 ศึกษาถึงเครื่องมือ หรือซอฟต์แวร์ที่น่ามาใช้ในการทำค้ำไมน์นิ่ง

ในการดำเนินการวิจัยครั้งนี้ เลือกใช้เครื่องมือ / ซอฟต์แวร์ที่ใช้ในการประมวลผล คือ โปรแกรม Clementine version 12.0 เป็นซอฟต์แวร์ที่ใช้ในการทำค้ำไมน์นิ่ง ที่มีฟังก์ชันการคำนวณทางสถิติจำนวนมาก ที่นำมาใช้ในการวิเคราะห์ข้อมูล โดยมีเครื่องมือในการสร้างโมเดล และมี Algorithm หรือเทคนิคการทำค้ำไมน์นิ่งเป็นจำนวนมากให้ใช้งาน นอกจากนี้โปรแกรม Clementine ยังมีฟังก์ชันในการเตรียมข้อมูลด้วย

## 3.2 รวบรวม คัดเลือก และจัดเตรียมข้อมูล

3.2.1 วัตถุประสงค์ในการเตรียมข้อมูล

1) เพื่อนำไปใช้เป็นข้อมูลในการสร้างโมเดลของการจัดกลุ่มข้อมูล (Classification) และการทำนายข้อมูล (Predictive)

2) เพื่อใช้ในการเปรียบเทียบกับค่าที่ได้จากการทำนายข้อมูล โดยนำค่าที่เปรียบเทียบนั้นไปพัฒนารูปแบบของโมเดลในการทำนายที่สร้างขึ้น ให้มีค่าความคลาดเคลื่อนน้อยที่สุด เพื่อค้นหาโมเดลการทำนายที่เหมาะสมที่สุด

### 3.2.2 ขั้นตอนในการเตรียมข้อมูล (Data Preparation)

การวิจัยในครั้งนี้ เป็นการใช้เทคนิคดาต้าไมน์นิ่งในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ ซึ่งในการเตรียมข้อมูลที่ใช้ในการทำนายสัมฤทธิ์ผลทางการเรียนนี้ มีขั้นตอนต่างๆ ดังนี้

#### 1) การคัดเลือกข้อมูล (Data Selection)

เป็นการคัดเลือกข้อมูลที่ต้องการใช้ในงานวิจัย โดยจะต้องทำการเก็บรวบรวมข้อมูลที่มีผลต่อสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ เพื่อนำไปใช้ในการทำนาย ซึ่งนำมาจากฐานข้อมูลของนักศึกษา คณะบริหารธุรกิจ วิทยาลัยราชพฤกษ์ ตั้งแต่ปี 2549 จนถึงปี 2551 รวมทั้งสิ้น 23,343 เรคคอร์ด จากนักศึกษา 1,154 คน แสดงได้ดังนี้

ตารางที่ 3-1 จำนวนนักศึกษาแยกตามสาขาวิชา

ชื่อสาขาวิชา	จำนวนนักศึกษาตั้งแต่ปี 2549-2551
สาขาคอมพิวเตอร์ธุรกิจ	416
สาขาการตลาด	338
สาขาการจัดการ	324
สาขาการจัดการการโรงแรมและการท่องเที่ยว	76
<b>รวม</b>	<b>1,154</b>

โดยข้อมูลชุดดังกล่าว ประกอบไปด้วยข้อมูล 2 ส่วน ส่วนแรกคือ ข้อมูลส่วนตัวของนักศึกษา โดยจะประกอบไปด้วยข้อมูลต่างๆ ดังนี้

- รหัสนักศึกษา
- ชื่อนักศึกษา
- นามสกุลนักศึกษา
- สาขาวิชาที่เรียน
- วันเดือนปีเกิด
- ภูมิลำเนา (จังหวัดที่แจ้ง)

- วุฒิการศึกษาเดิม
- ที่อยู่ปัจจุบัน

ตารางที่ 3-2 ตัวอย่างข้อมูลส่วนตัวของนักศึกษา

รหัส นักศึกษา	ชื่อ	สกุล	สาขาวิชา	วันเดือนปี เกิด	ภูมิลำเนา	วุฒิ เดิม	ที่อยู่ ปัจจุบัน
491010113	ชลธิชา	รอดชมภู	คอมพิวเตอร์ ธุรกิจ	21-ธ.ค.-88	เลข	03	นนทบุรี
491010099	สุกัญญา	เขียวแก้ว	คอมพิวเตอร์ ธุรกิจ	06-มี.ค.-83	ลพบุรี	03	นนทบุรี
491010135	จิติวุฒิ	แสงสว่าง	การตลาด	02-ม.ค.-85	กาญจนบุรี	03	นนทบุรี
491010103	ชนิษฐา	หมื่นไชยวงศ์	การจัดการ	27-ต.ค.-86	เชียงราย	00	กรุงเทพ มหานคร
.....	.....	.....	.....	.....	.....	....	.....

ส่วนที่สอง คือ ข้อมูลการลงทะเบียน และผลการเรียนของนักศึกษา โดยในการวิจัยครั้งนี้เป็นการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา เพื่อใช้ในการตัดสินใจให้คำแนะนำของอาจารย์ที่ปรึกษาเพื่อเลือกสาขาวิชาที่เหมาะสมให้กับนักศึกษาที่มีผลการเรียนต่ำ ดังนั้น จึงจำเป็นต้องเลือกรายวิชาที่ในแต่ละสาขาวิชา ทั้งสาขาคอมพิวเตอร์ธุรกิจ, สาขาการตลาด, สาขาการจัดการ และสาขาการจัดการโรงแรมและการท่องเที่ยว ทำการลงทะเบียนเรียนเหมือนกัน เพื่อนำไปใช้เป็นข้อมูลพื้นฐานในการทำนายผลการเรียนในส่วนจของรายวิชาเอกบังคับของแต่ละสาขาวิชา เพื่อหาความถนัดทางการเรียนในแต่ละสาขาวิชาให้แก่นักศึกษา โดยข้อมูลการลงทะเบียน และผลการเรียนของนักศึกษาประกอบไปด้วยข้อมูล ดังนี้

- รหัสนักศึกษา
- สาขาวิชา
- ปีการศึกษาที่ลงทะเบียน
- ภาคเรียนที่ลงทะเบียน
- รหัสรายวิชา
- ชื่อรายวิชา
- เกรด

ตารางที่ 3-3 ตัวอย่างข้อมูลการลงทะเบียน และผลการเรียนของนักศึกษา

รหัส นักศึกษา	สาขาวิชา	ปีที่ ลงทะเบียน	ภาคเรียนที่ ลงทะเบียน	รหัสรายวิชา	ชื่อรายวิชา	เกรด
491010006	การตลาด	2549	2	002004	ภาษาอังกฤษ 2	C+
491010006	การตลาด	2549	2	003001	คณิตศาสตร์ทั่วไป	D
491010006	การตลาด	2549	2	004001	ลีลาศ	B+
491010006	การตลาด	2549	2	100003	การบัญชีขั้นต้น2	C
491010006	การตลาด	2549	2	100008	หลักการตลาด	F
491010006	การตลาด	2549	2	100011	องค์การและการจัดการ	D
.....	.....	.....	.....	.....	.....	.....

โดยกลุ่มรายวิชาที่มีผลต่อการทำนายดังกล่าว คือ กลุ่มวิชาแกนในคณะบริหารธุรกิจ ซึ่งเปิดสอนใน 2 หลักสูตร คือ หลักสูตร 2 ปีต่อเนื่อง และหลักสูตร 4 ปี สามารถแสดงโครงสร้างหลักสูตรดังกล่าว ได้ดังนี้

ตารางที่ 3-4 โครงสร้างหลักสูตรของหลักสูตร 2 ปีต่อเนื่อง และหลักสูตร 4 ปี

หลักสูตร 2 ปีต่อเนื่อง	หลักสูตร 4 ปี
จำนวนหน่วยกิตรวมตลอดหลักสูตรไม่น้อยกว่า <b>84</b> หน่วยกิต	จำนวนหน่วยกิตรวมตลอดหลักสูตรไม่น้อยกว่า <b>136</b> หน่วยกิต
<b>องค์ประกอบของหลักสูตร</b>	<b>องค์ประกอบของหลักสูตร</b>
<b>หมวดวิชาศึกษาทั่วไป จำนวน 15 หน่วยกิต</b>	<b>หมวดวิชาศึกษาทั่วไป จำนวน 31 หน่วยกิต</b>
กลุ่มวิชาสังคมศาสตร์และมนุษยศาสตร์ 6 หน่วยกิต	กลุ่มวิชาสังคมศาสตร์และมนุษยศาสตร์ 9 หน่วยกิต
กลุ่มวิชาภาษาศาสตร์ 6 หน่วยกิต	กลุ่มวิชาภาษาศาสตร์ 12 หน่วยกิต
กลุ่มวิชาวิทยาศาสตร์และคณิตศาสตร์ 3 หน่วยกิต	กลุ่มวิชาวิทยาศาสตร์และคณิตศาสตร์ 9 หน่วยกิต
<b>หมวดวิชาเฉพาะ จำนวน 60 หน่วยกิต</b>	กลุ่มวิชาพลศึกษาและนันทนาการ 1 หน่วยกิต
กลุ่มวิชาแกน 21 หน่วยกิต	<b>หมวดวิชาเฉพาะ จำนวน 96 หน่วยกิต</b>
กลุ่มวิชาเอกบังคับ 30 หน่วยกิต	กลุ่มวิชาแกน 54 หน่วยกิต
กลุ่มวิชาเอกเลือก 9 หน่วยกิต	กลุ่มวิชาเอกบังคับ 30 หน่วยกิต
<b>หมวดวิชาเลือกเสรี จำนวน 9 หน่วยกิต</b>	กลุ่มวิชาเอกเลือก 12 หน่วยกิต
	<b>หมวดวิชาเลือกเสรี จำนวน 9 หน่วยกิต</b>

ซึ่งในแต่ละหลักสูตรนั้น มีรายวิชาในกลุ่มวิชาแกน ที่นำมาใช้ในการทำนายสัมฤทธิ์ผลทางการเรียน ดังนี้

ตารางที่ 3-5 รายวิชาแกนของแต่ละหลักสูตร

หลักสูตร 2 ปีต่อเนื่อง	หลักสูตร 4 ปี
1. หลักการบัญชี (100-001)	1. การบัญชีขั้นต้น 1 (100-002)
2. การเงินธุรกิจ (100-004)	2. การบัญชีขั้นต้น 2 (100-003)
3. การภาษีอากร (100-005)	3. การเงินธุรกิจ (100-004)
4. ระบบสารสนเทศเพื่อการจัดการ (100-007)	4. การภาษีอากร (100-005)
5. จริยธรรมธุรกิจ (100-013)	5. การใช้คอมพิวเตอร์ในงานธุรกิจ (100-006)
6. การจัดการการผลิตและการดำเนินงาน (100-014)	6. ระบบสารสนเทศเพื่อการจัดการ (100-007)
7. หลักเศรษฐศาสตร์ (100-016)	7. หลักการตลาด (100-008)
	8. วิจัยทางธุรกิจ (100-009)
	9. หลักการจัดการ (100-010)
	10. องค์การและการจัดการ (100-011)
	11. การจัดการเชิงกลยุทธ์ (100-012)
	12. จริยธรรมธุรกิจ (100-013)
	13. การจัดการการผลิตและการดำเนินงาน (100-014)
	14. การวิเคราะห์เชิงปริมาณ (100-015)
	15. หลักเศรษฐศาสตร์ (100-016)
	16. กฎหมายธุรกิจ (100-017)
	17. ภาษาอังกฤษเพื่อการสื่อสารทางธุรกิจ (100-018)
	18. การใช้ภาษาอังกฤษทางธุรกิจ (100-019)

ในการวิจัยครั้งนี้เป็นการทำนายผลการเรียน โดยนำรายวิชาแกนมาใช้ในการทำนายผลการเรียนของรายวิชาเอกบังคับในแต่ละสาขา เพื่อทำนายความถนัดทางการเรียนให้แก่นักศึกษาที่มีผลการเรียนต่ำ ซึ่งรายวิชาเอกบังคับของแต่ละสาขาที่นำมาใช้เป็นผลการทำนาย มีดังนี้

ตารางที่ 3-6 รายวิชาเอกบังคับของแต่ละสาขาวิชา

สาขาวิชา	รายวิชาเอกบังคับ
สาขาคอมพิวเตอร์ธุรกิจ	<ol style="list-style-type: none"> <li>1. ระบบไมโครคอมพิวเตอร์ (121-001)</li> <li>2. การเขียนโปรแกรมเบื้องต้น (121-002)</li> <li>3. โครงสร้างข้อมูล (121-003)</li> <li>4. ระบบการจัดการฐานข้อมูล (121-004)</li> <li>5. การเขียนโปรแกรมเชิงวัตถุ (121-005)</li> <li>6. การวิเคราะห์และออกแบบระบบ (121-006)</li> <li>7. การสื่อสารข้อมูลและเครือข่ายคอมพิวเตอร์ (121-007)</li> <li>8. เทคโนโลยีมัลติมีเดีย (121-008)</li> <li>9. การเขียนโปรแกรมบนเว็บ (121-009)</li> <li>10. โครงงานคอมพิวเตอร์ธุรกิจ (121-010)</li> </ol>
สาขาการตลาด	<ol style="list-style-type: none"> <li>1. พฤติกรรมผู้บริโภค (131-001)</li> <li>2. นโยบายผลิตภัณฑ์และราคา (131-002)</li> <li>3. การจัดการช่องทางการจัดจำหน่าย (131-003)</li> <li>4. การจัดการการจัดส่ง (131-004)</li> <li>5. การจัดการการตลาด (131-005)</li> <li>6. การจัดการการขาย (131-006)</li> <li>7. การตลาดระหว่างประเทศ (131-007)</li> <li>8. การพยากรณ์ทางการตลาด (131-008)</li> <li>9. การสื่อสารทางการตลาดแบบบูรณาการ (131-009)</li> <li>10. เทคโนโลยีสารสนเทศทางการตลาด (131-010)</li> <li>11. การวิจัยการตลาด (131-011)</li> <li>12. สัมมนากลยุทธ์ทางการตลาด (131-012)</li> </ol>
สาขาการจัดการ	<ol style="list-style-type: none"> <li>1. การจัดการทรัพยากรมนุษย์ (141-001)</li> <li>2. พฤติกรรมองค์การ (141-002)</li> <li>3. ทฤษฎีองค์การ (141-003)</li> <li>4. การเป็นผู้ประกอบการในธุรกิจสมัยใหม่ (141-004)</li> <li>5. การจัดการธุรกิจระหว่างประเทศ (141-005)</li> </ol>

สาขาวิชา	รายวิชาเอกบังคับ
	6. การจัดการส่งออกและนำเข้า (141-006) 7. การสื่อสารเพื่อการจัดการ (141-007) 8. การบริหารโครงการ (141-008) 9. การจัดการสภาพแวดล้อมทางธุรกิจ (141-009) 10. การวางแผนและการควบคุมการบริหาร (141-010) 11. การพยากรณ์ทางธุรกิจ (141-011) 12. สัมมนาการจัดการ (141-012)
สาขาการจัดการการโรงแรมและการท่องเที่ยว	1. อุตสาหกรรมการท่องเที่ยว (151-001) 2. จิตวิทยาการบริการ (151-002) 3. พฤติกรรมนักท่องเที่ยว (151-003) 4. การตลาดสำหรับธุรกิจโรงแรมและการท่องเที่ยว (151-004) 5. หลักการนำเที่ยว (151-005) 6. การพัฒนาการท่องเที่ยวแบบยั่งยืน (151-006) 7. กฎหมายเกี่ยวกับอุตสาหกรรมท่องเที่ยว (151-007) 8. การวิจัยทางการท่องเที่ยว (151-008) 9. การบริการและการจัดการด้านอาหารและเครื่องดื่ม (151-009) 10. การจัดการทรัพยากรบุคคลในธุรกิจโรงแรม (151-010)

## 2) การปรับปรุงข้อมูล (Data Preprocessing)

เมื่อผ่านการคัดเลือกข้อมูล (Data Selection) แล้ว จะเข้าสู่ขั้นตอนการปรับปรุงข้อมูล (Data Preprocessing) ซึ่งเป็นการประมวลผลเบื้องต้น สำหรับการปรับปรุงข้อมูลให้มีคุณภาพดีขึ้น โดยมีวิธีการต่างๆ ดังนี้

### 2.1) การทำความสะอาดข้อมูล (Data Cleaning)

เป็นขั้นตอนในการตรวจสอบข้อมูลให้มีความถูกต้องครบถ้วน และทำการแก้ไขข้อมูลที่ผิดพลาด โดยในการวิจัยครั้งนี้ ได้ทำความสะอาดของข้อมูล ดังนี้

- ทำการตรวจสอบและทำการแก้ไขข้อมูลที่ขาดหาย ซึ่งผลการลงทะเบียนของนักศึกษาในรายวิชาของภาคเรียนที่ 2/2552 ยังไม่ปรากฏผลการเรียน เนื่องจากช่วงเวลาที่นำข้อมูลมาทำการวิจัยนั้น ยังไม่ถึงช่วงสอบปลายภาค ดังนั้นจึงต้องทำการแก้ไข โดยการลบเรคคอร์ดนั้น

- ทำการตรวจสอบข้อมูลที่ไม่สามารถนำมาใช้ในการทำนายผลการเรียนได้ เช่น ผลการเรียนที่ติด W (นักศึกษาถอนรายวิชา) ต้องทำการแก้ไข โดยการลบเรคคอร์ดนั้น

- ทำการตรวจสอบข้อมูลการลงทะเบียน และผลการเรียนของนักศึกษาที่มีเกรดเฉลี่ย 0.00 หรือติด F ทุกวิชา เนื่องจากลงทะเบียนเรียนแล้ว แต่ไม่เคยมาเรียน และไม่มีตัวตนในปัจจุบัน โดยการตรวจสอบจากฝ่ายทะเบียนของทางวิทยาลัยฯ ซึ่งข้อมูลดังกล่าวไม่สามารถนำมาใช้ในการทำนายได้ ต้องทำการแก้ไข โดยการลบเรคคอร์ดนั้น

2.2) การรวมข้อมูล (Data Integration) เนื่องจากการวิจัยครั้งนี้รวบรวมข้อมูลมาจากฐานข้อมูลกลางของนักศึกษาจากวิทยาลัยราชพฤกษ์ ลักษณะและรูปแบบในการจัดเก็บข้อมูลเป็นรูปแบบเดียวกันทั้งหมด จึงไม่ต้องมีการปรับปรุงข้อมูลให้อยู่ในมาตรฐานเดียวกัน

### 3) การแปลงข้อมูล (Data Transformation)

เมื่อผ่านขั้นตอนการปรับปรุงข้อมูล (Data Preprocessing) แล้ว ก่อนที่จะนำข้อมูลเข้าสู่กระบวนการทำค้ำไ่มนึ่งนั้น จะต้องทำการตรวจสอบข้อมูล และแปลงข้อมูลให้อยู่ในรูปแบบที่ตรงตามข้อจำกัดของโมเดลที่ได้เลือกใช้ ซึ่งในงานวิจัยครั้งนี้ ได้เลือกใช้เทคนิคของการจัดกลุ่มข้อมูล (Classification) โดยเลือกใช้โมเดลของ Decision Tree ที่ใช้อัลกอริทึม C4.5 และเทคนิคการทำนายข้อมูล (Predictive) เลือกใช้โมเดลของ Neural Network ซึ่งเป็นโมเดลที่สามารถใช้ข้อมูลที่เป็นตัวเลข (Numerical data) และข้อมูลที่เป็นตัวอักษร (Categorical data) ได้ โดยมีแปลงข้อมูลในส่วนต่างๆ ดังนี้

#### 3.1) ส่วนของกลุ่มรายวิชาแกนที่นำมาใช้ในการจัดกลุ่มและการทำนาย

เนื่องจากทางวิทยาลัยราชพฤกษ์มีหลักสูตรที่เปิดสอนทั้ง 2 ปีต่อเนื่อง และหลักสูตร 4 ปี ซึ่งมีจำนวนหน่วยกิตของกลุ่มวิชาแกนไม่เท่ากันคือ หลักสูตร 2 ปีต่อเนื่องเรียนกลุ่มวิชาแกนทั้งหมด 21 หน่วยกิต แต่หลักสูตร 4 ปีเรียนกลุ่มวิชาแกนทั้งหมด 54 หน่วยกิต การวิจัยครั้งนี้จึงต้องทำการจัดกลุ่มให้แก่รายวิชาในกลุ่มวิชาแกน เพื่อให้รายวิชาแกนที่เลือกเรียนทั้งหลักสูตร 2 ปีต่อเนื่อง และหลักสูตร 4 ปี มีความสอดคล้องกัน โดยได้แบ่งกลุ่มรายวิชาแกนออกเป็น 4 กลุ่ม คือ กลุ่มวิชาแกนทางการเงิน, กลุ่มวิชาแกนทางคอมพิวเตอร์, กลุ่มวิชาแกนทางธุรกิจ และกลุ่มวิชาแกนทางภาษา

ซึ่งในแต่ละหลักสูตรมีจำนวนกลุ่มวิชาแกนไม่เท่ากัน คือ หลักสูตร 4 ปีมีกลุ่มวิชาแกนทางด้านภาษา แต่หลักสูตร 2 ปีต่อเนื่องไม่มีกลุ่มวิชาแกนทางด้านภาษา จึงนำกลุ่มวิชาแกนมาใช้ในการทำนายผลสัมฤทธิ์เพียง 3 กลุ่ม โดยมีรายละเอียด ดังนี้



ตารางที่ 3-7 การจัดกลุ่มรายวิชาแกน

กลุ่มรายวิชาแกน	หลักสูตร 2 ปีต่อเนื่อง	หลักสูตร 4 ปี
วิชาแกนทาง การเงิน	1. หลักการบัญชี 2. ภาษีอากร 3. การเงินธุรกิจ 4. หลักเศรษฐศาสตร์	1. การบัญชีขั้นต้น 1 2. การบัญชีขั้นต้น 2 3. ภาษีอากร 4. การเงินธุรกิจ 5. หลักเศรษฐศาสตร์
วิชาแกนทาง คอมพิวเตอร์	1. ระบบสารสนเทศเพื่อการจัดการ	1. การใช้คอมพิวเตอร์ในงานธุรกิจ 2. ระบบสารสนเทศเพื่อการจัดการ
วิชาแกนทางธุรกิจ	1. จริยธรรมทางธุรกิจ 2. การจัดการการผลิตและการ ดำเนินการ	1. หลักการตลาด 2. วิธีการวิจัยทางธุรกิจ 3. หลักการจัดการ 4. องค์การและการจัดการ 5. การจัดการเชิงกลยุทธ์ 6. จริยธรรมทางธุรกิจ 7. การจัดการการผลิตและการ ดำเนินงาน 8. การวิเคราะห์เชิงปริมาณ

เมื่อทำการจัดกลุ่มรายวิชาแกนแล้ว จะนำผลการเรียนของแต่ละรายวิชา ที่อยู่ในแต่ละกลุ่มวิชาแกนมาทำการเฉลี่ย เพื่อให้ทราบถึงผลการเรียน หรือความถนัดในแต่ละด้านของกลุ่มวิชาแกน และนำผลที่ได้ไปเป็นพื้นฐานในการทำนายผลการเรียนของกลุ่มวิชาเอกบังคับของแต่ละสาขา เพื่อใช้ในการแนะนำสาขาวิชาที่เหมาะสมให้แก่นักศึกษา

3.2) ส่วนของผลการเรียนในแต่ละกลุ่มรายวิชาแกนและกลุ่มรายวิชาเอกบังคับของแต่ละสาขา หลักจากที่มีการกำหนดรายกลุ่มรายวิชาแกนที่นำมาใช้ในการทำนาย และกลุ่มวิชาเอกบังคับที่นำมาใช้เป็นผลการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษาแต่ละสาขาแล้วนั้น จะนำผลการเรียนในรายวิชาของแต่ละกลุ่มรายวิชาแกน และรายวิชาเอกบังคับทั้งหมดของนักศึกษา มาหาค่าเฉลี่ยของผลการเรียนรวม สำหรับแต่ละกลุ่มวิชาแกนและกลุ่มวิชาเอกบังคับของแต่ละสาขา และนำผลการเรียน/เกรดเฉลี่ยนั้น มาใช้เป็นข้อมูลการเรียนรู้ (Training data) และข้อมูลทดสอบ (Test data)

ตารางที่ 3-8 การเฉลี่ยผลการเรียนของกลุ่มวิชาแกนและวิชาเอกบังคับ

กลุ่มวิชาแกนของแต่ละสาขา	รหัสรายวิชา	ผลการเรียนเฉลี่ย
วิชาแกนทางการเงิน (Finance)	100-001 100-005 100-004 100-016 100-002 100-003	เกรดเฉลี่ยของ วิชาแกนการเงิน
วิชาแกนทางคอมพิวเตอร์ (Computer)	100-006 100-007	เกรดเฉลี่ยของ วิชาแกน คอมพิวเตอร์
วิชาแกนทางธุรกิจ (Business)	100-013 100-014 100-008 100-009 100-010 100-011 100-012 100-015	เกรดเฉลี่ยของ วิชาแกนธุรกิจ
สาขาคอมพิวเตอร์ธุรกิจ (Com)	121-001 121-002 121-003 121-004 121-005 121-006 121-007 121-008 121-009 121-010	เกรดเฉลี่ยของ วิชาเอกบังคับ คอมพิวเตอร์ธุรกิจ
สาขาการตลาด (Market)	131-001 131-002 131-003 131-004 131-005 131-006 131-007 131-008 131-009 131-010 131-011 131-012	เกรดเฉลี่ยของ วิชาเอกบังคับ การตลาด
สาขาการจัดการ (Management)	141-001 141-002 141-003 141-004 141-005 141-006 141-007 141-008 141-009 141-010 141-011 141-012	เกรดเฉลี่ยของ วิชาเอกบังคับ การจัดการ

กลุ่มวิชาเอกบังคับของแต่ละสาขา	รหัสรายวิชา	ผลการเรียนเฉลี่ย	
สาขาการจัดการการโรงแรมและ การท่องเที่ยว (Hotel)	151-001	151-002	151-003
	151-004	151-005	151-006
	151-007	151-008	151-009
	151-010		
		เกรดเฉลี่ยของ วิชาเอกบังคับ การจัดการการ โรงแรมและ การท่องเที่ยว	

หลังจากที่ทำการเฉลี่ยผลการเรียนเสร็จเรียบร้อยแล้ว จะเป็นการกำหนดกลุ่มให้แก่ผลการเรียนดังกล่าว ให้แต่ละกลุ่มรายวิชา โดยมีเกณฑ์การกำหนดกลุ่ม ดังนี้

- เกรดเฉลี่ย 2.75 – 4.00 ช่วงผลการเรียนเป็น Good (ดี)
- เกรดเฉลี่ย 2.00 – 2.74 ช่วงผลการเรียนเป็น Normal (ปานกลาง)
- เกรดเฉลี่ย 0.00 – 1.99 ช่วงผลการเรียนเป็น Bad (ต่ำ)

3.3) ในส่วนของข้อมูลส่วนตัวของนักศึกษา มีการแปลงสถานที่ของสถานศึกษาเดิม โดยการจัดกลุ่มตามจังหวัดที่อยู่ออกเป็น 3 กลุ่ม คือ

- Bangkok (กรุงเทพมหานคร)
- Upcountry (ต่างจังหวัด)
- NA (ข้อมูลที่ไม่ได้ระบุไว้)

3.4) ในส่วนของข้อมูลส่วนตัวของนักศึกษา มีการแปลงข้อมูลวันเกิดให้เป็นช่วงอายุ โดยทำการจัดกลุ่มช่วงอายุให้แก่ักศึษาออกเป็น 5 ช่วง ดังนี้

- ช่วงอายุต่ำกว่า 20
- ช่วงอายุระหว่าง 20-25
- ช่วงอายุระหว่าง 26-30
- ช่วงอายุระหว่าง 31-35
- ช่วงอายุมากกว่า 35 ขึ้นไป

ตารางที่ 3-9 ตัวอย่างข้อมูลที่ผ่านการเตรียมข้อมูล (Data Preparation)

STUDENT CODE	Name	study	PROGRAM NAME	Rank study group	age group	GPA old group	Rank grade major	Rank grade business	Rank grade computer	Rank grade finance
501010163	สุกัศ ศรีโพธิ์	BKK	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	Normal	Bad	Normal	Good	Normal
510112019	จิราณูช แก้วศิษฐ์	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	Normal	Bad	Normal	Bad	Normal
501010129	ณรงค์ภพ ช้างวงศ์	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	Normal	Normal	Normal	Good	Bad
501010052	วาสนา ดิยสง่า	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	Good	Normal	Normal	Good	Normal
510112051	ศิรินาถ เมืองมัน	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	Bad	Bad	Good	Good	Bad

### 3.3 การสร้างโมเดลในการจัดกลุ่มข้อมูล และการพยากรณ์ข้อมูล

ในการวิจัยครั้งนี้เป็นการใช้เทคนิคค่าต้นไม้ในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ โดยจะแบ่งส่วนเทคนิคที่นำมาใช้ออกเป็น 2 ส่วน คือ ส่วนของการจัดกลุ่มข้อมูล (Classification) โดยใช้ Decision Tree และการทำนาย/การพยากรณ์ข้อมูล (Predictive) โดยใช้ Neural Network สามารถอธิบายได้ดังนี้

#### 3.3.1 การจัดกลุ่มข้อมูล (Classification)

ในส่วนของการสร้างโมเดลการจัดกลุ่มข้อมูล โดยใช้เทคนิค Decision Tree อัลกอริทึม CHAID เพื่อแสดงให้เห็นถึงพฤติกรรมทางการเรียนของนักศึกษาในแต่ละสาขาวิชา โดยจะต้องมีการกำหนดข้อมูลปัจจัยที่นำเข้าสู่โมเดลเพื่อศึกษาพฤติกรรมทางการเรียน ดังตารางที่ 3-7 และ 3-8

ตารางที่ 3-10 คำอธิบายของข้อมูลที่นำมาใช้ในการจัดกลุ่ม (Classification)

ชื่อข้อมูล	คำอธิบาย	ลักษณะของข้อมูล
Study	สถานที่การศึกษาเดิม	Bangkok, Upcountry, NA
Rank study group	วุฒิการศึกษาเดิม	สายสามัญ สายอาชีพ
Age group	ช่วงอายุของนักศึกษา	<=20, 20-25, 26-30, 31-35 และ 35+

ชื่อข้อมูล	คำอธิบาย	ลักษณะของข้อมูล
Rank grade major	กลุ่มผลการเรียนในกลุ่มวิชาเอกบังคับของแต่ละสาขาวิชา	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)
Rank grade business	กลุ่มผลการเรียนในกลุ่มวิชาแกนธุรกิจ	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)
Rank grade computer	กลุ่มผลการเรียนในกลุ่มวิชาแกนทางคอมพิวเตอร์	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)
Rank grade finance	กลุ่มผลการเรียนในกลุ่มวิชาแกนทางการเงิน	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)

โมเดลการจัดกลุ่มข้อมูลในแบบ Decision Tree นี้ แต่ละกิ่งจะแทนลักษณะหรือค่าต่างๆ ของข้อมูลนักศึกษาที่นำเข้าสู่โมเดล ซึ่งโหนดปลายทางจะมี 3 คลาส กำหนดตามช่วงผลการเรียนของผลการเรียนเฉลี่ยในกลุ่มวิชาเอกบังคับที่กล่าวมาแล้วข้างต้น

จากวิธีการข้างต้น จะต้องทำการสร้างโมเดลจัดกลุ่มข้อมูล เพื่อแสดงพฤติกรรมทางการเรียนกับทุกสาขาวิชา ซึ่งจะใช้ข้อมูลทั้งหมดของแต่ละสาขา

### 3.3.2 การทำนาย/พยากรณ์ข้อมูล (Predictive)

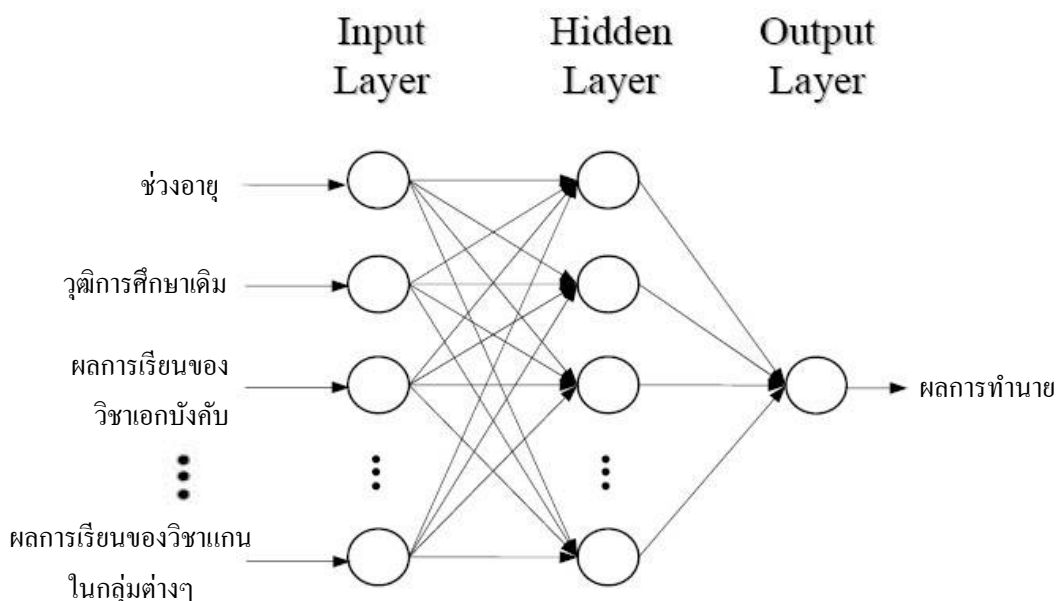
ในส่วนของการสร้างโมเดลในการทำนายผลสัมฤทธิ์ทางการเรียนของนักศึกษา โดยใช้เทคนิค Neural Network นั้น จะต้องมีการกำหนดรูปแบบการเรียนรู้เพื่อให้โมเดลเกิดการเรียนรู้พฤติกรรมการเรียนของนักศึกษาในแต่ละสาขาก่อน ได้แก่ สาขาคอมพิวเตอร์ธุรกิจ, สาขาการตลาด สาขาการจัดการ และสาขาการจัดการโรงแรมและการท่องเที่ยว โดยจะอาศัยข้อมูลการเรียนรู้ (Training data) ที่เป็นข้อมูลส่วนตัว และข้อมูลผลการเรียนของนักศึกษาจากแต่ละสาขาวิชาในอดีต

นอกจากนี้จะต้องมีการกำหนดรูปแบบของการป้อนข้อมูลเข้าสู่โมเดล (Network) เพื่อให้โมเดลเกิดการเรียนรู้ ซึ่งจะต้องมีการกำหนดว่าข้อมูลใดบ้างที่จะนำมาเป็นข้อมูลนำเข้า (Input data) ซึ่งในงานวิจัยครั้งนี้ได้กำหนดข้อมูลปัจจัยที่ใช้เป็นข้อมูลนำเข้า ดังตารางที่ 3-7 และตารางที่ 3-9

ตารางที่ 3-11 คำอธิบายของข้อมูลที่นำมาใช้ในการทำนาย (Predictive)

ชื่อข้อมูล	คำอธิบาย	ลักษณะของข้อมูล
Study	สถานที่การศึกษาเดิม	Bangkok, Upcountry, NA
Rank study group	วุฒิการศึกษาเดิม	สายสามัญ, สายอาชีพ
Age group	ช่วงอายุของนักศึกษา	<=20, 20-25, 26-30, 31-35 และ 35+
Cnt grade	จำนวนหน่วยกิตที่ลงทะเบียน	เป็นตัวเลขนับจำนวนเต็ม
GPA old group	กลุ่มผลการเรียนจากสถาบันการศึกษาเดิม	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)
Rank grade major	กลุ่มผลการเรียนในกลุ่มวิชาเอกบังคับของแต่ละสาขาวิชา	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)
Rank grade business	กลุ่มผลการเรียนในกลุ่มวิชาแกนธุรกิจ	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)
Rank grade computer	กลุ่มผลการเรียนในกลุ่มวิชาแกนทางคอมพิวเตอร์	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)
Rank grade finance	กลุ่มผลการเรียนในกลุ่มวิชาแกนทางการเงิน	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)
Rank grade total	กลุ่มผลการเรียนเฉลี่ยทั้งหมด	Good (ผลการเรียนดี) Normal (ผลการเรียนปานกลาง) Bad (ผลการเรียนต่ำ)

ในงานวิจัยครั้งนี้ได้เลือกใช้เทคนิค Neural Network แบบแพร่กระจายย้อนกลับ (Back-propagation) ซึ่งจะประกอบไปด้วย ชั้นรับข้อมูล (Input layer), ชั้นซ่อนเร้น (Hidden layer) และชั้นแสดงผล (Output layer) ดังนี้



ภาพที่ 3-1 ตัวอย่างโมเดลการทำนายผลการเรียน ด้วยเทคนิคของ Neural Network

และมีการกำหนดส่วนที่เป็นผลการทำนาย (Output data) ของโมเดลว่ามีกี่กลุ่ม นั่นก็คือผลการเรียนเฉลี่ยในรายวิชาเอกบังคับของแต่ละสาขา ซึ่งในงานวิจัยครั้งนี้ได้กำหนดผลการทำนายไว้เป็นกลุ่มๆ ดังนี้

- เกรดเฉลี่ย 2.75 – 4.00 ช่วงผลการเรียนเป็น Good (ดี)
- เกรดเฉลี่ย 2.00 – 2.74 ช่วงผลการเรียนเป็น Normal (ปานกลาง)
- เกรดเฉลี่ย 0.00 – 1.99 ช่วงผลการเรียนเป็น Bad (ต่ำ)

จากการวิจัยครั้งนี้ จะทำการสร้างโมเดลในการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา ออกเป็น 4 โมเดล ดังนี้

ข้อมูลที่ใช้ในการสร้างโมเดลทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา โดยใช้เทคนิค Neural Network จะต้องทำการแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดข้อมูลที่ใช้ในการเรียนรู้ (Training Data) และชุดข้อมูลที่ใช้ในการทดสอบ (Testing Data) โดยแต่ละสาขาวิชา ได้มีการแบ่งออกเป็น 2 ชุด ในสัดส่วนร้อยละ 80:20 ดังนี้

- 1) โมเดลการทำนายผลการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ
  - ชุดข้อมูลที่ใช้ในการเรียนรู้ (Training Data) 80% คิดเป็นจำนวน 333 คน
  - ชุดข้อมูลที่ใช้ในการทดสอบ (Testing Data) 20% คิดเป็นจำนวน 83 คน

- 2) โมเดลการทำนายผลการเรียนของนักศึกษาสาขาวิชาการตลาด
  - ชุดข้อมูลที่ใช้ในการเรียนรู้ (Training Data) 80% คิดเป็นจำนวน 271 คน
  - ชุดข้อมูลที่ใช้ในการทดสอบ (Testing Data) 20% คิดเป็นจำนวน 67 คน
- 3) โมเดลการทำนายผลการเรียนของนักศึกษาสาขาวิชาการจัดการ
  - ชุดข้อมูลที่ใช้ในการเรียนรู้ (Training Data) 80% คิดเป็นจำนวน 260 คน
  - ชุดข้อมูลที่ใช้ในการทดสอบ (Testing Data) 20% คิดเป็นจำนวน 64 คน
- 4) โมเดลการทำนายผลการเรียนของนักศึกษาสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว
  - ชุดข้อมูลที่ใช้ในการเรียนรู้ (Training Data) 80% คิดเป็นจำนวน 61 คน
  - ชุดข้อมูลที่ใช้ในการทดสอบ (Testing Data) 20% คิดเป็นจำนวน 15 คน

### 3.4 การทดสอบความถูกต้อง และความน่าเชื่อถือของโมเดล โดยการใช้ข้อมูลทดสอบ

เป็นขั้นตอนที่นำผลลัพธ์ที่ได้จากการนำข้อมูลเข้าสู่กระบวนการทำนายของโมเดลต่างๆ มาทำการทดสอบความถูกต้อง และความน่าเชื่อถือของโมเดลนั้นๆ โดยในงานวิจัยครั้งนี้ จะนำผลลัพธ์ที่แท้จริงของข้อมูลที่ใช้ทดสอบ (Testing data) มาเปรียบเทียบกับกลุ่มผลลัพธ์ที่หามาได้จากโมเดลการทำนาย เพื่อทดสอบความถูกต้องของแต่ละโมเดล โดยเลือกใช้สูตรหาค่าความถูกต้องของผลที่ได้จากการทำนาย (accuracy) สามารถอธิบายได้ตามหลักการ ดังนี้

	Prediction Good	Prediction Normal	Prediction Bad
Good	TG	FN1	FB1
Normal	FG1	TN	FB2
Bad	FG2	FN2	TB

จากตัวอย่างดังกล่าว สามารถอธิบายได้ดังนี้

- 1) ข้อมูลในแนวตั้ง (คอลัมน์) คือ ผลลัพธ์ที่ได้จากการทำนาย จะขึ้นอยู่กับโมเดลที่ใช้ในการทำนายผลการเรียนของนักศึกษาจากการวิจัย ว่ามีค่าผลลัพธ์ที่แท้จริงอะไรบ้าง จากตัวอย่างดังกล่าว ประกอบไปด้วยค่าของผลลัพธ์ที่ได้จากการทำนาย ดังนี้
  - Prediction Good คือ ผลลัพธ์จากการทำนายเป็นค่าผลการเรียนดี
  - Prediction Normal คือ ผลลัพธ์จากการทำนายเป็นค่าผลการเรียนปานกลาง
  - Prediction Bad คือ ผลลัพธ์จากการทำนายเป็นค่าผลการเรียนต่ำ



2) ข้อมูลในแนวนอน (แถว) คือ ผลลัพธ์ของข้อมูลที่แท้จริง จะขึ้นอยู่กับโมเดลที่ใช้ในการทำนายผลการเรียนของนักศึกษาจากการวิจัย ว่ามีการกำหนดกลุ่มของผลการเรียนไว้อย่างไร ซึ่งจากตัวอย่างดังกล่าวประกอบไปด้วยค่าของผลลัพธ์ ดังนี้

- Good คือ ผลลัพธ์ที่แท้จริงที่มีค่าเป็นผลการเรียนดี
- Normal คือ ผลลัพธ์ที่แท้จริงที่มีค่าเป็นผลการเรียนปานกลาง
- Bad คือ ผลลัพธ์ที่แท้จริงที่มีค่าเป็นผลการเรียนต่ำ

3) ในส่วนของผลลัพธ์ จะขึ้นอยู่กับการจัดกลุ่มของผลลัพธ์ ในการวิจัยครั้งนี้คือ ผลการเรียนของนักศึกษา ซึ่งจากตัวอย่างดังกล่าว สามารถอธิบายได้ดังนี้

*กรณีผลลัพธ์ที่แท้จริงเป็นผลการเรียนดี*

- TG หมายถึง จำนวนผลการทำนายถูก ว่ามีค่าเป็นผลการเรียนดี
- FN1 หมายถึง จำนวนผลการทำนายผิด ว่ามีค่าเป็นผลการเรียนปานกลาง
- FB1 หมายถึง จำนวนผลการทำนายผิด ว่ามีค่าเป็นผลการเรียนต่ำ

*กรณีผลลัพธ์ที่แท้จริงเป็นผลการเรียนปานกลาง*

- FG1 หมายถึง จำนวนผลการทำนายผิด ว่ามีค่าเป็นผลการเรียนดี
- TN หมายถึง จำนวนผลการทำนายถูก ว่ามีค่าเป็นผลการเรียนปานกลาง
- FB2 หมายถึง จำนวนผลการทำนายผิด ว่ามีค่าเป็นผลการเรียนต่ำ

*กรณีผลลัพธ์ที่แท้จริงเป็นผลการเรียนต่ำ*

- FG2 หมายถึง จำนวนผลการทำนายผิด ว่ามีค่าเป็นผลการเรียนดี
- FN2 หมายถึง จำนวนผลการทำนายผิด ว่ามีค่าเป็นผลการเรียนปานกลาง
- TB หมายถึง จำนวนผลการทำนายถูก ว่ามีค่าเป็นผลการเรียนต่ำ

จากคำอธิบายดังกล่าว สามารถนำมาเข้าสู่สูตรคำนวณเพื่อหาค่าความถูกต้อง หรือค่าความเชื่อมั่นของผลการทำนายที่ได้จากโมเดลการทำนายผลการเรียนของแต่ละสาขาวิชา ดังนี้

$\text{Accuracy value} = \frac{\text{TG} + \text{TN} + \text{TB}}{\text{TG} + \text{FN1} + \text{FB1} + \text{FG1} + \text{TN} + \text{FB2} + \text{FG2} + \text{FN2} + \text{TB}}$
---

จากสูตรดังกล่าวข้างต้น อธิบายได้ว่า ค่าความถูกต้องหรือค่าความเชื่อมั่นของโมเดลที่ใช้ในการทำนาย ได้มาจากนำจำนวนผลการทำนายที่ถูกทั้งหมด (ทั้งที่ทำนายถูกว่ามีค่าเป็นผลการเรียนดี, ผล

การเรียนปานกลาง และผลการเรียนต่ำ) หากด้วยจำนวนผลการทำนายทั้งหมด (ทั้งทำนายถูกและทำนายผิด)

ค่าความถูกต้องหรือค่าความเชื่อมั่น (Accuracy value) ที่ได้นั้น บอกถึงความแม่นยำของโมเดลที่ใช้ในการทำนาย ถ้าโมเดลที่ใช้ในการทำนายนั้น มีค่าความถูกต้องสูง หมายถึง โมเดลดังกล่าวสามารถทำนายผลลัพธ์ได้อย่างแม่นยำ

### 3.5 แปลความหมาย และนำความรู้ที่ได้มาใช้

เป็นขั้นตอนสุดท้าย สำหรับนักวิเคราะห์ข้อมูล ที่จะต้องนำผลลัพธ์ที่ได้จากการทำดาต้าไมน์นิ่ง มาสรุปความหมาย โดยผลที่ได้จะกลายเป็นข้อมูลความรู้ (Knowledge) ซึ่งจะนำไปเป็นสารสนเทศ ที่ช่วยในการตัดสินใจต่อไป

โดยการวิจัยในครั้งนี้ ได้นำเทคนิคดาต้าไมน์นิ่งมาใช้ในการสร้างฐานความรู้ เพื่อการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ ซึ่งจะนำผลการทำนายสัมฤทธิ์ผลทางการเรียนที่ได้ มาช่วยในกระบวนการตัดสินใจ ให้คำแนะนำของอาจารย์ที่ปรึกษา โดยการวิเคราะห์ความถนัด และจุดอ่อนทางการเรียนของนักศึกษาแต่ละท่าน เพื่อชี้แนะแนวทางในการเรียน ว่า นักศึกษาแต่ละท่านมีความเหมาะสมกับการเรียนในสาขาวิชาใดมากที่สุด ส่งผลให้กระบวนการพัฒนาคุณภาพการศึกษามีประสิทธิภาพมากยิ่งขึ้น

เมื่ออาจารย์ที่ปรึกษาต้องการข้อมูลเพื่อใช้ในการตัดสินใจให้คำแนะนำทางการเรียนแก่นักศึกษาภายในสาขาวิชาเดิมที่มีผลการเรียนต่ำ เมื่อเรียนในสาขาวิชาอื่นจะมีความถนัดมากขึ้น หรือทำให้มีผลการเรียนดีขึ้นหรือไม่ โดยจะนำข้อมูลส่วนตัวและข้อมูลผลการลงทะเบียนที่ทราบผลการเรียนแล้วของนักศึกษาคณดังกล่าว มาเข้าสู่กระบวนการทำงานของโมเดลการทำนายสัมฤทธิ์ผลทางการเรียนของแต่ละสาขาวิชา ซึ่งโมเดลของแต่ละสาขาวิชาจะมีการเรียนรู้พฤติกรรมทางการเรียนที่แตกต่างกัน โดยการเรียนรู้พฤติกรรมของแต่ละโมเดลในแต่ละสาขาวิชานั้น จะพิจารณาจากข้อมูลส่วนตัว และข้อมูลผลการเรียนของนักศึกษาในอดีตของแต่ละสาขาวิชา ซึ่งสัมฤทธิ์ผลทางการเรียนที่ได้จากการทำนายของนักศึกษาคณดังกล่าว อาจจะมีผลการเรียนที่ดีขึ้น เมื่อเรียนในสาขาวิชาอื่นมากกว่าสาขาวิชาเดิม

## บทที่ 4

### ผลการวิเคราะห์ข้อมูล

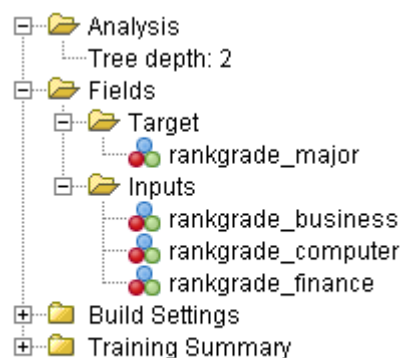
งานวิจัยในครั้งนี้ได้นำเทคนิคดาต้าไมนิงมาประยุกต์ใช้ในการทำนายสัมฤทธิ์ผลทางการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ โดยแบ่งการวิเคราะห์ข้อมูลเป็น 2 ส่วน คือ การจัดกลุ่มข้อมูล (Classification) เพื่อศึกษาพฤติกรรมทางการเรียนของนักศึกษาแต่ละสาขา และการทำนายผลการเรียน (Predictive) ซึ่งจะกล่าวถึงผลการดำเนินการวิเคราะห์ข้อมูลเป็นขั้นตอน ดังนี้

#### 4.1 ผลการจัดกลุ่มข้อมูล (Classification)

ดำเนินการจัดกลุ่มข้อมูล (Classification) โดยเลือกใช้โมเดล Decision Tree เพื่อศึกษาถึงพฤติกรรมทางการเรียนของนักศึกษาในแต่ละสาขาวิชา (Profile) คือ สาขาวิชาคอมพิวเตอร์ธุรกิจ, สาขาวิชาการตลาด, สาขาวิชาการจัดการ และสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว ซึ่งแสดงผลได้ดังนี้

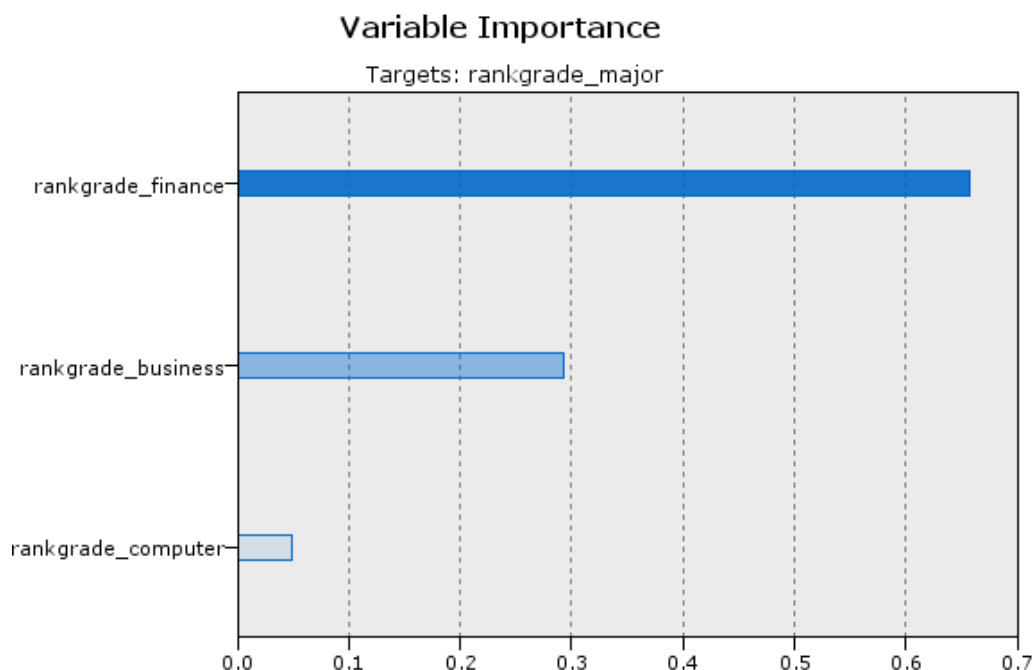
##### 4.1.1 ผลการจัดกลุ่มข้อมูล (Classification) ของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ

มีข้อมูลนักศึกษาจำนวน 416 เรคคอร์ด และมีจำนวนแอททริบิวต์ 7 แอททริบิวต์ ที่นำมาใช้ในการจัดกลุ่มข้อมูล ดังตารางที่ 3-8 ในขั้นตอนวิธีการดำเนินงานวิจัย เมื่อนำข้อมูลเหล่านี้เข้าสู่โมเดลในการจัดกลุ่มแล้ว สามารถแสดงผลได้ดังนี้



ภาพที่ 4-1 รายละเอียดการนำเข้าข้อมูล (สาขาวิชาคอมพิวเตอร์ธุรกิจ)

จากภาพที่ 4-1 มีการกำหนดให้แอททริบิวต์ rankgrade\_major เป็นแอททริบิวต์เป้าหมาย (Target) ซึ่งทางโมเดลมีการเลือกแอททริบิวต์นำเข้า (Input) เป็นจำนวน 3 แอททริบิวต์ ได้แก่ rankgrade\_business, rankgrade\_computer และ rankgrade\_finance โดยทางโมเดลได้มีการกำหนดความสำคัญของแอททริบิวต์ไว้ดังนี้



ภาพที่ 4-2 แสดงความสำคัญของแอททริบิวต์ของนักศึกษาสาขาคอมพิวเตอร์ธุรกิจ

จากภาพที่ 4-2 กล่าวได้ว่า ทางโมเดลได้กำหนดค่าให้แอททริบิวต์ rankgrade\_finance เป็นแอททริบิวต์ที่มีความสำคัญต่อการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจมากที่สุด และรองลงมาคือ rankgrade\_business และ rankgrade\_computer ตามลำดับ

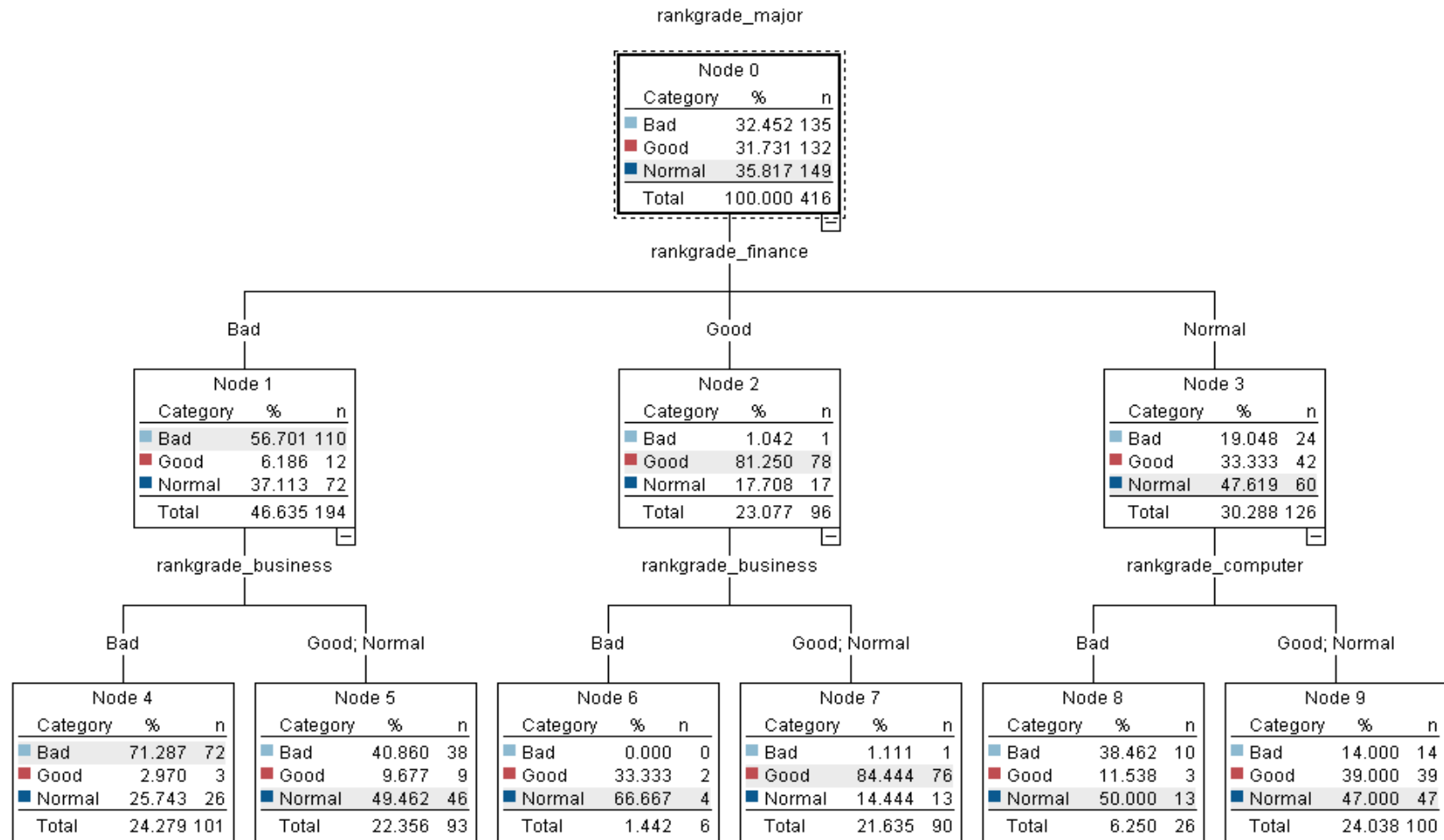
ซึ่งผลของการจัดกลุ่มข้อมูล (Classification) เพื่อใช้ในการศึกษาพฤติกรรมทางการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ ดังภาพที่ 4-3 กล่าวได้ดังนี้

1) นักศึกษาในสาขาคอมพิวเตอร์ธุรกิจ มีผลการเรียนในวิชาเอกบังคับเป็น Bad จำนวน 135 คน มีผลการเรียน Normal จำนวน 149 คน และมีผลการเรียน Good จำนวน 132 คน

2) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Bad มีจำนวน 194 คน โดยจะส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Bad ถึง 110 คน และในกลุ่มเดียวกันนี้ ถ้ามีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Bad จะส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Bad ถึง 72 คน

3) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Good มีจำนวน 96 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good ถึง 78 คน และในกลุ่มเดียวกันนี้ ถ้ามีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Good, Normal จะส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good ถึง 76 คน

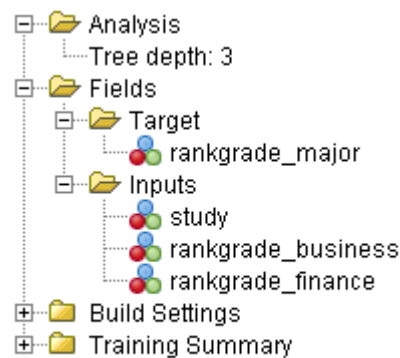
4) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Normal มีจำนวน 126 ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Normal ถึง 60 คน และในกลุ่มเดียวกันนี้ ถ้ามีผลการเรียนในกลุ่มวิชาแกนคอมพิวเตอร์เป็น Good, Normal ก็จะมีผลการเรียนในวิชาเอกบังคับเป็น Normal ถึง 47 คน



ภาพที่ 4-3 การจัดกลุ่มข้อมูล (Classification) ของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ

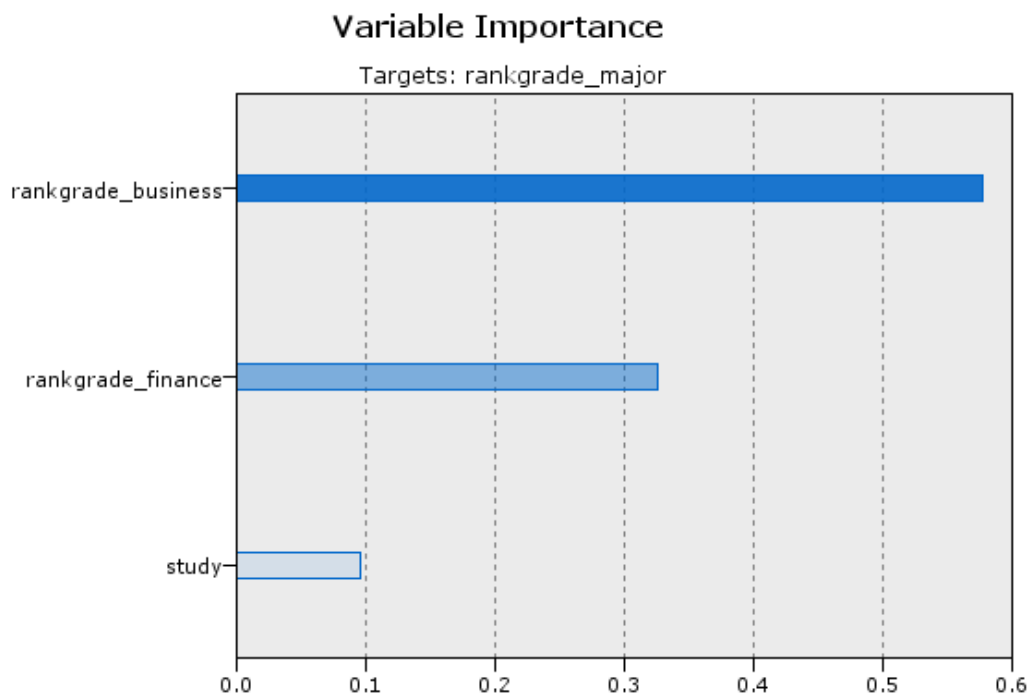
#### 4.1.2 ผลการจัดกลุ่มข้อมูล (Classification) ของนักศึกษาสาขาการตลาด

มีข้อมูลนักศึกษาจำนวน 338 เรคคอร์ด และมีจำนวนแอททริบิวต์ 7 แอททริบิวต์ ที่นำมาใช้ในการจัดกลุ่มข้อมูล ดังตารางที่ 3-8 ในขั้นตอนวิธีการดำเนินงานวิจัย เมื่อนำข้อมูลเหล่านี้เข้าสู่โมเดลในการจัดกลุ่มแล้ว สามารถแสดงผลได้ดังนี้

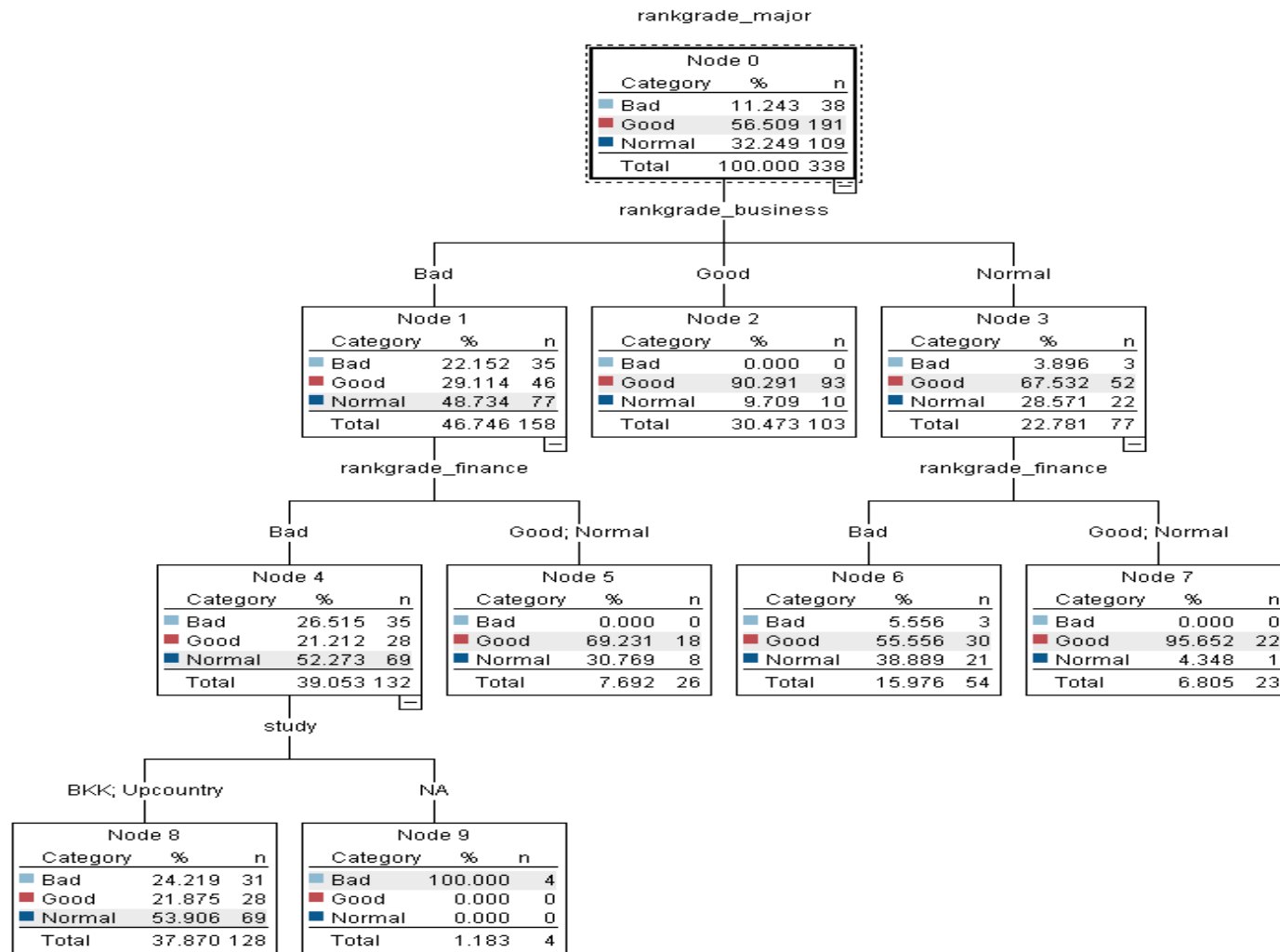


ภาพที่ 4-4 รายละเอียดการนำเข้าข้อมูล (สาขาการตลาด)

จากภาพที่ 4-4 มีการกำหนดให้แอททริบิวต์ `rankgrade_major` เป็นแอททริบิวต์เป้าหมาย (Target) ซึ่งทางโมเดลมีการเลือกแอททริบิวต์นำเข้า (Input) เป็นจำนวน 3 แอททริบิวต์ ได้แก่ `study`, `rankgrade_business` และ `rankgrade_finance` ซึ่งโมเดลได้มีการกำหนดความสำคัญของแอททริบิวต์ไว้ดังนี้



ภาพที่ 4-5 แสดงความสำคัญของแอททริบิวต์ของนักศึกษาสาขาการตลาด



ภาพที่ 4-6 การจัดกลุ่มข้อมูล (Classification) ของนักศึกษาสาขาวิชาการตลาด

จากภาพที่ 4-5 กล่าวได้ว่า ทางโมเดลได้กำหนดให้แอททริบิวต์ rankgrade\_business เป็นแอททริบิวต์ที่มีความสำคัญต่อการเรียนของนักศึกษาสาขาวิชาการตลาดมากที่สุด และรองลงมาคือ rankgrade\_finance และ study ตามลำดับ

ซึ่งผลของการจัดกลุ่มข้อมูล (Classification) เพื่อใช้ในการศึกษาพฤติกรรมทางการเรียนของนักศึกษาสาขาวิชาการตลาด ดังภาพที่ 4-6 กล่าวได้ดังนี้

1) นักศึกษาในสาขาการตลาด มีผลการเรียนในกลุ่มวิชาเอกบังคับเป็น Bad จำนวน 38 คน มีผลการเรียน Normal จำนวน 109 คน และมีผลการเรียน Good จำนวน 191 คน

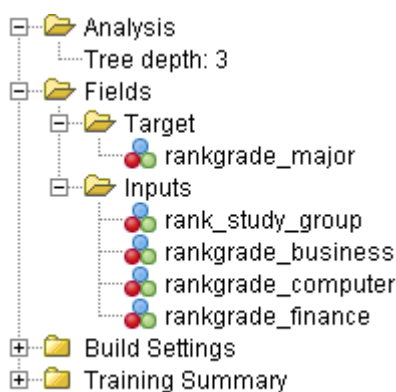
2) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Bad มีจำนวน 158 คน โดยจะส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Normal ถึง 77 คน และในกลุ่มเดียวกันนี้ มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Bad จะส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Normal ถึง 69 คน

3) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Good มีจำนวน 103 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good ถึง 93 คน

4) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Normal มีจำนวน 77 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good ถึง 52 คน และในกลุ่มเดียวกันนี้ ถ้ามีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Bad จะมีผลการเรียนในวิชาเอกบังคับเป็น Good ถึง 30 คน

#### 4.1.3 ผลการจัดกลุ่มข้อมูล (Classification) ของนักศึกษาสาขาวิชาการจัดการ

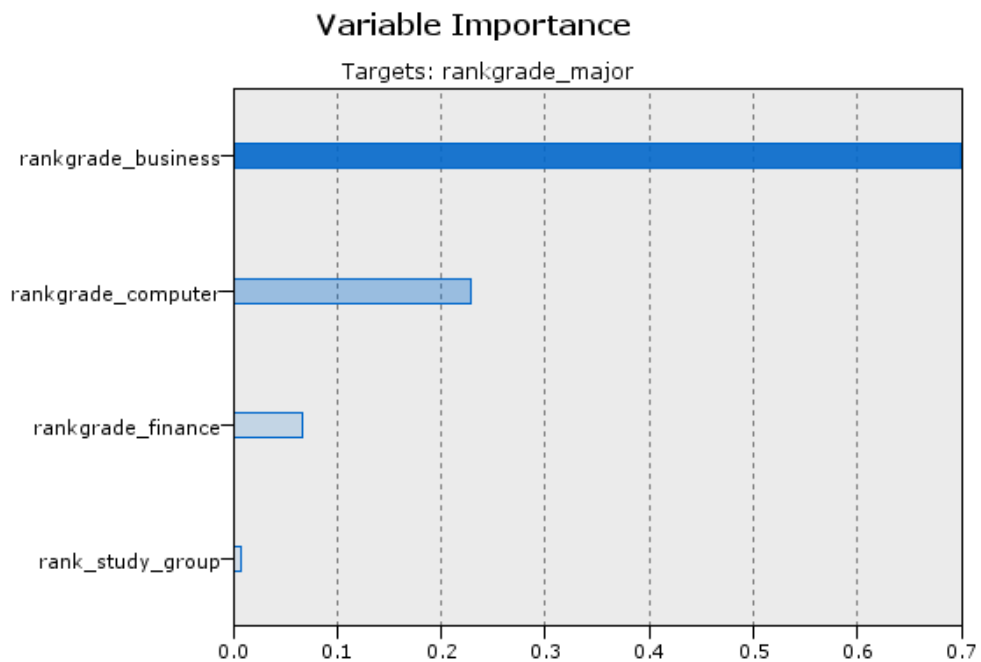
มีข้อมูลนักศึกษาจำนวน 324 เรคคอร์ด และมีจำนวนแอททริบิวต์ 7 แอททริบิวต์ ที่นำมาใช้ในการจัดกลุ่มข้อมูล ดังตารางที่ 3-8 ในขั้นตอนวิธีการดำเนินงานวิจัย เมื่อนำข้อมูลเหล่านี้เข้าสู่โมเดลในการจัดกลุ่มแล้ว สามารถแสดงผลได้ดังนี้



ภาพที่ 4-7 รายละเอียดการนำเข้าข้อมูล (สาขาวิชาการจัดการ)

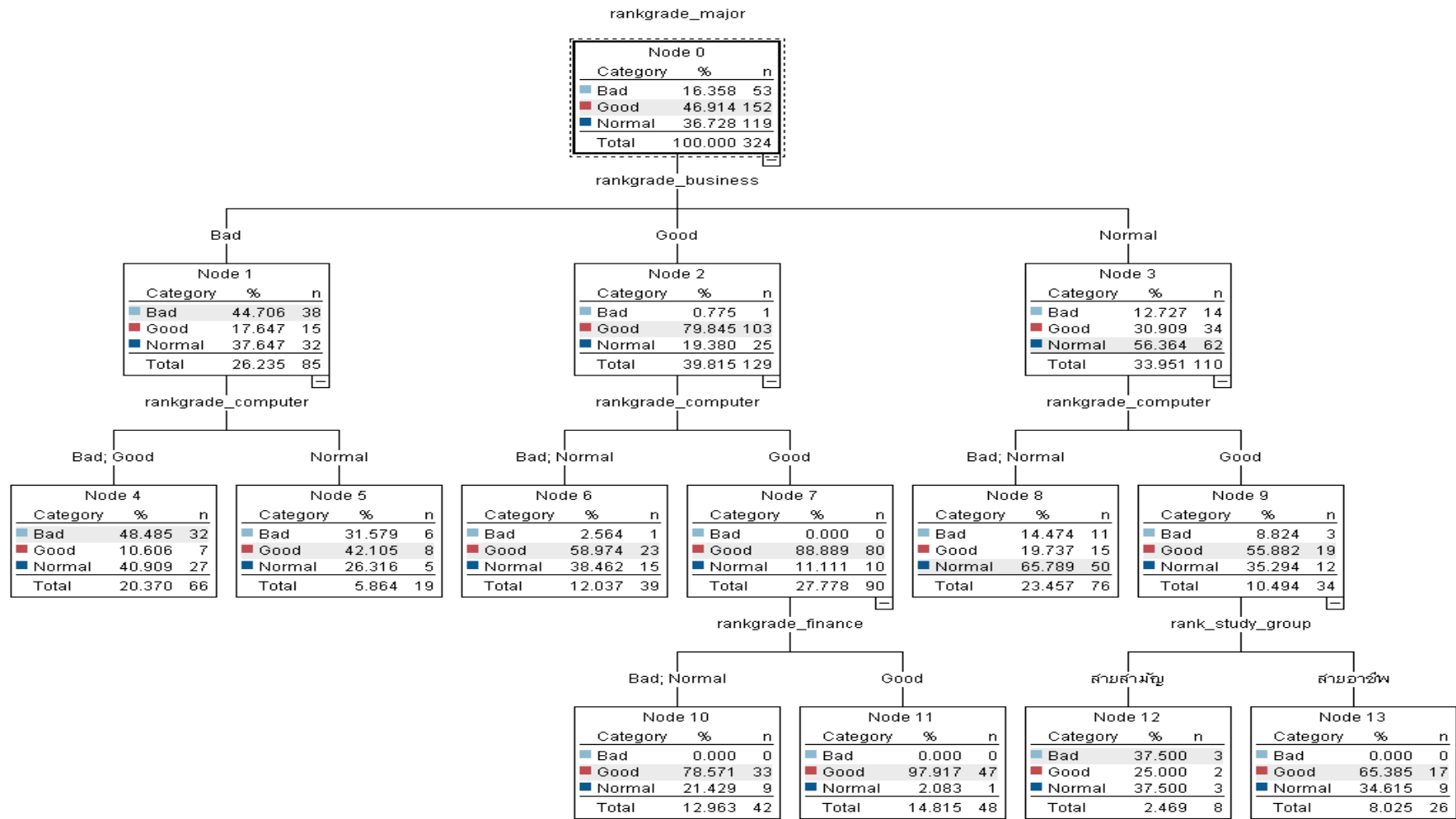


จากภาพที่ 4-7 มีการกำหนดให้แอททริบิวต์ rankgrade\_major เป็นแอททริบิวต์เป้าหมาย (Target) ซึ่งทางโมเดลมีการเลือกแอททริบิวต์นำเข้า (Input) เป็นจำนวน 4 แอททริบิวต์ ได้แก่ rank\_study\_group, rankgrade\_business, rankgrade\_computer และ rankgrade\_finance โดยทางโมเดลได้มีการกำหนดความสำคัญของแอททริบิวต์ ไว้ดังนี้



**ภาพที่ 4-8** แสดงความสำคัญของแอททริบิวต์ของนักศึกษาสาขาการจัดการ

จากภาพที่ 4-8 กล่าวได้ว่า ทางโมเดลได้กำหนดให้แอททริบิวต์ rankgrade\_business เป็นแอททริบิวต์ที่มีความสำคัญต่อการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจมากที่สุด และรองลงมาคือ rankgrade\_computer, rankgrade\_finance และ rank\_study\_group ตามลำดับ



ภาพที่ 4-9 การจัดกลุ่มข้อมูล (Classification) ของนักศึกษาสาขาวิชาการจัดการ

ซึ่งผลของการจัดกลุ่มข้อมูล (Classification) เพื่อใช้ในการศึกษาพฤติกรรมทางการเรียนของนักศึกษาสาขาวิชาการจัดการ ดังภาพที่ 4-9 กล่าวได้ดังนี้

1) นักศึกษาในสาขาการจัดการ มีผลการเรียนในกลุ่มวิชาเอกบังคับ Bad จำนวน 53 คน มีผลการเรียน Normal จำนวน 119 คน และมีผลการเรียน Good จำนวน 152 คน

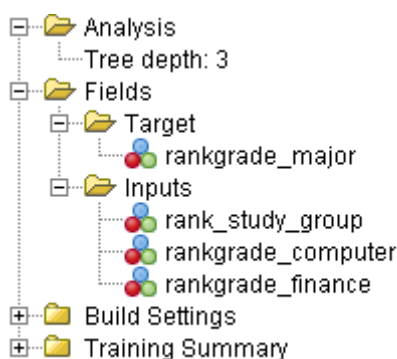
2) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Bad มีจำนวน 85 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Bad ถึง 38 คน และในกลุ่มเดียวกัน ถ้ามีผลการเรียนในกลุ่มวิชาแกนคอมพิวเตอร์เป็น Bad, Good ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Bad จำนวน 32 คน

3) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Good มีจำนวน 129 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good ถึง 103 คน และในกลุ่มเดียวกัน ถ้ามีผลการเรียนในกลุ่มวิชาแกนคอมพิวเตอร์เป็น Good ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good จำนวน 80 คน

4) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Normal มีจำนวน 110 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Normal ถึง 62 คน และในกลุ่มเดียวกัน ถ้ามีผลการเรียนในกลุ่มวิชาแกนคอมพิวเตอร์เป็น Bad, Normal ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Normal ถึง 50 คน แต่ถ้ามีผลการเรียนในกลุ่มวิชาแกนคอมพิวเตอร์เป็น Good ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good จำนวน 19 คน และในกลุ่มเดียวกันนี้ ส่วนใหญ่เป็นกลุ่มนักศึกษาที่จบการศึกษาทางด้านสายอาชีพ จะมีผลการเรียนในวิชาเอกบังคับเป็น Good ถึง 17 คน

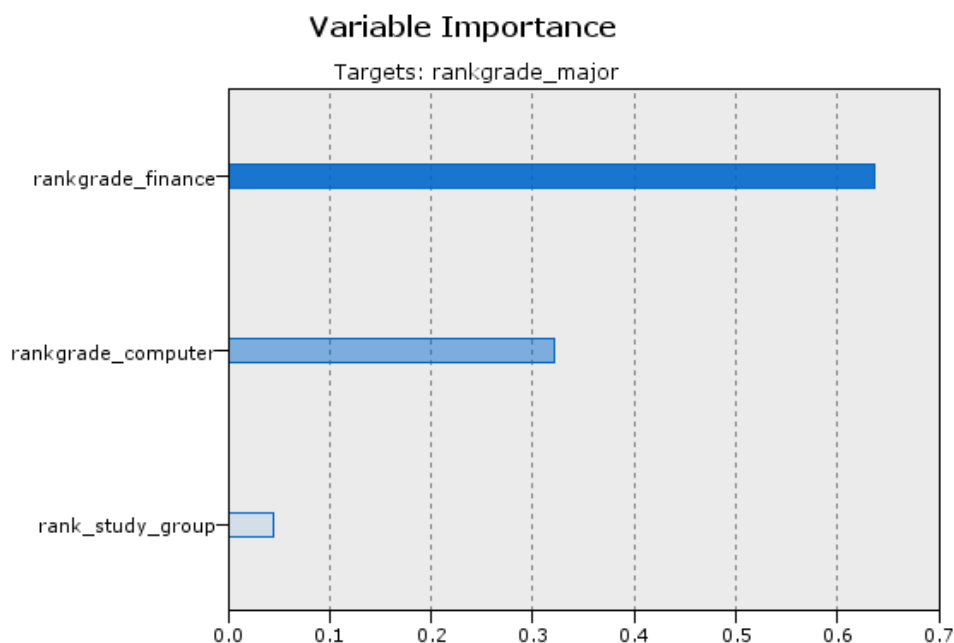
4.1.4 ผลการจัดกลุ่มข้อมูล (Classification) ของนักศึกษาสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว

มีข้อมูลนักศึกษาจำนวน 76 เรคคอร์ด และมีจำนวนแอททริบิวต์ 7 แอททริบิวต์ ที่นำมาใช้ในการจัดกลุ่มข้อมูล ดังตารางที่ 3-8 ในขั้นตอนวิธีการดำเนินงานวิจัย เมื่อนำข้อมูลเหล่านี้เข้าสู่โมเดลในการจัดกลุ่มแล้ว สามารถแสดงผลได้ดังนี้



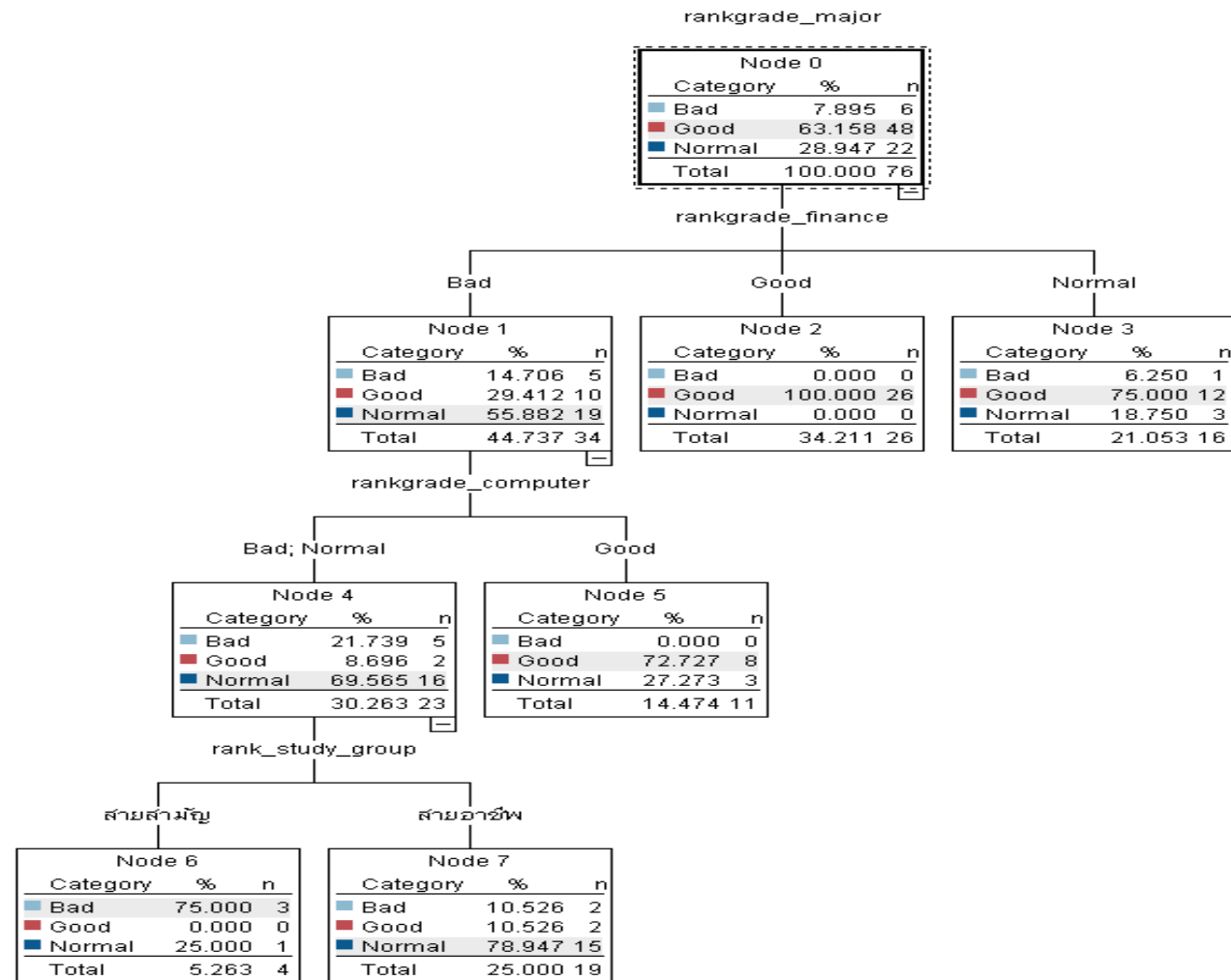
ภาพที่ 4-10 รายละเอียดการนำเข้าข้อมูล (สาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว)

จากภาพที่ 4-10 มีการกำหนดให้แอททริบิวต์ rankgrade\_major เป็นแอททริบิวต์เป้าหมาย (Target) ซึ่งทางโมเดลมีการเลือกแอททริบิวต์นำเข้า (Input) เป็นจำนวน 3 แอททริบิวต์ ได้แก่ rank\_study\_group, rankgrade\_computer และ rankgrade\_finance โดยทางโมเดลได้มีการกำหนดความสำคัญของแอททริบิวต์ไว้ดังนี้



**ภาพที่ 4-11** แสดงความสำคัญของแอททริบิวต์ของนักศึกษาสาขาการจัดการการโรงแรมฯ

จากภาพที่ 4-11 กล่าวได้ว่า ทางโมเดลได้กำหนดให้แอททริบิวต์ rankgrade\_finance เป็นแอททริบิวต์ที่มีความสำคัญต่อการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจมากที่สุด และรองลงมาคือ rankgrade\_computer และ rank\_study\_group ตามลำดับ



ภาพที่ 4-12 การจัดกลุ่มข้อมูล (Classification) ของนักศึกษาสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว

ซึ่งผลของการจัดกลุ่มข้อมูล (Classification) เพื่อใช้ในการศึกษาพฤติกรรมทางการเรียนของนักศึกษาสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว ดังภาพที่ 4-12 กล่าวได้ดังนี้

1) นักศึกษาในสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว มีผลการเรียนในกลุ่มวิชาเอกบังคับเป็น Bad จำนวน 6 คน มีผลการเรียน Normal จำนวน 22 คน และมีผลการเรียน Good จำนวน 48 คน

2) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Bad มีจำนวน 34 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Normal ถึง 19 คน และในกลุ่มเดียวกัน ถ้ามีผลการเรียนในกลุ่มวิชาแกนคอมพิวเตอร์เป็น Bad, Normal ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Normal ถึง 16 คน โดยส่วนใหญ่จะเป็นกลุ่มนักศึกษาที่จบการศึกษาทางด้านสายอาชีพ

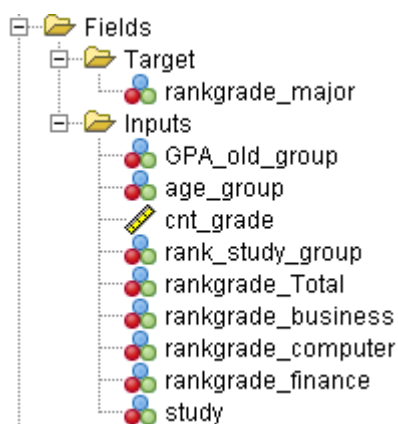
3) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Good มีจำนวน 16 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good ทั้งหมด 26 คน

4) นักศึกษาที่มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Normal มีจำนวน 26 คน ส่งผลให้มีผลการเรียนในวิชาเอกบังคับเป็น Good ถึง 12 คน

#### 4.2 ผลการทำนาย/พยากรณ์ผลการเรียน (Predictive)

ในงานวิจัยครั้งนี้ ได้เลือกเทคนิคการทำนาย/พยากรณ์ (Predictive) โดยใช้ Neural Network ในการสร้างโมเดลการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา จำนวน 4 โมเดล คือ สาขาวิชาคอมพิวเตอร์ธุรกิจ, สาขาวิชาการตลาด, สาขาวิชาการจัดการ และสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว

โดยอาศัยข้อมูลผลการเรียนและข้อมูลส่วนตัวของนักศึกษาในอดีต ดังตารางที่ 3-9 ซึ่งทำการกำหนดแอททริบิวต์นำเข้าเป็นจำนวน 9 แอททริบิวต์ และแอททริบิวต์เป้าหมาย (Target) จำนวน 1 แอททริบิวต์ ดังนี้



ภาพที่ 4-13 รายละเอียดการนำเข้าข้อมูลของแต่ละสาขาวิชา เพื่อสร้างโมเดลการทำนายผลการเรียน

จากภาพที่ 4-13 สามารถอธิบายได้ว่า การสร้างโมเดลการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา ได้ทำการเลือกแอททริบิวต์เป้าหมาย (Target) เป็น rankgrade\_major (ผลการเรียนเฉลี่ยในกลุ่มวิชาเอกบังคับของแต่ละสาขาวิชา) จัดว่าเป็นแอททริบิวต์ที่แสดงถึงความถนัดทางการเรียนของนักศึกษาในแต่ละสาขานั้นๆ ดังนั้น ในส่วนของแอททริบิวต์ที่เหลือ จะถูกกำหนดให้เป็นแอททริบิวต์นำเข้า (Input)

ซึ่งผลการวิเคราะห์ข้อมูล เพื่อสร้างโมเดลที่ใช้ในการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา (Predictive) มีดังนี้

#### 4.2.1 ผลการสร้างโมเดลในการทำนายผลการเรียนของสาขาวิชาคอมพิวเตอร์ธุรกิจ

สามารถแบ่งผลการวิเคราะห์ข้อมูลออกเป็นจำนวน 2 ครั้ง คือ

ครั้งที่ 1 เป็นการวิเคราะห์ข้อมูล เพื่อให้ Neurons เกิดการเรียนรู้พฤติกรรมทางการเรียนของสาขาวิชาคอมพิวเตอร์ธุรกิจ โดยใช้ข้อมูลเรียนรู้ (Training Data) แสดงได้ดังนี้

Results for output field rankgrade\_major

Comparing \$N-rankgrade\_major with rankgrade\_major

Correct	248	74.47%
Wrong	85	25.53%
Total	333	

Coincidence Matrix for \$N-rankgrade\_major (rows show actuals)

	Bad	Good	Normal
Bad	62	1	43
Good	1	91	13
Normal	9	18	95

Performance Evaluation

Confidence Values Report for \$NC-rankgrade\_major

ภาพที่ 4-14 ผลการวิเคราะห์ข้อมูลการทำนายด้วยข้อมูลเรียนรู้ สาขาวิชาคอมพิวเตอร์ธุรกิจ

จากภาพที่ 4-14 กล่าวได้ว่า ผลการวิเคราะห์ข้อมูลการทำนายผลการเรียนด้วยข้อมูลเรียนรู้ (Training Data) จำนวน 333 เรคคอร์ด มีความสามารถในการเรียนรู้ได้ถูกต้อง เป็นจำนวน 248 เรคคอร์ด คิดเป็น 74.47% และผลการเรียนรู้ไม่ถูกต้อง จำนวน 85 เรคคอร์ด คิดเป็น 25.53%

เมื่อแจกแจงรายละเอียดของผลการวิเคราะห์ข้อมูลเรียนรู้ (Training Data) จะได้ว่า

- ผลการเรียนจริงที่เป็น Bad มีจำนวน 106 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Bad จำนวน 62 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Good จำนวน 1 เรคคอร์ด เป็น Normal จำนวน 43 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Good มีจำนวน 105 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Good จำนวน 91 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Bad จำนวน 1 เรคคอร์ด เป็น Normal จำนวน 13 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Normal มีจำนวน 122 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Normal จำนวน 95 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Bad จำนวน 9 เรคคอร์ด เป็น Good จำนวน 18 เรคคอร์ด

ครั้งที่ 2 เป็นการวิเคราะห์ข้อมูล เพื่อทำการทดสอบความถูกต้องของโมเดลที่ใช้ในการทำนายผลการเรียน โดยใช้ข้อมูลการทดสอบ (Testing Data) แสดงได้ดังนี้

Results for output field rankgrade\_major

Comparing \$N-rankgrade\_major with rankgrade\_major

Correct	61	73.49%
Wrong	22	26.51%
Total	83	

Coincidence Matrix for \$N-rankgrade\_major (rows show actuals)

	Bad	Good	Normal
Bad	16	0	13
Good	0	21	6
Normal	1	2	24

Performance Evaluation

Confidence Values Report for \$NC-rankgrade\_major

ภาพที่ 4-15 ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ สาขาวิชาคอมพิวเตอร์ธุรกิจ

จากภาพที่ 4-15 กล่าวว่า ผลการทำนายผลการเรียน ด้วยข้อมูลทดสอบ (Testing Data) จำนวน 83 เรคคอร์ด มีความสามารถในการทำนายได้ถูกต้อง เป็นจำนวน 61 เรคคอร์ด คิดเป็น 73.49% และผลการทำนายไม่ถูกต้อง จำนวน 22 เรคคอร์ด คิดเป็น 26.51%

เมื่อแจกแจงรายละเอียดของผลการทำนายด้วยข้อมูลทดสอบ (Testing Data) จะได้ว่า

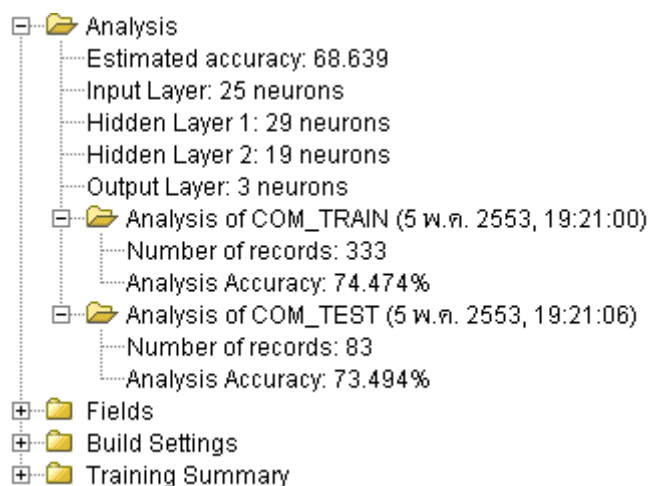
- ผลการเรียนจริงที่เป็น Bad มีจำนวน 29 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Bad จำนวน 16 เรคคอร์ด และผลการทำนายผิดเป็น Normal จำนวน 13 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Good มีจำนวน 27 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Good จำนวน 21 เรคคอร์ด และผลการทำนายผิดเป็น Normal จำนวน 6 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Normal มีจำนวน 27 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Normal จำนวน 24 เรคคอร์ด และผลการทำนายผิดเป็น Bad จำนวน 1 เรคคอร์ด และผลการทำนายผิดเป็น Good จำนวน 2 เรคคอร์ด



ซึ่งจากการวิเคราะห์ข้างต้น สามารถสรุปเป็นผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ ด้วย Neural Network ได้ดังนี้



ภาพที่ 4-16 สรุปผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียน สาขาวิชาคอมพิวเตอร์ธุรกิจ

จากภาพที่ 4-16 สรุปว่าผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ ด้วย Neural Network มี Input Layer จำนวน 25 Neurons มี Hidden Layer จำนวน 2 Layer คือ Hidden Layer 1 จำนวน 29 Neurons และ Hidden Layer 2 จำนวน 19 Neurons และมี Output Layer จำนวน 3 Neurons ซึ่งโมเดลดังกล่าว มีความถูกต้องในการทำนาย คิดเป็น 73.494%

#### 4.2.2 ผลการสร้างโมเดลในการทำนายผลการเรียนของสาขาวิชาการตลาด

สามารถแบ่งผลการวิเคราะห์ข้อมูลออกเป็นจำนวน 2 ครั้ง คือ

ครั้งที่ 1 เป็นการวิเคราะห์ข้อมูล เพื่อให้ Neurons เกิดการเรียนรู้พฤติกรรมทางการเรียนของสาขาวิชาการตลาด โดยใช้ข้อมูลเรียนรู้ (Training Data) ดังนี้

ผลการวิเคราะห์ข้อมูลการทำนายผลการเรียนด้วยข้อมูลเรียนรู้ (Training Data) จำนวน 271 เรคคอร์ด มีความสามารถในการเรียนรู้ได้ถูกต้อง เป็นจำนวน 219 เรคคอร์ด คิดเป็น 80.81% และผลการเรียนรู้ไม่ถูกต้อง จำนวน 52 เรคคอร์ด คิดเป็น 19.19% ดังภาพที่ 4-17

Results for output field rankgrade\_major

Comparing \$N-rankgrade\_major with rankgrade\_major

Correct	219	80.81%
Wrong	52	19.19%
Total	271	

Coincidence Matrix for \$N-rankgrade\_major (rows show actuals)

	Bad	Good	Normal
Bad	21	1	7
Good	0	142	12
Normal	11	21	56

Performance Evaluation

Confidence Values Report for \$NC-rankgrade\_major

ภาพที่ 4-17 ผลการวิเคราะห์ข้อมูลการทำนายด้วยข้อมูลเรียนรู้ สาขาวิชาการตลาด

เมื่อแจกแจงรายละเอียดของผลการวิเคราะห์ข้อมูลเรียนรู้ (Training Data) จะได้ว่า

- ผลการเรียนจริงที่เป็น Bad มีจำนวน 29 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Bad จำนวน 21 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Good จำนวน 1 เรคคอร์ด เป็น Normal จำนวน 7 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Good มีจำนวน 154 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Good จำนวน 142 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Normal จำนวน 12 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Normal มีจำนวน 88 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Normal จำนวน 56 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Bad จำนวน 11 เรคคอร์ด เป็น Good จำนวน 21 เรคคอร์ด

ครั้งที่ 2 เป็นการวิเคราะห์ข้อมูล เพื่อทำการทดสอบความถูกต้องของโมเดลที่ใช้ในการทำนายผลการเรียน โดยใช้ข้อมูลทดสอบ (Testing Data) แสดงได้ดังนี้

Results for output field rankgrade\_major

Comparing \$N-rankgrade\_major with rankgrade\_major

Correct	56	83.58%
Wrong	11	16.42%
Total	67	

Coincidence Matrix for \$N-rankgrade\_major (rows show actuals)

	Bad	Good	Normal
Bad	7	0	2
Good	0	33	4
Normal	2	3	16

Performance Evaluation

Confidence Values Report for \$NC-rankgrade\_major

ภาพที่ 4-18 ผลการทำนายผลการเรียน ด้วยข้อมูลทดสอบ สาขาวิชาการตลาด

จากภาพที่ 4-18 กล่าวว่า ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ (Testing Data) จำนวน 67 เรคคอร์ด มีความสามารถในการทำนายได้ถูกต้อง เป็นจำนวน 56 เรคคอร์ด คิดเป็น 83.58% และผลการทำนายไม่ถูกต้อง จำนวน 11 เรคคอร์ด คิดเป็น 16.42%

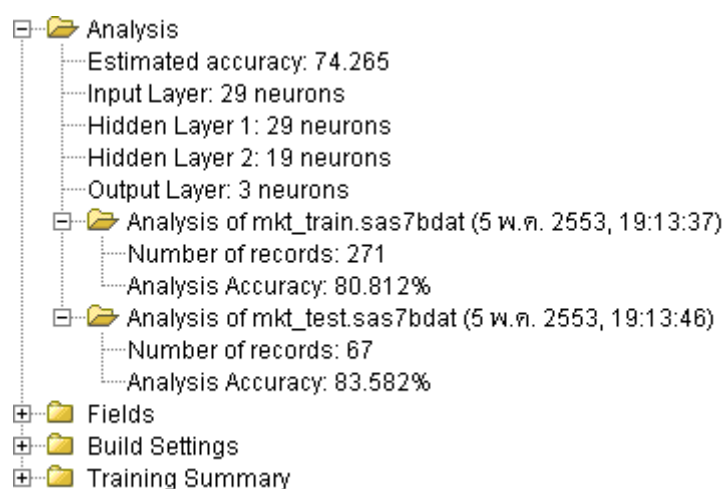
เมื่อแจกแจงรายละเอียดของผลการทำนายด้วยข้อมูลทดสอบ (Testing Data) จะได้ว่า

- ผลการเรียนจริงที่เป็น Bad มีจำนวน 9 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้อง เป็น Bad จำนวน 7 เรคคอร์ด และผลการทำนายผิดเป็น Normal จำนวน 2 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Good มีจำนวน 37 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Good จำนวน 33 เรคคอร์ด และผลการทำนายผิดเป็น Normal จำนวน 4 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Normal มีจำนวน 21 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Normal จำนวน 16 เรคคอร์ด และผลการทำนายผิดเป็น Bad จำนวน 2 เรคคอร์ด และผลการทำนายผิดเป็น Good จำนวน 3 เรคคอร์ด

ซึ่งจากการวิเคราะห์ข้างต้น สามารถสรุปเป็นผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาการตลาด ด้วย Neural Network ได้ดังนี้



ภาพที่ 4-19 สรุปผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียน สาขาวิชาการตลาด

จากภาพที่ 4-19 สรุปว่าผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาการตลาด ด้วย Neural Network มี Input Layer จำนวน 29 Neurons มี Hidden Layer จำนวน 2 Layer คือ Hidden Layer 1 จำนวน 29 Neurons และ Hidden Layer 2 จำนวน 19 Neurons และมี Output Layer จำนวน 3 Neurons ซึ่งโมเดลดังกล่าว มีความถูกต้องในการทำนาย คิดเป็น 83.582%

#### 4.2.3 ผลการสร้างโมเดลในการทำนายผลการเรียนของสาขาวิชาการจัดการ

สามารถแบ่งผลการวิเคราะห์ข้อมูลออกเป็นจำนวน 2 ครั้ง คือ

ครั้งที่ 1 เป็นการวิเคราะห์ข้อมูล เพื่อให้ Neurons เกิดการเรียนรู้พฤติกรรมทางการเรียนของสาขาวิชาการจัดการ โดยใช้ข้อมูลเรียนรู้ (Training Data) แสดงได้ดังนี้

Results for output field rankgrade\_major

Comparing \$N-rankgrade\_major with rankgrade\_major

Correct	216	83.08%
Wrong	44	16.92%
Total	260	

Coincidence Matrix for \$N-rankgrade\_major (rows show actuals)

	Bad	Good	Normal
Bad	29	0	12
Good	0	105	18
Normal	3	11	82

Performance Evaluation

Confidence Values Report for \$NC-rankgrade\_major

ภาพที่ 4-20 ผลการวิเคราะห์ข้อมูลการทำนายด้วยข้อมูลเรียนรู้ สาขาวิชาการจัดการ

จากภาพที่ 4-20 กล่าวว่า ผลการวิเคราะห์ข้อมูลการทำนายผลการเรียนด้วยข้อมูลเรียนรู้ (Training Data) จำนวน 260 เรคคอร์ด มีความสามารถในการเรียนรู้ได้ถูกต้อง เป็นจำนวน 216 เรคคอร์ด คิดเป็น 83.08% และผลการเรียนรู้ไม่ถูกต้อง จำนวน 44 เรคคอร์ด คิดเป็น 16.92%

เมื่อแจกแจงรายละเอียดของผลการวิเคราะห์ข้อมูลเรียนรู้ (Training Data) จะได้ว่า

- ผลการเรียนจริงที่เป็น Bad มีจำนวน 41 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Bad จำนวน 29 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Normal จำนวน 12 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Good มีจำนวน 123 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Good จำนวน 105 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Normal จำนวน 18 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Normal มีจำนวน 96 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Normal จำนวน 82 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิดเป็น Bad จำนวน 3 เรคคอร์ด เป็น Good จำนวน 11 เรคคอร์ด

ครั้งที่ 2 เป็นการวิเคราะห์ข้อมูล เพื่อทำการทดสอบความถูกต้องของโมเดลที่ใช้ในการทำนายผลการเรียน โดยใช้ข้อมูลการทดสอบ (Testing Data) แสดงได้ดังนี้

Results for output field rankgrade\_major

Comparing \$N-rankgrade\_major with rankgrade\_major

Correct	50	78.12%
Wrong	14	21.88%
Total	64	

Coincidence Matrix for \$N-rankgrade\_major (rows show actuals)

	Bad	Good	Normal
Bad	6	0	6
Good	0	23	6
Normal	0	2	21

Performance Evaluation

Confidence Values Report for \$NC-rankgrade\_major

ภาพที่ 4-21 ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ สาขาวิชาการจัดการ

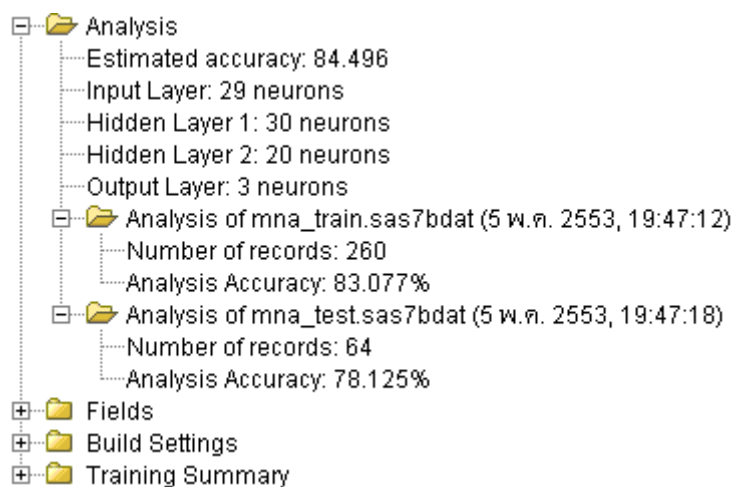
จากภาพที่ 4-21 กล่าวว่า ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ (Testing Data) จำนวน 64 เรคคอร์ด มีความสามารถในการทำนายได้ถูกต้อง เป็นจำนวน 50 เรคคอร์ด คิดเป็น 78.12% และผลการทำนายไม่ถูกต้อง จำนวน 14 เรคคอร์ด คิดเป็น 21.88%

เมื่อแจกแจงรายละเอียดของผลการทำนายด้วยข้อมูลทดสอบ (Testing Data) จะได้ว่า

- ผลการเรียนจริงที่เป็น Bad มีจำนวน 12 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Bad จำนวน 6 เรคคอร์ด และผลการทำนายผิดเป็น Normal จำนวน 6 เรคคอร์ด
- ผลการเรียนจริงที่เป็น Good มีจำนวน 29 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Good จำนวน 23 เรคคอร์ด และผลการทำนายผิดเป็น Normal จำนวน 6 เรคคอร์ด
- ผลการเรียนจริงที่เป็น Normal มีจำนวน 23 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Normal จำนวน 21 เรคคอร์ด และผลการทำนายผิดเป็น Good จำนวน 2 เรคคอร์ด

ซึ่งจากการวิเคราะห์ข้างต้น สามารถสรุปเป็นผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาการจัดการ ด้วย Neural Network ได้ดังนี้

ผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาการจัดการ ด้วย Neural Network มี Input Layer จำนวน 29 Neurons มี Hidden Layer จำนวน 2 Layer คือ Hidden Layer 1 จำนวน 30 Neurons และ Hidden Layer 2 จำนวน 20 Neurons และมี Output Layer จำนวน 3 Neurons ซึ่งโมเดลดังกล่าว มีความถูกต้องในการทำนาย คิดเป็น 78.125% ดังภาพที่ 4-22

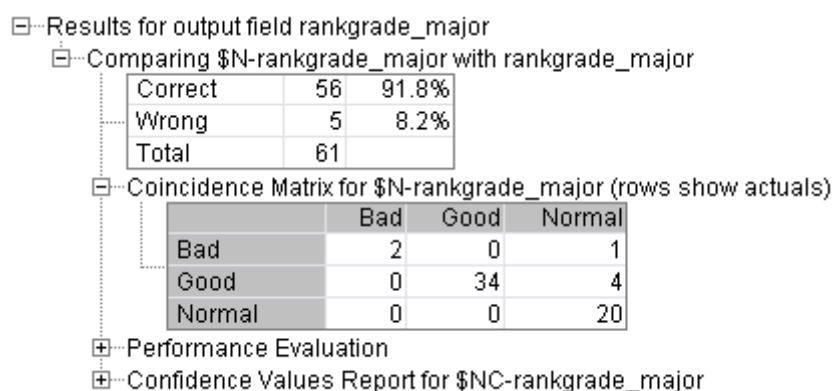


ภาพที่ 4-22 สรุปผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียน สาขาวิชาการจัดการ

#### 4.2.3 ผลการสร้างโมเดลในการทำนายผลการเรียนของสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว

สามารถแบ่งผลการวิเคราะห์ข้อมูลออกเป็นจำนวน 2 ครั้ง คือ

ครั้งที่ 1 เป็นการวิเคราะห์ข้อมูล เพื่อให้ Neurons เกิดการเรียนรู้พฤติกรรมทางการเรียนของสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว โดยใช้ข้อมูลเรียนรู้ (Training Data) แสดงได้ดังนี้



ภาพที่ 4-23 ผลการวิเคราะห์ข้อมูลการทำนายด้วยข้อมูลเรียนรู้ สาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว

จากภาพที่ 4-23 กล่าวว่า ผลการวิเคราะห์ข้อมูลการทำนายผลการเรียนด้วยข้อมูลเรียนรู้ (Training Data) จำนวน 61 เรคคอร์ด มีความสามารถในการเรียนรู้ได้ถูกต้อง จำนวน 56 เรคคอร์ด คิดเป็น 91.8% และผลการเรียนรู้ไม่ถูกต้อง จำนวน 5 เรคคอร์ด คิดเป็น 8.2%

เมื่อแจกแจงรายละเอียดของผลการวิเคราะห์ข้อมูลเรียนรู้ (Training Data) จะได้ว่า

- ผลการเรียนจริงที่เป็น Bad มีจำนวน 3 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Bad จำนวน 2 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิด เป็น Normal จำนวน 1 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Good มีจำนวน 38 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Good จำนวน 34 เรคคอร์ด และผลการวิเคราะห์ในการทำนายผิด เป็น Normal จำนวน 4 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Normal มีจำนวน 20 เรคคอร์ด มีผลการวิเคราะห์ข้อมูลในการทำนายผลการเรียนได้ถูกต้องเป็น Normal ทั้งหมดจำนวน 20 เรคคอร์ด

ครั้งที่ 2 เป็นการวิเคราะห์ข้อมูล เพื่อทำการทดสอบความถูกต้องของโมเดลที่ใช้ในการทำนายผลการเรียน โดยใช้ข้อมูลการทดสอบ (Testing Data) แสดงได้ดังนี้

Results for output field rankgrade\_major

Comparing \$N-rankgrade\_major with rankgrade\_major

Correct	13	86.67%
Wrong	2	13.33%
Total	15	

Coincidence Matrix for \$N-rankgrade\_major (rows show actuals)

	Bad	Good	Normal
Bad	1	0	2
Good	0	10	0
Normal	0	0	2

Performance Evaluation

Confidence Values Report for \$NC-rankgrade\_major

ภาพที่ 4-24 ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ สาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว

จากภาพที่ 4-24 กล่าวว่า ผลการทำนายผลการเรียนด้วยข้อมูลทดสอบ (Testing Data) จำนวน 15 เรคคอร์ด มีความสามารถในการทำนายได้ถูกต้อง เป็นจำนวน 13 เรคคอร์ด คิดเป็น 86.67% และผลการทำนายไม่ถูกต้อง จำนวน 2 เรคคอร์ด คิดเป็น 13.33%

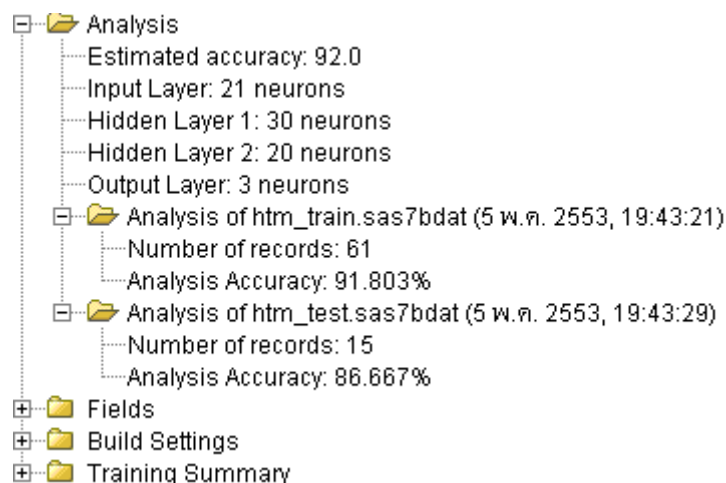
เมื่อแจกแจงรายละเอียดของผลการทำนายด้วยข้อมูลทดสอบ (Testing Data) จะได้ว่า

- ผลการเรียนจริงที่เป็น Bad มีจำนวน 3 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Bad จำนวน 1 เรคคอร์ด และผลการทำนายผิดเป็น Normal จำนวน 2 เรคคอร์ด

- ผลการเรียนจริงที่เป็น Good มีจำนวน 10 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Good ทั้งหมดจำนวน 10 เรคคอร์ด

- ผลการเรียนรู้ที่เป็น Normal มีจำนวน 2 เรคคอร์ด มีผลการทำนายผลการเรียนได้ถูกต้องเป็น Normal ทั้งหมดจำนวน 2 เรคคอร์ด

ซึ่งจากการวิเคราะห์ข้างต้น สามารถสรุปเป็นผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว ด้วย Neural Network ได้ดังนี้



ภาพที่ 4-25 สรุปผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียน สาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว

จากภาพที่ 4-25 สรุปว่าผลการวิเคราะห์ข้อมูลของโมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว ด้วย Neural Network มี Input Layer จำนวน 21 Neurons มี Hidden Layer จำนวน 2 Layer คือ Hidden Layer 1 จำนวน 30 Neurons และ Hidden Layer 2 จำนวน 20 Neurons และมี Output Layer จำนวน 3 Neurons ซึ่งโมเดลดังกล่าว มีความถูกต้องในการทำนาย คิด เป็น 86.667%



## บทที่ 5

### สรุป อภิปรายผล และข้อเสนอแนะ

#### 5.1 สรุปผลการวิจัย

การผลิิตบัณฑิตให้เป็นคนเก่ง มีความรู้ ทักษะ ความชำนาญในแต่ละสาขาวิชาได้นั้น ต้องอาศัยปัจจัยหลายอย่าง ทั้งทางด้านความรู้ในเนื้อหาวิชา คุณธรรม และจริยธรรมแล้ว การดูแลและให้คำปรึกษาในเรื่องการเรียนแก่นักศึกษา โดยอาจารย์ที่ปรึกษา จัดได้ว่าเป็นปัจจัยสำคัญที่มีส่วนช่วยพัฒนาคุณภาพการศึกษา ซึ่งงานวิจัยในครั้งนี้ได้เห็นถึงความสำคัญดังกล่าว จึงเสนอถึงขั้นตอนและวิธีการในการสร้างโมเดลการทำนายผลการเรียนของนักศึกษา วิทยาลัยราชพฤกษ์ โดยใช้เทคนิคดาต้าไมน์นิ่ง เพื่อนำโมเดลที่ได้นั้นมาใช้ในการประกอบการตัดสินใจให้คำแนะนำในการเรียนในสาขาวิชาที่เหมาะสมแก่นักศึกษาของอาจารย์ที่ปรึกษา ซึ่งได้ทำการสร้างโมเดลอยู่ 2 ประเภท คือ โมเดลในการศึกษาพฤติกรรมทางการเรียนของนักศึกษาในแต่ละสาขาวิชา โดยใช้เทคนิค Decision Tree จำนวน 4 โมเดล และโมเดลในการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา โดยใช้เทคนิค Neural Network จำนวน 4 โมเดล ซึ่งสามารถสรุปผลการสร้างโมเดลต่างๆ ได้ดังนี้

##### 5.1.1 ผลการสร้างโมเดลในการศึกษาพฤติกรรมทางการเรียนของนักศึกษาในแต่ละสาขาวิชา

- โมเดลศึกษาพฤติกรรมทางการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ สรุปผลว่าปัจจัยที่มีผลต่อการเรียนของสาขาวิชาคอมพิวเตอร์ธุรกิจคือ กลุ่มวิชาแกนการเงิน, วิชาแกนธุรกิจ และวิชาแกนทางคอมพิวเตอร์ ซึ่งกลุ่มวิชาแกนการเงิน มีผลต่อการเรียนในวิชาเอกบังคับของสาขาวิชาคอมพิวเตอร์ธุรกิจมากที่สุด โดยมีนักศึกษาจำนวน 110 คนจาก 194 คนที่มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Bad แล้วส่งผลให้มีผลการเรียนในกลุ่มวิชาเอกบังคับ เป็น Bad ไปด้วยคิดเป็น 56.701%

- โมเดลศึกษาพฤติกรรมทางการเรียนของนักศึกษาสาขาวิชาการตลาด สรุปผลว่าปัจจัยที่มีผลต่อการเรียนของสาขาวิชาการตลาดคือ กลุ่มวิชาแกนธุรกิจ, วิชาแกนการเงิน และสถานที่ตั้งของสถาบันเดิม ซึ่งกลุ่มวิชาแกนธุรกิจมีผลต่อการเรียนในวิชาเอกบังคับของสาขาวิชาการตลาดมากที่สุด โดยมีนักศึกษาจำนวน 177 คนจาก 180 คน ที่มีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Good, Normal แล้วส่งผลให้มีผลการเรียนในกลุ่มวิชาบังคับเป็น Good, Normal ซึ่งคิดเป็น 98.33%

- โมเดลศึกษาพฤติกรรมทางการเรียนของนักศึกษาสาขาวิชาการจัดการ สรุปผลว่าปัจจัยที่มีผลต่อการเรียนของสาขาวิชาการจัดการคือ กลุ่มวิชาแกนธุรกิจ, วิชาแกนทางคอมพิวเตอร์, วิชาแกนการเงิน และวุฒิการศึกษาเดิม ซึ่งกลุ่มวิชาแกนธุรกิจนั้น มีผลต่อการเรียนในวิชาเอกบังคับของ

สาขาวิชาการจัดการมากที่สุด โดยมีนักศึกษาจำนวน 224 คนจาก 239 คน ที่มีผลการเรียนในกลุ่มวิชาแกนธุรกิจเป็น Good, Normal แล้วส่งผลให้มีผลการเรียนในกลุ่มวิชาบังคับเป็น Good, Normal ซึ่งคิดเป็น 93.72%

- โมเดลศึกษาพฤติกรรมทางการเรียนของนักศึกษาสาขาวิชาการจัดการการโรงแรมและการท่องเที่ยว สรุปผลว่าปัจจัยที่มีผลต่อการเรียนของสาขาวิชาการจัดการการโรงแรมฯ คือ กลุ่มวิชาแกนการเงิน, วิชาแกนทางคอมพิวเตอร์ และวุฒิการศึกษาเดิม ซึ่งกลุ่มวิชาแกนการเงินมีผลต่อการเรียนในวิชาเอกบังคับของสาขาวิชาการจัดการการโรงแรมฯ มากที่สุดโดยมีนักศึกษาจำนวน 41 คนจาก 42 คน ที่มีผลการเรียนในกลุ่มวิชาแกนการเงินเป็น Good, Normal แล้วส่งผลให้มีผลการเรียนในกลุ่มวิชาบังคับเป็น Good, Normal ซึ่งคิดเป็น 97.62%

#### 5.1.2 ผลการสร้างโมเดลการทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา

- โมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาคอมพิวเตอร์ธุรกิจ สรุปผลการเรียนรู้ข้อมูล (Training) ของโมเดลทำนายผลการเรียน สาขาวิชาคอมพิวเตอร์ธุรกิจ มีจำนวน 333 คน มีค่าที่เรียนรู้ได้ถูกต้องคิดเป็น 74.47% เมื่อนำข้อมูลที่ใช้ในการทดสอบความถูกต้อง (Testing) ของโมเดลการทำนายผลการเรียน มีจำนวน 83 คน มีค่าความถูกต้องในการทำนายคิดเป็น 73.49%

- โมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาการตลาด สรุปผลในการเรียนรู้ข้อมูล (Training) ของโมเดลทำนายผลการเรียน สาขาวิชาการตลาด มีจำนวน 271 คน มีค่าที่เรียนรู้ได้ถูกต้องคิดเป็น 80.81% เมื่อนำข้อมูลที่ใช้ในการทดสอบความถูกต้อง (Testing) ของโมเดลการทำนายผลการเรียน มีจำนวน 67 คน มีค่าความถูกต้องในการทำนายคิดเป็น 83.58%

- โมเดลทำนายผลการเรียนของนักศึกษาสาขาวิชาการจัดการ สรุปผลการเรียนรู้ข้อมูล (Training) ของโมเดลทำนายผลการเรียน สาขาวิชาการจัดการ มีจำนวน 216 คน มีค่าที่เรียนรู้ได้ถูกต้องคิดเป็น 83.08% เมื่อนำข้อมูลที่ใช้ในการทดสอบความถูกต้อง (Testing) ของโมเดลการทำนายผลการเรียน มีจำนวน 64 คน มีค่าความถูกต้องในการทำนายคิดเป็น 78.12%

- โมเดลทำนายผลการเรียนของนักศึกษา ในสาขาวิชาการจัดการการโรงแรม และการท่องเที่ยว สรุปผลการเรียนรู้ข้อมูล (Training) ของโมเดลทำนายผลการเรียน สาขาวิชาการจัดการการโรงแรมฯ มีจำนวน 61 คน มีค่าที่เรียนรู้ได้ถูกต้องคิดเป็น 91.80% เมื่อนำข้อมูลที่ใช้ในการทดสอบความถูกต้อง (Testing) ของโมเดลการทำนายผลการเรียน มีจำนวน 15 คน มีค่าความถูกต้องในการทำนายคิดเป็น 86.67%

## 5.2 อภิปรายผลการวิจัย

ในงานวิจัยครั้งนี้ นำเสนอถึงการนำเทคนิคดาต้าไมน์นิ่งมาประยุกต์ใช้ในการทำนายผลการเรียนของนักศึกษา ซึ่งนำข้อมูลจากฐานข้อมูลนักศึกษา คณะบริหารธุรกิจ ตั้งแต่ปี 2549-2551 มาทำการวิเคราะห์หาสารสนเทศที่เป็นประโยชน์มาประกอบการตัดสินใจ ในการให้คำปรึกษาทางการเรียนของอาจารย์ที่ปรึกษา โดยมีการวิเคราะห์เพื่อศึกษาถึงพฤติกรรมทางการเรียน (Classification) ของนักศึกษาในแต่ละสาขาวิชา ทำให้ได้สารสนเทศที่สามารถบอกถึงทักษะที่จำเป็นในการเรียนของแต่ละสาขาวิชา โดยอาศัยข้อมูลปัจจัยต่างๆ เช่น ข้อมูลเกี่ยวกับผลการเรียนในกลุ่มวิชาแกนทางการเงิน แกนธุรกิจ และแกนคอมพิวเตอร์ เป็นต้น ซึ่งสามารถนำผลที่ได้ไปเป็นแนวทางในการพัฒนาการเรียนของนักศึกษา เช่น การพัฒนาการจัดการเรียนการสอน โดยมีการเพิ่มเติมทักษะทางด้านการคำนวณ การคิด และการแก้ไขปัญหาในรายวิชาที่สอน เพื่อเป็นการเสริมความรู้พื้นฐานทางการเรียนของแต่ละสาขาวิชาให้มีความเหมาะสมยิ่งขึ้น

นอกจากนี้ ยังได้นำเทคนิคดาต้าไมน์นิ่งมาใช้ในการวิเคราะห์ข้อมูล สำหรับทำนายผลการเรียนของนักศึกษาในแต่ละสาขาวิชา ซึ่งผลที่ได้จากการทำนายของโมเดลที่สร้างขึ้นนั้น มีเปอร์เซ็นต์ความถูกต้องค่อนข้างสูง โดยผลการทำนายถูกนำเสนอเป็นแนวโน้มของผลการเรียนในกลุ่มวิชาเอก บังคับของแต่ละสาขาวิชาว่าจะออกมาดี ปานกลาง หรือไม่ดี เพื่อนำไปเป็นข้อมูลที่ช่วยประกอบการตัดสินใจในการให้คำแนะนำ คำปรึกษาทางการเรียนของนักศึกษาจากอาจารย์ที่ปรึกษา ในการแนะนำสาขา วิชาที่เหมาะสมกับความถนัด และความสามารถของนักศึกษาที่ได้ผลการเรียนต่ำ ในวิชาเอกบังคับในสาขาวิชานั้นๆ ถ้าอาจารย์ที่ปรึกษาต้องการทราบสาขาวิชาที่เหมาะสมกับนักศึกษาคนหนึ่ง ก็จะนำข้อมูลของนักศึกษานั้นๆ เข้าโมเดลการทำนายผลการเรียนของสาขาวิชาต่างๆ เพื่อดูว่าผลการเรียนในกลุ่มวิชาเอกบังคับของสาขาวิชาไหนบ้าง ที่ออกมาดี หรือปานกลาง จัดได้ว่าเป็นข้อมูลหนึ่งที่มีความสำคัญมากในการแนะนำแนวทางการเรียนให้เหมาะสมกับนักศึกษา ส่งผลให้การผลิตบัณฑิตของวิทยาลัยมีประสิทธิภาพมากขึ้น

ซึ่งผลการทำนายที่ได้ มิได้หมายความว่านักศึกษาที่ผ่านการทำนาย จะมีผลการเรียนตามผลการทำนาย 100% แต่หมายความว่าข้อมูลของนักศึกษานั้นๆ มีผลการเรียนในรูปแบบเดียวกันกับกลุ่มนักศึกษาในสาขาวิชานั้นๆ ซึ่งเป็นเพียงข้อมูลที่ประกอบการตัดสินใจ นอกจากนี้จะต้องอาศัยข้อมูลประกอบการตัดสินใจอื่นๆ เช่น ความเอาใจใส่ในการเรียน การมาเรียน การส่งงาน เป็นต้น

### 5.3 ข้อเสนอแนะ

#### 5.3.1 ข้อเสนอแนะในการนำผลการวิจัยไปใช้

5.3.1.1 ในอนาคต สามารถปรับโมเดลการทำนายผลการเรียนของนักศึกษา ให้สามารถนำไปใช้ในหน่วยงานอื่นภายในวิทยาลัยได้ เช่น การเปลี่ยนข้อมูลนำเข้า เพื่อนำไปใช้ในการทำนายสาขาวิชาที่เหมาะสมให้แก่ักศึกษาใหม่ เพื่อเป็นข้อมูลในการประกอบการตัดสินใจเลือกสาขาวิชาที่เหมาะสม

5.3.1.2 ผลการวิเคราะห์ข้อมูลจากงานวิจัยครั้งนี้ เป็นเพียงข้อมูลที่ใช้ในการประกอบการตัดสินใจให้คำปรึกษาแก่อาจารย์ที่ปรึกษาเท่านั้น เพื่อให้การแนะนำมีประสิทธิภาพมากขึ้น จะต้องคำนึงถึงปัจจัยต่างๆ ที่มีผลกระทบต่อผลการเรียนของนักศึกษาด้วย

#### 5.3.2 ข้อเสนอแนะในการวิจัยครั้งต่อไป

เนื่องจากความถูกต้องของการวิเคราะห์ข้อมูล โดยใช้เทคนิคค้ำไ่มันั้นในงานวิจัยครั้งนี้ ขึ้นอยู่กับข้อมูลที่นำมาสร้างโมเดล ดังนั้น ข้อเสนอแนะจะมุ่งเน้นไปที่การรวบรวมและจัดเตรียมข้อมูล ดังนี้

5.3.2.1 เนื่องจากทางวิทยาลัยราชพฤกษ์ เป็นวิทยาลัยที่เพิ่งเปิดทำการเรียนการสอนในปี 2549 จึงทำให้มีข้อมูลที่ใช้ในการวิเคราะห์มีจำนวนน้อยเกินไป ดังนั้น ในอนาคตควรมีการรวบรวมข้อมูลเพื่อทำการวิเคราะห์อีกครั้ง เพื่อให้ได้ผลของการวิเคราะห์ข้อมูลมีความสมบูรณ์มากขึ้น

5.3.2.2 ในอนาคตควรมีการรวบรวม และจัดเตรียมข้อมูลนำเข้า ที่นำไปใช้ในการทำนายผลการเรียน โดยให้มีรายละเอียดในเรื่องของปัจจัยที่มีผลต่อผลการเรียนของนักศึกษาในแต่ละสาขาวิชามากขึ้น เพื่อให้การทำนายผลการเรียนมีความถูกต้องมากขึ้น

## บรรณานุกรม

- กฤษณะ ไวยมัย, ชิดชนก ส่งศิริ และชนาวินท์ รัชธรรมานนท์, การใช้เทคนิคดาต้าไมนิ่งเพื่อพัฒนาคุณภาพการศึกษาคณะวิศวกรรมศาสตร์, NECTEC technical Journal vol. III, No. 11 หน้า 134-142, 2544.
- การนิคมอุตสาหกรรมแห่งประเทศไทย, การจัดการความรู้ (Knowledge Management), ครั้งที่ 1, 10 เมษายน 2551. [Online Available]:  
<http://www.ieat.go.th/cms.php?lang=th&action=view&item=903>
- จามรกุล เหล่าเกียรติกุล และจิรารัตน์ สิทธิวรชาติ, ตัวแบบพยากรณ์ลักษณะความเหมาะสมของนักศึกษาใหม่สาขาวิชาเทคโนโลยีสารสนเทศ ด้วยกฎการจำแนกประเภทเชิงความสัมพันธ์, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2551. [Online Available]:  
<http://ora.chandra.ac.th/~jamornkul/publications/nccit2008-ET004.pdf>
- จิราพร ยิ่งกว่าชาติ, บุญเสริม กิจศิริกุล และ ประสงค์ปรมณีตพลกรัง, การทำนายผลสำเร็จการศึกษาของนักศึกษาระดับอุดมศึกษาด้วยการเรียนรู้แบบเบย์และการทำเหมืองข้อมูล, มหาวิทยาลัย ศรีปทุม, 2549. [Online Available]:  
[http://www.spu.ac.th/research\\_files/03\\_05.pdf](http://www.spu.ac.th/research_files/03_05.pdf)
- จีระนันท์ ตังคะวัชระ, การพยากรณ์ยอดขายอุตสาหกรรมแหวนโดยใช้เทคนิคการทำเหมืองข้อมูล, วิทยานพนธ์ปริญาวิทยาสตรมหาบัณฑิต, มหาวิทยาลัยขอนแก่น, 2550.
- จดหมายข่าวจากทีมงาน KMUTI\_KM, การจัดการความรู้, ฉบับที่ 1, 1 กันยายน 2549.  
 [Online Available]: [http://erp.rmuti.ac.th/km/manual/Letter/การจัดการความรู้\\_4.pdf](http://erp.rmuti.ac.th/km/manual/Letter/การจัดการความรู้_4.pdf)
- ณัฐริน เจริญเกียรติบวร, การใช้เทคนิคเหมืองข้อมูลเพื่อช่วยในการแนะแนวการศึกษาต่อระดับอุดมศึกษา, ภาควิชาคอมพิวเตอร์, มหาวิทยาลัยศิลปากร, 2549.
- ธวัช เกษสดีชัย, การประยุกต์ใช้โครงข่ายประสาทเทียมเพื่อพยากรณ์ความเข้มการแผ่รังสีพลังงานแสงอาทิตย์, ภาควิชาครุศาสตร์ไฟฟ้า, สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549.
- บุษรา ลีมีพิพัฒนางกูร, การพยากรณ์โอกาสการสำเร็จการศึกษาของนักศึกษา โดยใช้จินตึกอัลกอริทึมแบบหลายวัตถุประสงค์ เปรียบเทียบกับวิธีทางระบบโครงข่ายประสาทเทียม : กรณีศึกษานักศึกษาภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ, ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ, สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549.

ปรีชา ยามันสะบีตั้น, บุญเสริม กิจศิริกุล, ปิยะวัฒน์ จิระพงษ์สุวรรณ และประสงค์ ปราณีตพลกรัง,  
การประยุกต์ใช้ดาต้าไมนิ่งในการบริหารลูกค้าสัมพันธ์สำหรับนักศึกษาระดับอุดมศึกษา,  
มหาวิทยาลัยศรีปทุม, 2549. [Online Available]:

[http://www.spu.ac.th/research\\_files/03\\_15.pdf](http://www.spu.ac.th/research_files/03_15.pdf)

วิชัย กองศาสนะ, การวิเคราะห์เจาะลึกข้อมูล เพื่อการตัดสินใจด้วย BI, Vertical Industries  
Consultant, eLeader, กันยายน 2546. [Online Available]:

<http://intranet.dip.go.th/article/datafile/e-leader13.pdf>

วิทยาลัยราชพฤกษ์, คู่มืออาจารย์ที่ปรึกษา วิทยาลัยราชพฤกษ์, หน้า 2-3, วิทยาลัยราชพฤกษ์, 2550.  
เอกสารประกอบรายวิชาระเบียบวิธีวิจัยทางเทคโนโลยีสารสนเทศ, **Web mining**, ระบบห้องสมุด  
(E-library) คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น. [Online Available]:

[http://dlibed.kku.ac.th/e\\_lib2/exxe/586.pdf](http://dlibed.kku.ac.th/e_lib2/exxe/586.pdf)

## ภาคผนวก

ตัวอย่างข้อมูลนักศึกษา คณะบริหารธุรกิจ วิทยาลัยราชพฤกษ์  
ปีการศึกษาที่ 2549-2551 จำนวน 250 คนจาก 1,154 คน

STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
501010164	Upcountry	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	2_Normal	5	Good	Normal	Good	Bad	Normal
501010163	BKK	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	2_Normal	6	Bad	Normal	Good	Normal	Normal
510112019	Upcountry	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	1_Bad	6	Bad	Bad	Bad	Bad	Bad
501010129	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	2_Normal	6	Bad	Bad	Good	Bad	Normal
501010052	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	3_Good	6	Normal	Normal	Good	Bad	Normal
510124178	BKK	การจัดการ	สายอาชีพ	20-25	1_Bad	7	Bad	Bad	Bad	Bad	Bad
510112051	Upcountry	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	1_Bad	7	Bad	Good	Good	Bad	Normal
510112052	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	1_Bad	7	Bad	Bad	Good	Bad	Bad
510112012	BKK	คอมพิวเตอรืธุรกิจ	สายสามัญ	25-30	2_Normal	8	Bad	Bad	Good	Bad	Normal
510114016	BKK	การจัดการ	สายอาชีพ	20-25	3_Good	8	Good	Good	Normal	Good	Good
491010117	BKK	การจัดการ	สายอาชีพ	<=20	1_Bad	8	Bad	Bad	Bad	Bad	Bad
501010216	Upcountry	การตลาด	สายสามัญ	<=20	1_Bad	8	Normal	Bad	Bad	Bad	Bad
501010057	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	8	Bad	Good	Bad	Bad	Bad
510114006	Upcountry	การจัดการ	สายสามัญ	20-25	3_Good	8	Good	Good	Normal	Normal	Good
510123191	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	8	Normal	Bad	Good	Good	Good
510122184	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	3_Good	8	Good	Good	Good	Normal	Good
510122189	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	8	Normal	Normal	Bad	Bad	Bad
510122185	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	35+	3_Good	8	Good	Good	Good	Normal	Good
510122186	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	2_Normal	8	Normal	Good	Bad	Bad	Normal
510122192	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	3_Good	8	Good	Good	Good	Bad	Good
502010499	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	9	Bad	Bad	Bad	Bad	Bad
510113035	BKK	การตลาด	สายสามัญ	20-25	2_Normal	9	Good	Normal	Good	Bad	Normal
510113034	BKK	การตลาด	สายอาชีพ	<=20	2_Normal	9	Good	Normal	Good	Good	Good
510114052	Upcountry	การจัดการ	สายอาชีพ	<=20	2_Normal	9	Good	Normal	Good	Bad	Normal
510114039	Upcountry	การจัดการ	สายสามัญ	25-30	3_Good	9	Good	Normal	Good	Bad	Good
510114047	Upcountry	การจัดการ	สายอาชีพ	20-25	3_Good	9	Good	Good	Good	Good	Good
510114053	BKK	การจัดการ	สายสามัญ	20-25	2_Normal	9	Good	Bad	Good	Normal	Normal



STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510114055	BKK	การจัดการ	สายอาชีพ	25-30	3_Good	9	Good	Good	Good	Good	Good
510114036	Upcountry	การจัดการ	สายสามัญ	25-30	2_Normal	9	Good	Bad	Bad	Bad	Bad
510114018	BKK	การจัดการ	สายสามัญ	25-30	3_Good	9	Normal	Normal	Normal	Normal	Normal
510114001	BKK	การจัดการ	สายสามัญ	<=20	2_Normal	9	Bad	Bad	Normal	Normal	Normal
510114044	BKK	การจัดการ	สายอาชีพ	25-30	3_Good	9	Good	Good	Good	Good	Good
510114019	BKK	การจัดการ	สายสามัญ	35+	3_Good	9	Good	Normal	Bad	Normal	Normal
510114040	Upcountry	การจัดการ	สายอาชีพ	<=20	1_Bad	9	Bad	Bad	Normal	Bad	Bad
510124028	BKK	การจัดการ	สายอาชีพ	<=20	2_Normal	10	Normal	Bad	Good	Bad	Normal
510123042	NA	การตลาด	N/A	N/A	N/A	10	Bad	Bad	Bad	Bad	Bad
510112001	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	3_Good	10	Good	Good	Good	Good	Good
510112026	BKK	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	2_Normal	10	Normal	Normal	Good	Bad	Normal
510112027	BKK	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	2_Normal	10	Good	Bad	Normal	Normal	Normal
510112020	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	2_Normal	10	Bad	Normal	Bad	Bad	Bad
510112009	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	3_Good	10	Good	Good	Good	Good	Good
510112003	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	3_Good	10	Good	Good	Bad	Good	Good
510112035	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	2_Normal	10	Bad	Normal	Good	Bad	Normal
510112041	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	2_Normal	10	Normal	Bad	Bad	Bad	Bad
510112005	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	1_Bad	10	Bad	Bad	Bad	Bad	Bad
510112044	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	20-25	2_Normal	10	Bad	Bad	Bad	Bad	Bad
510112037	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	2_Normal	10	Normal	Normal	Good	Bad	Normal
510112010	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	3_Good	10	Bad	Good	Good	Good	Good
510112024	BKK	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	3_Good	10	Good	Normal	Good	Good	Good
510112045	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	1_Bad	10	Bad	Bad	Bad	Normal	Bad
510112034	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	1_Bad	10	Bad	Bad	Bad	Normal	Bad
510112004	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	<=20	1_Bad	10	Bad	Bad	Bad	Bad	Bad
510112046	BKK	คอมพิวเตอร์ธุรกิจ	สายสามัญ	20-25	1_Bad	10	Bad	Bad	Bad	Bad	Bad

STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510112047	Upcountry	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	2_Normal	10	Normal	Good	Bad	Good	Normal
510112011	BKK	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	3_Good	10	Good	Good	Good	Good	Good
510112022	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	2_Normal	10	Normal	Normal	Good	Good	Good
510112013	BKK	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	2_Normal	10	Bad	Good	Normal	Bad	Bad
510112048	BKK	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	2_Normal	10	Normal	Normal	Good	Normal	Normal
510112008	BKK	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	3_Good	10	Good	Good	Bad	Good	Good
510112014	Upcountry	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	3_Good	10	Good	Good	Bad	Good	Good
510112049	Upcountry	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	3_Good	10	Good	Good	Bad	Good	Good
510112015	BKK	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	2_Normal	10	Normal	Normal	Bad	Bad	Bad
510115051	Upcountry	การจัดการโรงแรมฯ	สายสามัญ	<=20	1_Bad	10	Good	Bad	Good	Bad	Normal
510124096	BKK	การจัดการ	สายอาชีพ	20-25	2_Normal	10	Normal	Bad	Bad	Bad	Bad
510123190	BKK	การตลาด	สายอาชีพ	20-25	2_Normal	10	Good	Bad	Normal	Bad	Normal
510122136	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	2_Normal	10	Bad	Good	Good	Normal	Normal
510123144	Upcountry	การตลาด	สายอาชีพ	<=20	1_Bad	10	Normal	Bad	Bad	Bad	Bad
510123089	BKK	การตลาด	สายอาชีพ	<=20	1_Bad	11	Bad	Bad	Bad	Bad	Bad
501010109	BKK	คอมพิวเตอรืธุรกิจ	สายสามัญ	<=20	1_Bad	11	Bad	Bad	Bad	Bad	Bad
510113012	BKK	การตลาด	สายสามัญ	<=20	3_Good	11	Good	Good	Good	Good	Good
510113020	Upcountry	การตลาด	สายสามัญ	<=20	2_Normal	11	Bad	Normal	Bad	Bad	Bad
510113024	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	11	Good	Normal	Good	Bad	Normal
502010208	NA	คอมพิวเตอรืธุรกิจ	N/A	N/A	N/A	11	Bad	Bad	Bad	Bad	Bad
502010154	BKK	การตลาด	สายอาชีพ	<=20	1_Bad	11	Bad	Bad	Bad	Bad	Bad
491010020	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	1_Bad	11	Bad	Bad	Bad	Bad	Bad
510122061	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	25-30	2_Normal	11	Bad	Bad	Good	Normal	Normal
510122060	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	30-35	1_Bad	11	Bad	Bad	Good	Bad	Normal
510122022	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	2_Normal	12	Normal	Good	Normal	Bad	Normal
510123069	Upcountry	การตลาด	สายอาชีพ	<=20	2_Normal	12	Good	Normal	Bad	Bad	Normal

STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510113006	BKK	การตลาด	สายสามัญ	<=20	1_Bad	12	Normal	Bad	Bad	Bad	Bad
510113016	BKK	การตลาด	สายสามัญ	<=20	2_Normal	12	Good	Normal	Good	Good	Good
510113023	Upcountry	การตลาด	สายสามัญ	35+	1_Bad	12	Good	Normal	Bad	Bad	Bad
510113013	Upcountry	การตลาด	สายอาชีพ	<=20	2_Normal	12	Normal	Bad	Bad	Bad	Bad
510113022	Upcountry	การตลาด	สายสามัญ	<=20	2_Normal	12	Good	Normal	Good	Bad	Normal
510113030	Upcountry	การตลาด	สายสามัญ	20-25	1_Bad	12	Normal	Normal	Normal	Bad	Normal
510113019	Upcountry	การตลาด	สายสามัญ	<=20	1_Bad	12	Bad	Bad	Bad	Bad	Bad
502010198	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	30-35	1_Bad	12	Bad	Bad	Bad	Bad	Bad
510124032	BKK	การจัดการ	สายอาชีพ	20-25	3_Good	12	Good	Good	Good	Bad	Good
510125012	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	<=20	3_Good	12	Good	Normal	Good	Good	Good
510125013	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	<=20	3_Good	12	Good	Normal	Good	Good	Good
510125006	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Good	Good	Good	Good
510125011	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Normal	Good	Bad	Normal
510125008	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Normal	Good	Good	Good
510125026	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	2_Normal	12	Good	Bad	Bad	Normal	Normal
510125007	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	2_Normal	12	Good	Normal	Good	Bad	Good
510125022	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Good	Good	Good	Good
510125015	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	2_Normal	12	Good	Bad	Normal	Bad	Normal
510125005	BKK	การจัดการโรงแรมฯ	สายอาชีพ	<=20	3_Good	12	Good	Good	Good	Good	Good
510125010	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Bad	Good	Bad	Normal
510125003	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	20-25	1_Bad	12	Normal	Bad	Bad	Bad	Bad
510125019	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Normal	Good	Good	Good
510125024	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Normal	Good	Bad	Normal
510125020	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	<=20	3_Good	12	Good	Normal	Good	Normal	Good
510125002	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Good	Good	Good	Good
510125001	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	1_Bad	12	Normal	Bad	Bad	Bad	Bad

STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510125004	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	2_Normal	12	Good	Bad	Good	Bad	Normal
510122148	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	1_Bad	12	Bad	Bad	Bad	Bad	Bad
510125027	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	20-25	2_Normal	12	Normal	Bad	Bad	Bad	Bad
510123172	BKK	การตลาด	สายอาชีพ	20-25	2_Normal	12	Normal	Normal	Bad	Bad	Normal
510125023	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	2_Normal	12	Normal	Bad	Bad	Bad	Bad
510125025	BKK	การจัดการโรงแรมฯ	สายอาชีพ	25-30	3_Good	12	Good	Good	Good	Good	Good
510125029	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	20-25	2_Normal	12	Good	Normal	Good	Normal	Normal
510125016	BKK	การจัดการโรงแรมฯ	สายอาชีพ	20-25	3_Good	12	Good	Good	Good	Good	Good
510125032	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	<=20	3_Good	12	Good	Good	Good	Normal	Good
510124031	Upcountry	การจัดการ	สายอาชีพ	20-25	3_Good	12	Good	Good	Good	Good	Good
510124181	Upcountry	การจัดการ	สายอาชีพ	20-25	1_Bad	12	Bad	Bad	Bad	Bad	Bad
510124158	BKK	การจัดการ	สายอาชีพ	20-25	2_Normal	12	Bad	Normal	Bad	Bad	Normal
502010338	Upcountry	การจัดการ	สายอาชีพ	20-25	1_Bad	12	Bad	Normal	Bad	Bad	Bad
510124079	Upcountry	การจัดการ	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510123045	BKK	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Good	Normal	Bad	Normal
510123007	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	13	Normal	Bad	Bad	Bad	Bad
510123081	BKK	การตลาด	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510123014	BKK	การตลาด	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510123121	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	13	Normal	Bad	Bad	Bad	Bad
510123015	BKK	การตลาด	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510123057	Upcountry	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Normal	Good	Normal	Normal
510123073	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	13	Normal	Normal	Good	Bad	Normal
510123056	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	13	Good	Bad	Bad	Normal	Normal
510123136	BKK	การตลาด	สายอาชีพ	<=20	2_Normal	13	Good	Good	Bad	Normal	Normal
510123046	BKK	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Normal	Normal	Normal	Normal
510123148	BKK	การตลาด	สายอาชีพ	25-30	3_Good	13	Good	Good	Good	Good	Good

STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510123047	BKK	การตลาด	สายอาชีพ	<=20	3_Good	13	Good	Bad	Good	Good	Good
510123163	Upcountry	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Bad	Normal	Normal	Normal
510123080	BKK	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Normal	Bad	Normal	Normal
510123050	Upcountry	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Good	Good	Normal	Good
510123161	BKK	การตลาด	สายอาชีพ	25-30	3_Good	13	Good	Good	Good	Bad	Good
510123049	BKK	การตลาด	สายอาชีพ	<=20	2_Normal	13	Normal	Bad	Bad	Bad	Normal
510123109	Upcountry	การตลาด	สายอาชีพ	<=20	2_Normal	13	Good	Bad	Good	Bad	Normal
510123110	BKK	การตลาด	สายอาชีพ	20-25	2_Normal	13	Bad	Bad	Bad	Bad	Bad
510123051	BKK	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Normal	Bad	Normal	Normal
510123155	BKK	การตลาด	สายอาชีพ	25-30	3_Good	13	Good	Good	Good	Good	Good
510123054	BKK	การตลาด	สายอาชีพ	20-25	2_Normal	13	Normal	Bad	Good	Bad	Normal
510123115	Upcountry	การตลาด	สายอาชีพ	30-35	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510123085	Upcountry	การตลาด	สายอาชีพ	<=20	2_Normal	13	Normal	Bad	Bad	Bad	Bad
510123127	BKK	การตลาด	สายอาชีพ	25-30	3_Good	13	Good	Normal	Good	Bad	Good
510123137	BKK	การตลาด	สายอาชีพ	<=20	2_Normal	13	Normal	Bad	Bad	Bad	Bad
510123160	Upcountry	การตลาด	สายอาชีพ	25-30	3_Good	13	Good	Bad	Normal	Good	Normal
510123130	Upcountry	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Good	Good	Good	Good
510122010	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	3_Good	13	Normal	Good	Good	Normal	Good
510122006	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	3_Good	13	Good	Bad	Good	Normal	Good
510122036	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	3_Good	13	Normal	Good	Good	Normal	Good
510122037	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Normal	Good	Normal	Normal
510122039	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	3_Good	13	Normal	Normal	Good	Normal	Good
510122011	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Normal	Good	Normal	Normal
510122008	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	2_Normal	13	Normal	Normal	Good	Bad	Normal
510122023	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	2_Normal	13	Normal	Bad	Good	Normal	Normal
510122049	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Bad	Bad	Normal	Normal

STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510122038	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	1_Bad	13	Bad	Bad	Normal	Bad	Bad
510122032	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	3_Good	13	Good	Good	Good	Good	Good
510122001	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	3_Good	13	Good	Bad	Good	Normal	Good
510122009	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	3_Good	13	Good	Good	Good	Good	Good
501010143	Upcountry	การจัดการ	สายอาชีพ	25-30	3_Good	13	Good	Good	Bad	Bad	Normal
510114048	BKK	การจัดการ	สายสามัญ	20-25	1_Bad	13	Normal	Normal	Bad	Bad	Normal
510123177	Upcountry	การตลาด	สายอาชีพ	25-30	3_Good	13	Good	Normal	Good	Normal	Normal
510123126	Upcountry	การตลาด	สายอาชีพ	20-25	3_Good	13	Good	Good	Good	Normal	Good
501010144	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	20-25	1_Bad	13	Normal	Bad	Bad	Bad	Bad
510123145	Upcountry	การตลาด	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510123120	Upcountry	การตลาด	สายอาชีพ	25-30	2_Normal	13	Normal	Normal	Bad	Bad	Bad
510123122	Upcountry	การตลาด	สายอาชีพ	30-35	3_Good	13	Good	Good	Normal	Normal	Good
502010469	BKK	การตลาด	สายอาชีพ	20-25	1_Bad	13	Normal	Bad	Bad	Bad	Bad
510123169	BKK	การตลาด	สายอาชีพ	35+	3_Good	13	Good	Bad	Good	Good	Good
510123133	Upcountry	การตลาด	สายอาชีพ	<=20	2_Normal	13	Good	Bad	Bad	Bad	Normal
510122194	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510122168	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Normal	Good	Bad	Normal
510122177	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	2_Normal	13	Normal	Bad	Good	Bad	Normal
510122174	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510122165	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Normal	Bad	Bad
510122171	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	3_Good	13	Good	Normal	Good	Good	Good
510122164	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Normal	Good	Bad	Normal
510122163	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Normal	Normal	Bad	Normal
510122172	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	30-35	2_Normal	13	Normal	Good	Good	Good	Good
510122180	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Normal	Bad	Bad	Bad
510122175	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Normal	Normal	Bad	Bad

STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510114033	BKK	การจัดการ	สายสามัญ	20-25	2_Normal	13	Normal	Good	Bad	Normal	Normal
501010093	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	1_Bad	13	Bad	Bad	Normal	Bad	Bad
510122126	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Good	Good	Bad	Normal
510122094	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Good	Good	Normal	Good
510122140	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Bad	Normal	Good	Normal	Normal
510122134	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Normal	Good	Normal	Normal
510122065	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	2_Normal	13	Normal	Bad	Bad	Good	Normal
510122078	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	30-35	3_Good	13	Normal	Good	Good	Good	Good
510122085	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	3_Good	13	Good	Good	Good	Good	Good
510122093	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	25-30	2_Normal	13	Normal	Good	Good	Normal	Good
510122125	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Good	Good	Bad	Normal
510122084	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510122103	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Bad	Bad	Bad	Bad	Bad
510122069	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	3_Good	13	Normal	Good	Good	Good	Good
510122067	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	1_Bad	13	Normal	Bad	Bad	Bad	Bad
510122079	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	25-30	2_Normal	13	Bad	Good	Normal	Normal	Normal
510122095	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	3_Good	13	Good	Good	Good	Normal	Good
510122153	BKK	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Bad	Bad	Normal	Bad	Bad
510122114	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	3_Good	13	Normal	Bad	Good	Good	Normal
510122130	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	25-30	2_Normal	13	Normal	Bad	Good	Normal	Normal
510122066	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	20-25	2_Normal	13	Normal	Normal	Good	Normal	Normal
510122113	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	3_Good	13	Good	Good	Good	Good	Good
491010103	NA	การจัดการ	N/A	N/A	N/A	13	Bad	Bad	Bad	Bad	Bad
501010035	BKK	การจัดการโรงแรมฯ	สายอาชีพ	<=20	2_Normal	13	Good	Normal	Good	Normal	Good
502010326	Upcountry	การจัดการโรงแรมฯ	สายอาชีพ	20-25	1_Bad	14	Bad	Bad	Bad	Bad	Bad
501010156	Upcountry	การจัดการ	สายอาชีพ	20-25	3_Good	14	Good	Good	Good	Good	Good

STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510122051	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	2_Normal	14	Normal	Normal	Normal	Normal	Normal
510122059	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	2_Normal	14	Normal	Bad	Good	Bad	Normal
510122053	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	3_Good	14	Good	Good	Good	Normal	Good
510122056	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	3_Good	14	Normal	Good	Good	Good	Good
510122002	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	3_Good	14	Good	Good	Good	Good	Good
510122055	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	3_Good	14	Good	Good	Good	Good	Good
510122014	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	3_Good	14	Good	Good	Good	Good	Good
510122045	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	3_Good	14	Good	Normal	Good	Normal	Good
510122062	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	14	Bad	Bad	Normal	Bad	Bad
510122058	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	2_Normal	14	Good	Bad	Good	Bad	Normal
510122046	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	2_Normal	14	Good	Normal	Good	Bad	Normal
510122048	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	1_Bad	14	Bad	Bad	Bad	Bad	Bad
510122054	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	2_Normal	14	Bad	Bad	Normal	Bad	Bad
510122047	Upcountry	คอมพิวเตอรืธุรกิจ	สายอาชีพ	<=20	2_Normal	14	Normal	Normal	Normal	Bad	Normal
510122003	BKK	คอมพิวเตอรืธุรกิจ	สายอาชีพ	20-25	3_Good	14	Normal	Normal	Good	Good	Good
510123003	BKK	การตลาด	สายอาชีพ	<=20	1_Bad	14	Normal	Bad	Normal	Bad	Bad
510123096	BKK	การตลาด	สายอาชีพ	35+	3_Good	14	Good	Bad	Bad	Normal	Normal
510123083	BKK	การตลาด	สายอาชีพ	<=20	1_Bad	14	Bad	Bad	Normal	Bad	Bad
510123038	BKK	การตลาด	สายอาชีพ	<=20	2_Normal	14	Normal	Bad	Bad	Bad	Bad
510123099	BKK	การตลาด	สายอาชีพ	35+	3_Good	14	Good	Normal	Bad	Good	Normal
510123005	BKK	การตลาด	สายอาชีพ	20-25	2_Normal	14	Good	Bad	Normal	Bad	Normal
510123100	BKK	การตลาด	สายอาชีพ	25-30	2_Normal	14	Normal	Bad	Bad	Bad	Bad
510123017	Upcountry	การตลาด	สายอาชีพ	<=20	2_Normal	14	Normal	Bad	Normal	Bad	Bad
510123044	BKK	การตลาด	สายอาชีพ	<=20	2_Normal	14	Normal	Bad	Bad	Bad	Bad
510123021	BKK	การตลาด	สายอาชีพ	20-25	2_Normal	14	Normal	Bad	Normal	Bad	Bad
510123125	Upcountry	การตลาด	สายอาชีพ	30-35	2_Normal	14	Good	Bad	Bad	Bad	Bad



STUDENTCODE	study	PROGRAMNAME	rank_study_group	age_group	GPA_old_group	cnt_grade	rankgrade_major	rankgrade_business	rankgrade_computer	rankgrade_finance	rankgrade_Total
510123001	BKK	การตลาด	สายอาชีพ	20-25	3_Good	14	Good	Good	Bad	Normal	Normal
510123026	BKK	การตลาด	สายอาชีพ	20-25	2_Normal	14	Normal	Bad	Bad	Bad	Bad
510123043	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	14	Normal	Bad	Normal	Bad	Bad
510123018	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	14	Normal	Bad	Bad	Bad	Bad
510112054	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	1_Bad	14	Normal	Bad	Normal	Bad	Bad
501010198	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	2_Normal	14	Normal	Bad	Normal	Bad	Normal
501010169	Upcountry	คอมพิวเตอร์ธุรกิจ	สายสามัญ	20-25	1_Bad	14	Bad	Bad	Bad	Bad	Bad
501010199	Upcountry	คอมพิวเตอร์ธุรกิจ	สายอาชีพ	<=20	3_Good	14	Good	Good	Good	Good	Good
510124012	BKK	การจัดการ	สายอาชีพ	20-25	1_Bad	14	Bad	Bad	Bad	Bad	Bad
501010055	Upcountry	การจัดการ	สายสามัญ	35+	3_Good	14	Normal	Good	Good	Normal	Good
510124084	BKK	การจัดการ	สายอาชีพ	35+	3_Good	14	Good	Good	Good	Good	Good
510123147	Upcountry	การตลาด	สายอาชีพ	35+	2_Normal	14	Normal	Bad	Bad	Bad	Bad
510123178	Upcountry	การตลาด	สายอาชีพ	20-25	2_Normal	14	Good	Bad	Bad	Bad	Bad
510123151	Upcountry	การตลาด	สายอาชีพ	20-25	3_Good	14	Good	Normal	Good	Normal	Good
510123176	BKK	การตลาด	สายอาชีพ	25-30	3_Good	14	Good	Bad	Bad	Normal	Normal