สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

ระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

DECISION SUPPORT SYSTEM FOR SALES AT ORISOFT CO., LTD.

T144195

โดย

ณัฏฐพัชร กิตติชัยศรี

NATTAPAT KITTICHAISRI

อาจารย์ที่ปรึกษา

ดร.สิงหะ ฉวีสุข

QA76.3 .ณ3187 2557
เลขหมู่.....
เลขทะเบียน.....144195
วัน,เดือน,ปี...09 ก.ย. 2559
b.00264390
b.12817399
i.....

รายงานฉบับนี้เป็นส่วนหนึ่งของวิชาศึกษาอิสระ 2

หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2557

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DECISION SUPPORT SYSTEM FOR SALES AT ORISOFT CO., LTD.

NATTAPAT KITTICHAISRI

A REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS OF THE COURSE

INDEPENDENT STUDY 2

MASTER OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2/2014

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2015

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'SINSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบรับรองการศึกษาอิสระ 2 (Independent Study 2)

เรื่อง

ระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

DECISION SUPPORT SYSTEM FOR SALES AT ORISOFT CO., LTD.

ณัฏฐพัชร กิตติชัยศรี

รหัสประจำตัว 56606035

ขอรับรองว่ารายงานฉบับนี้ ข้าพเจ้าไม่ได้คัดลอกมาจากที่ใด

รายงานฉบับนี้ได้รับการตรวจสอบและอนุมัติให้เป็นส่วนหนึ่งของ

การศึกษาวิชาการศึกษาอิสระ หลักสูตรวิทยาศาสตรมหาบัณฑิต (เทคโนโลยีสารสนเทศ)

ภาคเรียนที่ 2 ปีการศึกษา 2557

.....**อาจารย์ที่ปรึกษา**

(ดร.สิงหะ ฉวีสุข)

.....**กรรมการสอบ**

(รศ.ดร.อาริต ธรรมโน)

.....**กรรมการสอบ**

(ดร.ณฐพล พันธุวงศ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	ระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์
นักศึกษา	นางสาวณัฏฐพัชร กิตติชัยศรี
รหัสนักศึกษา	56606035
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	เทคโนโลยีระบบสารสนเทศ
ปีการศึกษา	2557
อาจารย์ที่ปรึกษา	ดร. สิงหะ ฉวีสุข

บทคัดย่อ

ในปัจจุบันการบริหารจัดการของฝ่ายบุคคลมีการนำเอาเทคโนโลยีทางด้านสารสนเทศมาช่วยในการบริหารจัดการให้มีประสิทธิภาพเพิ่มมากขึ้น ซึ่งก็ส่งผลทำให้บริษัทที่เป็นผู้ผลิตซอฟต์แวร์ทางด้านการบริหารจัดการฝ่ายบุคคลจึงมีเพิ่มมากขึ้นตามความต้องการของบริษัทต่างๆที่ต้องการนำเทคโนโลยีมาใช้เหล่านี้มาใช้ในบริษัทของตน ดังนั้นการที่บริษัทเจ้าของซอฟต์แวร์เหล่านี้ ต้องการที่จะประสบความสำเร็จทางด้านยอดขาย จำเป็นต้องมีระบบวิเคราะห์ข้อมูลที่ประสิทธิภาพ ซึ่งจะช่วยวิเคราะห์ความต้องการของลูกค้าว่าซอฟต์แวร์หรือบริการเสริมรูปแบบไหนที่จะเป็นที่ต้องการของลูกค้า และสามารถที่จะนำข้อมูลเหล่านี้ไปใช้เป็นตัวชี้วัดว่าควรจะเน้นการขายซอฟต์แวร์หรือบริการเสริมรูปแบบไหน หรือสามารถที่จะแนะนำลูกค้าได้ว่าซอฟต์แวร์รูปแบบไหนเหมาะกับความต้องการของลูกค้า โดยที่การศึกษาวิชาอิสระชิ้นนี้สร้างขึ้นเพื่อพัฒนาระบบการวิเคราะห์ข้อมูลการขาย สามารถทำนายแนวโน้มของการขาย รวมทั้งหาความสัมพันธ์ของข้อมูลทางด้านการขายซอฟท์แวร์หรือบริการที่เกี่ยวข้องการบริหารจัดการฝ่ายบุคคลอีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Decision Support System for Sales at Orisoft (Thailand) Co., Ltd.
Student	Miss Nattapat Kittichaisri
Student ID.	56606035
Degree	Master of Science
Program	Information Technology
Major	Information System Technology
Academic Year	2014
Advisor	Dr. Singha Chawesuk

ABSTRACT

Nowadays, human resource management brings the information technology to help to manage the organization more efficiently. That will affect to the company which develop the software about human resource management have more company following to the demand of the companies which need to use this technology. So if these software companies need to have the good circulation, it will need to have efficient data analysis system which it will help to analysis the customer demand and can use these data to indicate that it should emphasize to sell which software or services. Furthermore, it can advise the customer which software suitable for their company. The main purpose of this research and development, it will develop the sales data analysis system and can predict the software sales circulation. Furthermore, it also can calculate data relationship about human resource management software or service sales information.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

รายงานการศึกษาอิสระฉบับนี้สำเร็จลุล่วงตามเป้าหมายได้ด้วยความอนุเคราะห์ของบุคคลหลายท่าน ซึ่งไม่อาจจะนำมากล่าวได้ทั้งหมด ซึ่งผู้มีพระคุณท่านแรกที่ข้าพเจ้าใคร่ขอบกราบขอบพระคุณคือ ดร.สิงหะ ฉวีสุข อาจารย์ที่ปรึกษา ที่ได้กรุณาสละเวลาให้คำแนะนำ ให้ข้อคิดเห็นเพื่อแก้ไขข้อบกพร่องต่างๆ ด้วยความใส่ใจข้าพเจ้าด้วยดีเสมอมา จึงขอขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอบพระคุณบริษัทออริซอฟท์ที่อนุญาตให้ใช้ข้อมูลการซื้อขายซอฟต์แวร์ย้อนหลังมาใช้ในการวิเคราะห์ และให้คำปรึกษาซึ่งเป็นส่วนสำคัญต่อการพัฒนาระบบนี้ พร้อมทั้งกำลังใจและความช่วยเหลือตลอดมา

สุดท้ายนี้ขอขอบพระคุณบิดา มารดา ครอบครัวและเพื่อนๆ ของข้าพเจ้าที่อยู่เบื้องหลังในความสำเร็จ คอยเป็นกำลังใจที่ดี สนับสนุนในทุกๆเรื่อง จนทำให้การค้นคว้าอิสระฉบับนี้สำเร็จลุล่วงด้วยดี

ประโยชน์ที่ได้รับจากการค้นคว้าอิสระฉบับนี้ข้าพเจ้าของมอบให้ผู้สนใจศึกษาต่อไป

ณัฏฐพัชร กิตติชัยศรี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อ I

ABSTRACT II

กิตติกรรมประกาศ III

สารบัญ IV

สารบัญตาราง VI

สารบัญรูป VIII

บทที่ 1 บทนำ 1

1.1 ความเป็นมาของระบบ 1

1.2 วัตถุประสงค์ 2

1.3 ขอบเขตของงาน 2

1.4 แนวทางการศึกษาและพัฒนา 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

บทที่ 2 ทฤษฎีและเทคโนโลยีที่เกี่ยวข้อง 4

2.1 กระบวนการที่เกี่ยวข้องในการพัฒนาระบบ 4

2.1.1 วงจรการพัฒนาระบบ (SDLC : System Development Life Cycle) 4

2.1.2 กระบวนการ CRISP-DM 6

2.2 การออกแบบระบบโดยใช้ยูเอ็มแอล 8

2.3 การออกแบบฐานข้อมูล 9

2.4 ทฤษฎีที่เกี่ยวข้องในการพัฒนาระบบ 9

2.4.1 การจำแนกประเภทข้อมูล (Data Classification) 10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.4.2 การจำแนกประเภทข้อมูลด้วยวิธี Naive Bayes 12

2.4.3 การจำแนกประเภทด้วยวิธี K-Nearest Neighbors 14

2.4.4 การจำแนกประเภทด้วยวิธี Decision Tree 17

2.4.5 การจัดกลุ่ม (Clustering) 24

2.4.6 ระบบลูกค้าสัมพันธ์ (CRM) 26

2.4.7 ระบบสนับสนุนการตัดสินใจ(DSS) 27

2.5 เทคโนโลยีที่เกี่ยวข้องกับการพัฒนาระบบ 32

2.5.1 ภาษา C# (C# Programming Language) 32

2.5.2 Visual Studio Express 32

2.5.3 Microsoft Excel 33

2.5.4 Rapidminer Studio 6 34

บทที่ 3 การวิเคราะห์ระบบปัจจุบัน 35

3.1 การทำงานของระบบงานปัจจุบัน 35

3.2 แนวทางการแก้ปัญหา 37

บทที่ 4 การวิเคราะห์และออกแบบระบบใหม่ 38

4.1 ยูสเคสไดอะแกรม (Use Case Diagram) 38

4.1.1 แผนภาพยูสเคสไดอะแกรม (Use Case Diagram) 38

4.1.2 รายละเอียดยูสเคส 39

4.2 แอคติวิตี้ไดอะแกรม (Activity Diagram) 46

4.3 ซีเควนไดอะแกรม (Sequence Diagram) 52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

บทที่ 5 การออกแบบฐานข้อมูล 55

5.1 อีอาร์ไดอะแกรม (ER Diagram) 55

5.2 พจนานุกรมข้อมูล 56

5.3 คลาสไดอะแกรม (Class Diagram) 57

บทที่ 6 การออกแบบส่วนต่อประสานผู้ใช้ 55

6.1 ส่วนต่อประสานผู้ใช้สำหรับใส่รายละเอียดข้อมูล 60

6.2 ส่วนต่อประสานผู้ใช้สำหรับการวิเคราะห์ข้อมูล 61

บทที่ 7 สรุปผลและข้อเสนอแนะ 62

7.1 สรุปโครงการ 62

7.2 ประโยชน์ที่ได้รับจากการออกแบบและพัฒนาระบบ 71

7.3 ข้อเสนอแนะและแนวทางในการพัฒนาระบบเพิ่มเติม 72

บรรณานุกรม 73

ประวัติผู้เขียน 74

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

4.1 ยูสเคส โหลดข้อมูล....36

4.2 ยูสเคส ส่งผ่านข้อมูล....36

4.3 ยูสเคส สร้างข้อมูล....37

4.4 ยูสเคส แสดงข้อมูล....37

4.5 ยูสเคส จัดเตรียมข้อมูล....38

4.6 ยูสเคส สร้างตัวแบบข้อมูล....39

4.7 ยูสเคส ดูผลการวิเคราะห์ข้อมูล....39

4.8 ยูสเคส นำผลที่ได้ไปใช้ประกอบการแนะนำลูกค้า....40

4.9 ยูสเคส นำผลที่ได้ไปใช้ในการทำนาย....41

5.7 ความหมายคลาสของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัท ออริซอฟท์ 56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ หน้า

2-1 แสดงวงจรการทำงานของ SDLC 4

2-2 แสดงกระบวนการ CRISP-DM 6

2-4 แสดงภาพจำลองกระบวนการสร้างตัวโมเดลจำแนกประเภทข้อมูล 11

2-5 ตารางแสดงสภาพอากาศของตัวอย่าง Naive Bayes 11

2-6 ตารางแสดงประเภทของ Iris 14

2-7 แสดงการกระจายตัวของข้อมูลจากตารางแสดงประเภท Iris 15

2-8 แสดงโครงสร้างของต้นไม้ตัดสินใจ 15

2-9 ตารางแสดงการตัดสินใจในการซื้อคอมพิวเตอร์ 15

2-10 แสดงโหนดของการตัดสินใจในการซื้อคอมพิวเตอร์ 22

2-11 แสดงโหนดของการตัดสินใจในการซื้อคอมพิวเตอร์ในรูปแบบโครงสร้าง 24

2-12 แสดงภาพจำลองการจัดกลุ่ม 24

2-13 แสดงโปรแกรม Microsoft Visual Studio Express 31

2-14 แสดงโปรแกรม Microsoft Excel 31

2-15 แสดงโปรแกรม Rapidminer Studio 6 32

3-1 แสดงขั้นตอนการทำงานภายในองค์กร 35

4-1 แสดงยูสเคสไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์ บริษัทออริซอฟท์ 36

4-2 แสดงแอคติวิตี้ไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์ บริษัทออริซอฟท์ 44

4-3 แสดงแอคติวิตี้ไดอะแกรมของการสร้างฐานข้อมูลการซื้อขายและข้อมูลลูกค้า 44

4-4 แสดงแอคติวิตี้ไดอะแกรมของการจัดเตรียมข้อมูลเพื่อนำไปใช้ในการวิเคราะห์ 45

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ หน้า

4-5 แสดงแอคติวิตี้ไดอะแกรมของการทำนายและแสดงผลสัดส่วนด้วยแผนภาพวงกลม.....46

4-6 แสดงแอคติวิตี้ไดอะแกรมของการทำนายผล.....47

4-7 แสดงแอคติวิตี้ไดอะแกรมของการแนะนำลูกค้า.....48

4-8 แสดงซีเควนไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์ บริษัทออริซอฟท์.....50

4-9 แสดงซีเควนไดอะแกรมของการจัดเตรียมข้อมูลเพื่อนำไปใช้ในการวิเคราะห์.....51

4-10 แสดงซีเควนไดอะแกรมของการนำข้อมูลที่วิเคราะห์ได้ไปทำนายผล.....52

5-1 แสดงอีอาร์ไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์ บริษัทออริซอฟท์.....53

6-1 แสดงส่วนต่อประสานผู้ใช้ระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์ บริษัทออริซอฟท์ในส่วนของการใส่ข้อมูลการซื้อขายและข้อมูลลูกค้า.....58

6-2 แสดงส่วนต่อประสานผู้ใช้ระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์ บริษัทออริซอฟท์ในส่วนของการวิเคราะห์ข้อมูลและแสดงผลลัพธ์การทำนาย.....58

7-1 แสดงอัตราส่วนตาม Package ที่ลูกค้าเลือกซื้อ.....60

7-2 แสดงอัตราส่วนตาม Module ที่ลูกค้าเลือกซื้อ.....62

7-3 แสดงอัตราส่วนตามประเภทของซอฟต์แวร์ที่ลูกค้าเลือกซื้อ.....62

7-4 แสดงอัตราส่วนตามประเภทของลูกค้า.....63

7-5 แสดงความสัมพันธ์ทั้งหมดของซอฟต์แวร์ที่ลูกค้ามักเลือกซื้อพร้อมกัน.....64

7-6 แสดงผลการแบ่งกลุ่มของลูกค้าประเภทโรงงาน.....66

7-7 แสดงผลการทำนายความน่าจะเป็นที่จะขายได้ในปี 2015 โดยใช้เทคนิค K-NN.....67

7-8 แสดงผลการทำนายความน่าจะเป็นที่จะขายได้ในปี 2015 โดยใช้เทคนิค Naïve Bayes.....68

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ หน้า

7-9 แสดงผลการทำนายความน่าจะเป็นที่จะขายได้ในปี 2015 โดยใช้เทคนิค Decision Tree 69

7-10 แสดงผลลัพธ์ค่าความถูกต้องของความน่าจะเป็น โดยใช้เทคนิค K-NN 69

7-11 แสดงผลลัพธ์ค่าความถูกต้องของความน่าจะเป็น โดยใช้เทคนิค Naïve Bayes 70

7-12 แสดงผลลัพธ์ค่าความถูกต้องของความน่าจะเป็น โดยใช้เทคนิค Decision Tree 70

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาของระบบ

บริษัทออริซอฟท์ ก่อตั้งในปี พ.ศ.2530 โดยเป็นผู้ให้บริการด้านการจัดการทุนมนุษย์ มีสินค้าและบริการชั้นมากมายเช่น ระบบจัดการทรัพยากรบุคคล ระบบจัดการเวลาทำงาน ระบบจัดทำบัญชีเงินเดือนและโซลูชั่นให้พนักงานจัดการผ่านเวบสำหรับตลาดทั้งในประเทศและต่างประเทศ มีลูกค้าอยู่ในหลายภาคธุรกิจตั้งแต่ธุรกิจโรงแรม หน่วยงานรัฐบาล ก่อสร้าง วิศวกรรม โทรคมนาคม ค้าปลีก การขนส่งและการผลิต

การขายซอฟต์แวร์บริษัทออริซอฟท์นั้นจะแยกเป็น 3 ประเภทหลัก ได้แก่

1. ขายระบบซอฟต์แวร์โดยติดตั้งที่ไซต์ลูกค้า ในระบบซอฟต์แวร์นี้จะมีแบ่งแยกย่อยได้อีก 4 โมดูล ได้แก่

- HCM เป็นระบบจัดเก็บประวัติพนักงาน ประวัติการลา
- TMS เป็นระบบจัดเก็บและประมวลเวลาการทำงานของพนักงาน
- Payroll เป็นระบบเงินเดือนพนักงาน พร้อมทั้งพิมพ์รายงานต่างๆสำหรับนำส่ง
- ESS เป็นระบบ online ให้พนักงานสามารถทำรายการลาของตนเองได้

ซึ่งลูกค้าจะเลือกซื้อระบบที่มีความจำเป็นต่อการใช้งานภายในองค์กรและงบประมาณ

2.ขายค่าซ่อมบำรุง หลังจากที่ลูกค้าได้ทำการซื้อระบบซอฟต์แวร์ไปติดตั้งแล้ว เมื่อเวลาผ่านไป 1 ปี ทางบริษัทจะเรียกเก็บค่าบำรุงซอฟต์แวร์ ซึ่งรวมไปถึงการให้ความช่วยเหลือในเวลาที่ลูกค้ามีปัญหา แต่ทั้งนี้ถ้าไม่ต้องการซื้อค่าซ่อมบำรุง ลูกค้าจะยังคงใช้ระบบต่างๆต่อไปได้เพียงค่าอาจไม่ได้รับการดูแล

3.ขายบริการทำเงินเดือน คือ ลูกค้าไม่ต้องการทำการจ่ายเงินพนักงานด้วยตนเอง แต่จ้างให้ทางบริษัททำให้ จะมีแค่ระแบบทำเงินเดือนเท่านั้น

แต่ทั้งนี้ การขายซอฟต์แวร์และบริการต่างๆจะอ้างอิงตามจำนวนพนักงานของลูกค้า ถ้าลูกค้ามีจำนวนมากราคาที่ขายจะสูงกว่าบริษัทที่มีลูกค้าน้อย ทางบริษัทเองก็ได้จัดทำเป็นรูปแบบแพคเกจตามจำนวนพนักงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์

จากการศึกษาปัญหาการขายซอฟต์แวร์และการให้บริการลูกค้า ผู้จัดทำจึงได้ทำการวิเคราะห์และพัฒนาระบบสนับสนุนการตัดสินใจโดยมีวัตถุประสงค์ในการพัฒนา ดังนี้

1. เพื่อศึกษาระบบการขายซอฟต์แวร์บริษัทออริซอฟท์ (ประเทศไทย)
2. เพื่อวิเคราะห์ปัญหาระบบการขายซอฟต์แวร์บริษัทออริซอฟท์ (ประเทศไทย)
3. เพื่อพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

1.3 ขอบเขตของงาน

ระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์ พัฒนาขึ้นเพื่อปรับปรุงการจัดเก็บการขายซอฟต์แวร์และนำรายละเอียดการขายนั้นมาวิเคราะห์ ทำนายผล ขอบเขตในการพัฒนาระบบมีดังนี้

1. ระบบจัดเก็บข้อมูลขายซอฟต์แวร์บริษัทออริซอฟท์ (ประเทศไทย)
2. ระบบจัดการข้อมูล ลดความผิดพลาดของข้อมูล
3. ระบบวิเคราะห์การขายแสดงผลในรูปแบบของการทำนาย

1.4 แนวทางการศึกษาและพัฒนา

1. ศึกษาธุรกิจ การตลาด และการขาย ของบริษัทออริซอฟท์ (ประเทศไทย)
2. ศึกษาความต้องการของผู้บริหาร
3. วิเคราะห์และออกแบบระบบสนับสนุนการตัดสินใจ
4. ใช้ Visual Studio Express 2013 ในการออกแบบส่วนต่อประสานผู้ใช้
5. ใช้ภาษา C# ในการพัฒนาส่วนต่อประสานผู้ใช้
6. ใช้ Microsoft Excel เก็บข้อมูลที่อยู่ในรูปแบบไฟล์ .csv
7. ใช้ RapidMiner Studio 6.0 ในการกำจัดข้อมูลซ้ำซ้อนและข้อมูลที่ผิดพลาด
8. ใช้ RapidMiner Studio 6.0 ในการสร้างรูปแบบ Algorithm สำหรับวิเคราะห์ข้อมูล
9. ใช้ RapidMiner Studio 6.0 ในการสร้างผลการทำนาย
10. ทำการสรุปผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อวิเคราะห์ความต้องการของลูกค้าในอุตสาหกรรมที่แตกต่างกัน
2. เพื่อใช้เป็นแนวทางในการแนะนำซอฟต์แวร์ที่เหมาะสมให้แก่ลูกค้า
3. เพื่อใช้เป็นแนวทางในการเพิ่มยอดขาย
4. เพื่อลดค่าใช้จ่ายเรื่องเอกสารและสถานที่เก็บเอกสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและเทคโนโลยีที่เกี่ยวข้อง

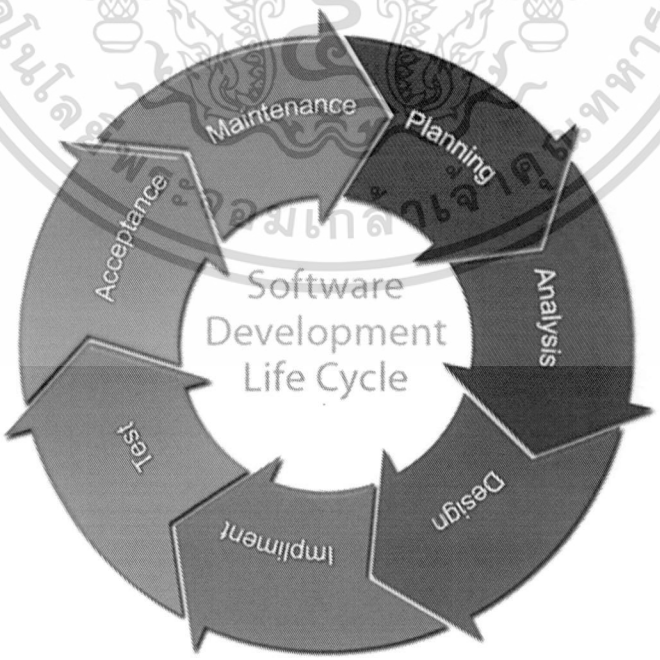
การออกแบบและพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์ บริษัทออริซอฟต์(ประเทศไทย) ได้ใช้แนวคิดของเดสก์ทอปแอพพลิเคชั่น มีการเก็บข้องมูลลงฐานข้อมูล และนำข้อมูลเหล่านั้นไปวิเคราะห์เพื่อใช้เป็นข้อมูลประกอบการตัดสินใจ โดยการออกแบบและพัฒนาระบบได้นำหลักการทฤษฎีพื้นฐานและเทคโนโลยีที่เกี่ยวข้อง ดังต่อไปนี้

2.1 กระบวนการที่เกี่ยวข้องในการพัฒนาระบบ

ในการออกแบบและพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์ บริษัทออริซอฟท์(ประเทศไทย) ได้ใช้กระบวนการในการพัฒนาระบบ ดังนี้

2.1.1 วงจรการพัฒนาระบบ (SDLC : System Development Life Cycle)

ในการออกแบบและพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์จำเป็นต้องต้องมีมาตรฐานในการพัฒนาระบบ เพื่อให้ได้ซึ่งระบบที่มีประสิทธิภาพมากที่สุด ดังนั้นจึงได้มีการนำวงจรการพัฒนาระบบงานขึ้นเพื่อตอบสนองความต้องการของระบบ ซึ่งในส่วนนี้ได้ทำการแสดงถึงรายละเอียดของกิจกรรมต่างๆของวงจรการพัฒนาระบบงานได้ดังนี้



รูปที่ 2-1 แสดงวงจรการทำงานของ SDLC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- วางแผนโครงการ (Project Planning)

การวางแผนโครงงานเป็นกระบวนการขั้นแรก โดยจะคำนึงประโยชน์ที่ได้จากระบบสารสนเทศ มีการกำหนดทีมงานที่จะสร้างระบบและศึกษาความเป็นไปได้ ผู้บริหารและคณะกรรมการเป็นผู้สนับสนุนและให้การรับรองวิเคราะห์ระบบ (Analysis)

- การวิเคราะห์ระบบ (Analysis)

จากขั้นวางแผนโครงงานและได้รับการสนับสนุนจากผู้บริหารและคณะกรรมการ กระบวนการต่อมาคือการวิเคราะห์ระบบเพื่อศึกษาความต้องการขององค์หรือหน่วยงาน โดยรวบรวมข้อมูลจากการทำงานด้วยวิธี การสังเกตุ การสัมภาษณ์ การทำแบบสอบถาม และนำข้อมูลที่ได้มาวิเคราะห์เพื่อกำหนดเป้าหมายแนวทางที่ต้องการเพื่อสร้างระบบใหม่ออกแบบระบบ (Design)

- การออกแบบระบบ (Design)

การออกแบบระบบคือการนำข้อมูลที่วิเคราะห์มาออกแบบเพื่อให้ได้ระบบงานตามที่ต้องการและสามารถเปลี่ยนแปลงได้เมื่อความต้องการระบบหลักมีการเปลี่ยนแปลง นักวิเคราะห์จะออกแบบระบบโดยการสร้างแบบจำลองกระบวนการไหลของข้อมูล (Data flow Diagram : DFD) สร้างแบบจำลองข้อมูล (Entity Relationship Diagram : ERD) สร้างสถาปัตยกรรมของระบบ (Architecture Design) การสร้างต้นแบบ เป็นต้น ซึ่งสิ่งเหล่านี้ ทางผู้พัฒนาระบบจะนำไปพัฒนาในกระบวนการต่อไป

- พัฒนาและติดตั้งระบบ (Implementation)

เป็นขั้นตอนของการพัฒนาโปรแกรมให้ระบบทำงานได้ตามผลการวิเคราะห์และออกแบบระบบจะประกอบด้วยการทดสอบแบบเพื่อให้ได้ส่วนของโปรแกรมที่พร้อมติดตั้งที่ทำงานได้ตามความต้องการของระบบรวมถึงผู้ใช้ระบบต้องมีความรู้ได้รับการฝึกอบรมในการใช้งานระบบสารสนเทศ เมื่อการใช้งานโปรแกรมได้ระยะหนึ่งต้องมีการประเมินผลการทำงานของโปรแกรมว่าเป็นไปตามเป้าหมายที่กำหนดไว้หรือไม่เพียงใดและผู้ใช้มีความพึงพอใจมากน้อยเพียงใด

- การทดสอบระบบ (Testing)

การทดสอบระบบบางครั้งอาจอยู่ในขั้นพัฒนาและติดตั้งระบบ เป็นการตรวจสอบความถูกต้องของระบบงานที่สร้างขึ้นมาว่าตรงกับความต้องการจริงหรือไม่

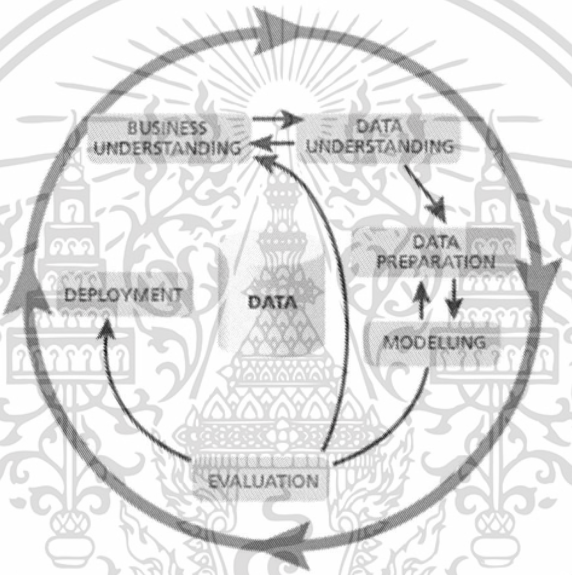
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- การบำรุงรักษาระบบ (Maintenance)

หลังจากที่ระบบได้ติดตั้งและใช้งานได้ในระยะหนึ่ง จำเป็นต้องมีการบำรุงรักษาระบบอย่างสม่ำเสมอ โดยความคาดหวังของบริษัทต้องการให้ระบบใช้งานได้ยาวนานและสามารถรองรับเทคโนโลยีใหม่ๆในอนาคตได้

2.1.2 กระบวนการ CRISP-DM

ในการออกแบบและพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์ได้นำกระบวนการ CRISP-DM มาใช้ในส่วนของการวิเคราะห์ข้อมูลการซื้อขาย CRISP-DM ย่อมาจาก Cross-Industry Standard Process for Data Mining ประกอบด้วย 6 ขั้นตอน ดังนี้



รูปที่ 2-2 แสดงกระบวนการ CRISP-DM

- Business Understanding

เป็นขั้นตอนแรกในกระบวนการ CRISP-DM เป็นกระบวนการทำความเข้าใจ ระบุปัญหาหรือโอกาสเชิงธุรกิจ จากนั้นกำหนดรูปแบบที่เหมะสมต่อการนำมาวิเคราะห์ข้อมูล

- Data Understanding

เนื่องจากข้อมูลเป็นปัจจัยที่สำคัญที่สุดที่ขาดไม่ได้ ในขั้นตอนนี้เป็นการรวบรวมข้อมูลที่เกี่ยวข้อง เพื่อใช้ในการวิเคราะห์ ในการรวบรวมข้อมูลนั้นควรพิจารณาด้วยว่าเป็นข้อมูลที่ได้มาจากแหล่งข้อมูลที่ถูกต้องน่าเชื่อถือ ข้อมูลที่ได้มีปริมาณมากพอ และเป็นข้อมูลที่เหมาะสม มีรายละเอียดเพียงพอต่อการนำไปใช้ในการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Data Preparation

ขั้นตอนการเตรียมข้อมูลเป็นขั้นตอนที่ใช้เวลานานที่สุด เนื่องจากโมเดลจะให้ผลลัพธ์ที่ถูกต้องหรือไม่นั้น ขึ้นอยู่กับคุณภาพของข้อมูลที่ใช้ กล่าวคือถ้าข้อมูลที่ใช้นั้นไม่ถูกต้อง มีผิดพลาด ย่อมสะท้อนถึงผลลัพธ์ที่ได้ ซึ่งอาจทำให้ตีความผลลัพธ์ได้คลาดเคลื่อนเช่นกัน โดยการเตรียมข้อมูลนั้น สามารถแบ่งออกได้เป็น 3 ขั้นตอนย่อยคือ

1. ทำการคัดเลือกข้อมูล (Data Selection) ควรที่จะต้องกำหนดเป้าหมายก่อนว่าเราจะทำการวิเคราะห์อะไร แล้วจึงเลือกใช้เฉพาะข้อมูลที่เกี่ยวข้องกับสิ่งที่เราจะทำการวิเคราะห์
2. การกลั่นกรองข้อมูล (Data Cleaning) ในบางกรณีอาจพบข้อมูลที่ไม่ถูกต้องอันเนื่องมาจากปัญหาในระหว่างการจัดเก็บข้อมูล เช่นการกรอกข้อมูลไม่ครบบ้าง
   กรอกข้อมูลซ้ำซ้อนบ้าง ในขั้นตอนนี้เราจะทำการกรองข้อมูลที่ไม่ถูกต้องหรือซ้ำซ้อนออก หรืออาจทำการซ่อมข้อมูลที่ขาดหายไปด้วยวิธีการบางอย่าง เช่นการพิจารณาจากค่าเฉลี่ยของข้อมูลส่วนใหญ่เป็นต้น
3. การแปลงรูปข้อมูล (Data Transformation) เป็นขั้นตอนการเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมนำไปใช้ในการวิเคราะห์ตามอัลกอริทึมของดาต้าไมน์นิ่งที่เลือกใช้

- Modeling

เป็นขั้นตอนการวิเคราะห์ข้อมูลใช้เทคนิคดาต้าไมน์นิ่ง เช่นการสร้างตัวทำนาย (prediction model) เพื่อให้ได้ผลลัพธ์ที่ดีที่สุด แต่ถ้ามีความผิดพลาดก็ต้องย้อนกลับไปที่ขั้นตอนการเตรียมข้อมูล เพื่อแปลงข้อมูลที่ได้มาให้เหมาะสมกับเทคนิค ขั้นตอนนี้ต้องประเมินโมเดล ข้อมูลที่ได้ให้มีความถูกต้อง และนำไปสู่ตัวบ่งชี้ที่น่าเชื่อถือของโมเดลได้

- Evaluation

การประเมินหรือวัดประสิทธิภาพในขั้นตอนModeling เป็นเพียงการวัดความน่าเชื่อถือของโมเดล แต่ขั้นตอนนี้เป็นการประเมินประสิทธิภาพของผลลัพธ์จากโมเดลวิเคราะห์ข้อมูลว่าครอบคลุมและสามารถตอบโจทย์ทางธุรกิจที่ตั้งไว้ในขั้นตอนแรกหรือไม่ ถ้ามีการสร้างโมเดลวิเคราะห์ข้อมูลหลายโมเดล ในขั้นตอนนี้ต้องทำการประเมินจุดเด่นจุดด้วยของแต่ละโมเดลด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเลือกใช้โมเดลที่เหมาะสม ต้องมีทักษะในการวิเคราะห์ข้อมูลและธุรกิจ เพื่อลดเวลาและเพิ่มความสะดวกรวดเร็วจึงมีการใช้เครื่องทางด้านกราฟฟิคมาช่วยในขั้นตอนนี้เช่น การแสดงผลการวิเคราะห์ด้วยกราฟ การรายงานผลรูปแบบต่างๆ

- Deployment

ผลลัพทธ์หรือองค์ความรู้ที่ได้จากการวิเคราะห์จะไม่มีประโยชน์เลย ถ้าไม่ถูกนำไปใช้งานจริง ตัวอย่างเช่น การนำองค์ความรู้ที่ได้ไปใช้ในการจัดโปรโมชั่นส่งเสริมการขายสินค้า ใช้ในการทำนายแนวโน้มการทุจริตในระบบการเงินของธนาคาร หรือตรวจจับความผิดปกติในการซื้อขายหุ้นในตลาดหลักทรัพย์เป็นต้น ในงานศึกษาอิสระนี้จะนำผลลัพธ์ของการวิเคราะห์ข้อมูลไปใช้ในการแนะนำซอฟต์แวร์ที่เหมาะสนให้แก่ลูกค้า เป็นการเพิ่มยอดขายให้กับบริษัทด้วย

2.2 การออกแบบระบบโดยใช้ยูเอ็มแอล

ยูเอ็มแอลเป็นการนำเสนอแผนภาพ จะใช้อธิบายและนำเสนอแนวความคิดให้เห็นภาพรวมการทำงานของระบบไปในทิศทางเดียวกัน แผนภาพยูเอ็มแอลมีทั้งหมดหลายรูปแบบ แต่ที่ใช้ในรายงานฉบับนี้จะมีทั้งหมด 4 แบบ แบ่งออกเป็น 2 กลุ่ม ดังนี้

1. กลุ่มแผนภาพแสดงโครงสร้าง

1.1 ยูสเคสไดอะแกรม ใช้แสดงการทำงานหลักของระบบ โดยมีองค์ประกอบดังนี้

- แอกเตอร์ ใช้แทนผู้ติดต่อกับระบบ เช่น คน หรือระบบ
- ยูสเคส คือ การทำงานต่างๆของระบบ หรือสิ่งที่ระบบต้องทำจากแอกเตอร์
- ความสัมพันธ์ คือ เส้นแสดงความสัมพันธ์ระหว่างแอกเตอร์กับยูสเคส ยูสเคสกับยูสเคส หรือแอกเตอร์กับแอกเตอร์

1.2 คลาสไดอะแกรม ใช้แสดงคลาส และความสัมพันธ์ของคลาส มีส่วนประกอบสำคัญ 3 ส่วน คือ ชื่อของคลาส คุณลักษณะของคลาส และการทำงานของคลาส

2. กลุ่มแผนภาพแสดงพฤติกรรม

2.1 ซีเควนซ์ไดอะแกรม เป็นแผนภาพที่ใช้อธิบายการทำงานของยูสเคส เพื่อแสดงถึงขั้นตอนการทำงานและลำดับของการสื่อสารระหว่างอ็อบเจกต์ที่ตอบโต้กัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 แอกทีวิตีไดอะแกรม ใช้แสดงลำดับของกิจกรรมการทำงานของแต่ละขั้นตอนในระบบ จะเป็นตัวอธิบายขั้นตอนการทำงานของระบบ และแบ่งหน้าที่ความรับผิดชอบในแต่ละกิจกรรมด้วยช่องสวิมเลน

2.3 การออกแบบฐานข้อมูล

ระบบจัดการการขายสินเชื่อรายย่อยจะมีการเก็บข้อมูลต่างๆของระบบลงฐานข้อมูล เช่น ข้อมูลลูกค้า ข้อมูลรถยนต์ ข้อมูลประกันภัย จึงต้องมีการออกแบบการเก็บข้อมูล และมีการปรับแก้ไขฐานข้อมูลตามหลักการนอร์มัลไลเซชันระดับ 3 เป็นอย่างน้อย โดยการออกแบบฐานข้อมูลของระบบจะมีแบบจำลองที่เกี่ยวข้องดังนี้

1. แผนภาพอีอาร์

ใช้แสดงแบบจำลองข้อมูลของระบบ มีรูปแบบการเขียนที่นิยมอยู่ 2 แบบ คือแบบ Chen Model และแบบ Crow's Foot ซึ่งแบบจำลองของโครงการนี้จะเลือกใช้แบบ Crow's Foot โดยแผนภาพอีอาร์จะมีองค์ประกอบดังนี้

- เอนทิตี คือ บุคคล สถานที่ สิ่งของ หรือเหตุการณ์ที่สนใจจัดเก็บภายใต้ขอบเขตของระบบ เช่น พนักงาน ลูกค้า รถยนต์ เป็นต้น
- แอตทริบิวต์ คือ สิ่งที่อธิบายคุณลักษณะของเอนทิตี เช่น เอนทิตีของพนักงาน ประกอบไปด้วยแอตทริบิวต์ รหัสพนักงาน ชื่อพนักงาน ตำแหน่งพนักงาน เป็นต้น
- ความสัมพันธ์ คือ ความเกี่ยวข้องกันระหว่างเอนทิตีต่างๆในระบบ ประกอบไปด้วยประเภทความสัมพันธ์ เช่น แบบหนึ่งต่อหนึ่ง แบบหนึ่งต่อกลุ่ม และแบบกลุ่มต่อกลุ่ม คาร์ดินัลลิตี คือ จำนวนของเอนทิตีที่สามารถมีความสัมพันธ์กับอีกเอนทิตี และการมีส่วนร่วมในความสัมพันธ์ เช่น แบบเลือกได้ และแบบบังคับ

2. พจนานุกรมข้อมูล

ใช้อธิบายรายละเอียดของเอนทิตี และสร้างความเข้าใจระหว่างผู้ออกแบบและผู้พัฒนาฐานข้อมูล ในพจนานุกรมข้อมูลจะมีองค์ประกอบ เช่น ชื่อเอนทิตี ชื่อแอตทริบิวต์ คำอธิบายของแอตทริบิวต์ ชนิดของแอตทริบิวต์ รูปแบบของแอตทริบิวต์ เป็นต้น

2.4 ทฤษฎีที่เกี่ยวข้องในการพัฒนาระบบ

ทฤษฎีที่ใช้ในการพัฒนาระบบ มีดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.1 การจำแนกประเภทข้อมูล (Data Classification)

การจำแนกประเภทข้อมูลคือกระบวนการสร้างโมเดลจำแนกประเภทข้อมูล (Data Classification Model) เพื่อทำนายกลุ่มของข้อมูลใหม่ (Unseen data) ตัวอย่างของกลุ่มเช่น กลุ่มของลูกค้าที่ซื้อคอมพิวเตอร์-ไม่ซื้อคอมพิวเตอร์ กลุ่มของลูกค้าที่ฐานะดี-ปานกลาง-แย่ กลุ่มของการผลิตสินค้า ผ่านเกณฑ์-ไม่ผ่านเกณฑ์ ในที่นี้คำว่ากลุ่มจะเรียกว่า class ของข้อมูล ซึ่งใน class เดียวกันนั้นจะต้องมีข้อมูลที่มีความเหมือนหรือคล้ายคลึงกันมากกว่าข้อมูลที่อยู่ใน class ที่แตกต่างกัน

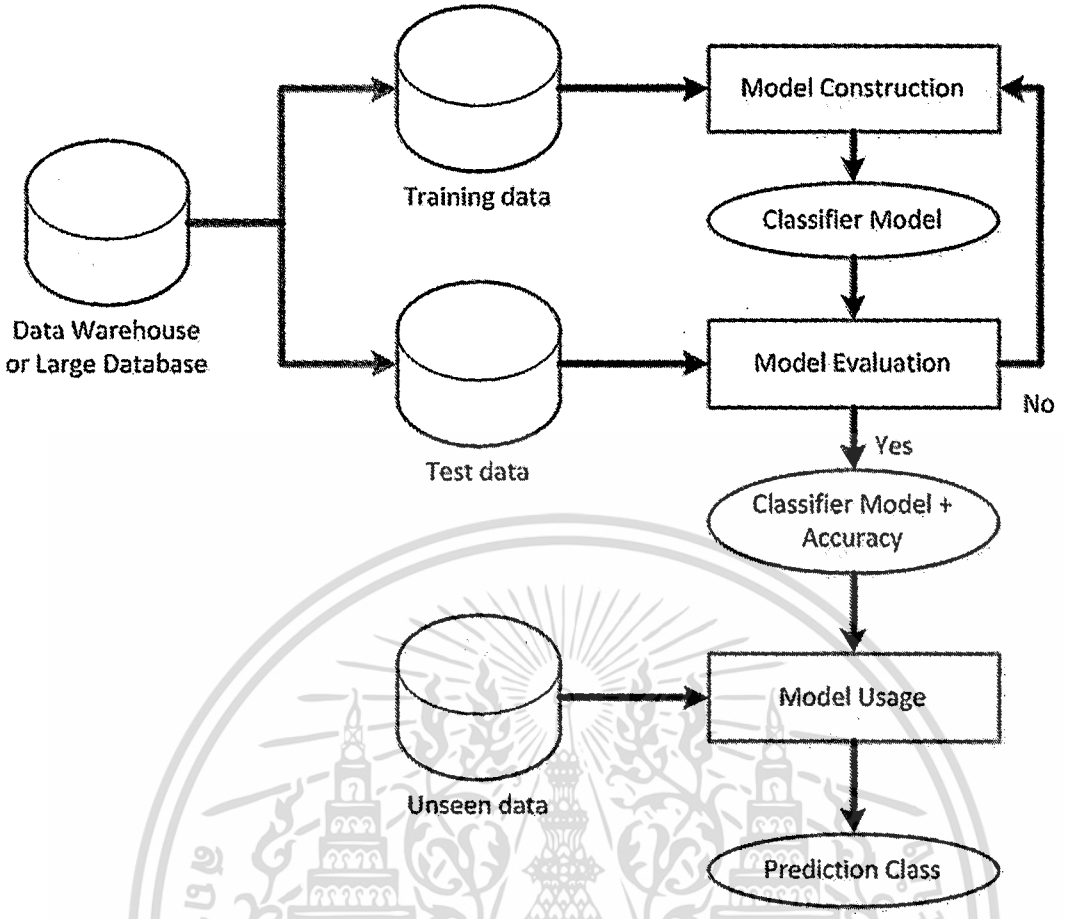
การสร้างโมเดลจำแนกประเภทข้อมูล จะเกิดขึ้นมาจากการหาความสัมพันธ์ของข้อมูลในฐานข้อมูลขนาดใหญ่ โดยข้อมูลทั้งหมดจะมีการแบ่งออกเป็น 2 กลุ่มคือกลุ่มข้อมูลเรียนรู้ (Training set) เป็นชุดข้อมูลที่มีบทบาทในการสร้างโมเดลจำแนกประเภทข้อมูลขึ้นมา และมีกลุ่มข้อมูลทดสอบ (Test set) เป็นชุดข้อมูลประเมินความถูกต้องของโมเดลจำแนกประเภทข้อมูล

โมเดลจำแนกประเภทข้อมูลได้ถูกนำมาประยุกต์ใช้งานหลายๆ ด้าน ไม่ว่าจะเป็นการวิเคราะห์หุ้น เพื่อหาว่าหุ้นแต่ละบริษัทมีคุณภาพเป็นอย่างไร เมื่อมีปัจจัยที่เกี่ยวข้อง ไม่ว่าจะเป็นการเติบโตของรายได้ ความสามารถในการควบคุมต้นทุน ความผันผวนของรายได้และกำไร และผู้บริหาร หรือจะเป็นการพยากรณ์อากาศ การจัดสรรกฎหมายที่เหมาะสมในการพิจารณาคดีความ การจัดการความสัมพันธ์ของลูกค้า (Customer relationship management) และอื่นๆ

<u>กระบวนการสร้างตัวโมเดลจำแนกประเภทข้อมูล</u>

แบ่งออกเป็น 3 ขั้นตอน ซึ่งภาพรวมของกระบวนการสร้างโมเดลจำแนกประเภทข้อมูลแสดงได้ดังรูป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2-4 แสดงภาพจำลองกระบวนการสร้างตัวโมเดลจำแนกประเภท

กระบวนการของแต่ละขั้นตอนมีดังนี้

1. Model Construction (Learning) เป็นขั้นตอนการสร้างโมเดลจำแนกประเภท โดยอาศัยการเรียนรู้จากข้อมูลที่ได้กำหนด class ไว้เรียบร้อยแล้วหรือเรียกว่าข้อมูลเรียนรู้ (Training data) ซึ่งโมเดลจำแนกประเภทที่ได้จะแสดงด้วยวิธีการพื้นฐานทางเหมืองข้อมูล (Data mining) ยกตัวอย่างเช่น ต้นไม้ตัดสินใจ (Decision Tree) โมเดลจำแนกประเภทที่ได้จะมีลักษณะคล้ายต้นไม้จริงกลับหัวที่มีโหนดรากอยู่ด้านบนสุดและโหนดใบอยู่ล่างสุดของต้นไม้ แต่ละโหนดบนต้นไม้จะมีคุณลักษณะ (attribute) เป็นตัวเลือกทดสอบ ซึ่งจะมีกิ่งซึ่งเป็นค่าที่เป็นไปได้ของคุณลักษณะ (attribute value) ที่ถูกเลือกทดสอบไว้ และมีโหนดใบแสดง class ที่กำหนดไว้

2. Model Evaluation (Accuracy) เป็นขั้นตอนตรวจสอบความถูกต้อง โดยอาศัยข้อมูลที่ใช้สำหรับทดสอบเรียกว่าข้อมูลทดสอบ (Testing data) ซึ่งกลุ่มที่แท้จริงของข้อมูลที่ใช้ทดสอบจะถูกนำมาเปรียบเทียบกับกลุ่มที่หามาได้จากโมเดลจำแนกประเภท เพื่อทดสอบว่าโมเดลจำแนกประเภทนี้สามารถจัดกลุ่มประเภทข้อมูลได้อย่างถูกต้องมากน้อยเพียงใด และมีการปรับปรุงโมเดลจำแนกประเภทจนกว่าจะได้ค่าความถูกต้องในระดับที่ยอมรับได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. Model Usage (Classification) เป็นขั้นตอนการนำโมเดลจำแนกประเภทที่สร้างขึ้นมาใช้กับข้อมูลที่ไม่เคยเห็นมาก่อน (unseen data) เพื่อทำนายและกำหนดกลุ่มให้กับข้อมูลนั้น

2.4.2 การจำแนกประเภทข้อมูลด้วยวิธี Naive Bayes

การจำแนกประเภทข้อมูลด้วยวิธี Naive Bayes นั้น ถือว่าเป็นวิธีที่ได้รับความนิยมเป็นอย่างมาก เนื่องจากมีโมเดลที่สร้างง่ายและไม่ซับซ้อน โดยอาศัยทฤษฎีความน่าจะเป็น (probability) เป็นหลักสมการของการหาความน่าจะเป็น คือ

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{2.1}$$

โดยที่

- P(B|A) คือ ค่า conditional probability หรือค่าความน่าจะเป็นที่เกิดเหตุการณ์ B ขึ้นก่อนและจะมีเหตุการณ์ A ตามมา
- P(A∩B) คือ ค่า joint probability หรือค่าความน่าจะเป็นที่เหตุการณ์ A และเหตุการณ์ B เกิดขึ้นร่วมกัน
- P(A) คือ ค่าความน่าจะเป็นที่เหตุการณ์ A เกิดขึ้น

ตัวอย่างของการทำนายความน่าจะเป็นที่สามารถเข้าใจได้ง่าย เช่น การทำนายสภาพจราจรเมื่อเกิดเหตุการณ์ฝนตก จะสามารถทำนายสภาพจราจรได้ดังสมการที่นำเหตุการณ์ไปแทนค่าในสมการข้างต้นได้ดังต่อไปนี้

$$P(\text{สภาพจราจร} \mid \text{ฝนตก}) = P(\text{ฝนตก} \mid \text{สภาพจราจร}) \times P(\text{สภาพจราจร})/P(\text{ฝนตก})$$

จากตัวอย่างนี้ เราสามารถทำนายสภาพจราจรโดยสังเกตุฝนตก อย่างไรก็ตามเหตุการณ์ที่นำมาใช้ในการทำนายนั้นต้องสอดคล้องกับเหตุการณ์ที่จะทำนาย เช่น ถ้าหากเราต้องการทำนายสภาพจราจร เราก็จะไม่ใช้เหตุการณ์ไฟดับมาพิจารณา เพราะเหตุการณ์ทั้งสองไม่มีความเกี่ยวข้องกัน

ตัวอย่างการทำนายความน่าจะเป็นการตีเทนนิส จากเหตุการณ์ 4 เหตุการณ์ สภาพอากาศ (outlook), อุณหภูมิ (temperature), ความชื้น (humidity), และ แรงลม (windy) ดังตารางต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Outlook	Temperature	Humidity	Windy	Play Tennis
sunny	hot	high	TRUE	no
overcast	hot	high	FALSE	yes
rainy	mild	high	FALSE	yes
rainy	cold	normal	FALSE	yes
rainy	cold	normal	TRUE	no
overcast	cold	normal	TRUE	yes
sunny	mild	high	FALSE	no
sunny	cold	normal	FALSE	yes
rainy	mild	normal	FALSE	yes
sunny	mild	normal	TRUE	yes
overcast	mild	high	TRUE	yes
overcast	hot	normal	FALSE	yes
rainy	mild	high	TRUE	no

รูปที่ 2-5 ตารางแสดงสภาพอากาศของตัวอย่าง Naive Bayes

สมมติในกรณีที่มีข้อมูลใหม่ที่เราต้องการทำนายการเล่นเทนนิส สภาพอากศ = overcast, อุณภูมิ = cold, ความชื้น = normal, และ แรงลม = true จะสามารถทำนายการเล่นเทนนิส ดังสมการดังต่อไปนี้

P(play tennis = yes | outlook = overcast, temperature = cold, humidity = normal, windy = true) = P(outlook = overcast | play tennis=yes) x P(temperature = cold | play tennis=yes) x P(humidity = normal | play tennis=yes) x P(windy = true | play tennis=yes) x P(play tennis=yes)

P(play tennis = yes | outlook = overcast, temperature = cold, humidity = normal, windy = true) = (0.444 x 0.222 x 0.666 x 0.333 x 0.692) = 0.015

และสามารถทำนายการไม่เล่นเทนนิส ดังสมการดังต่อไปนี้

P(play tennis = no | outlook = overcast, temperature = cold, humidity = normal, windy = true) = P(outlook = overcast | play tennis=no) x P(temperature = cold | play tennis=no) x P(humidity = normal | play tennis=yes) x P(windy = true | play tennis=no) x P(play tennis=no)

P(play tennis = no | outlook = overcast, temperature = cold, humidity = normal, windy = true) = (0 x 0.250 x 0.250 x 0.750 x 0.307) = 0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นในกรณี สภาพอากศ = overcast, อุณภูมิ = cold, ความชื้น = normal, และ แรงลม = true จะสามารถทำนายการเล่นเทนนิส = yes

2.4.3 การจำแนกประเภทด้วยวิธี K-Nearest Neighbors

การจำแนกประเภทด้วยวิธี K-Nearest Neighbors คือ วิธีการในการแบ่งกลุ่มของข้อมูล เทคนิคนี้จะตัดสินใจ ว่าคลาสใดที่จะแทนเงื่อนไขหรือกรณีใหม่ๆได้บ้าง ด้วยการตรวจสอบจำนวนบางจำนวน ("K" ใน K-nearest neighbor) ของเงื่อนไขที่เหมือนกันหรือใกล้เคียงกันมากที่สุด โดยจะหาผลรวม (Count Up) ของจำนวนเงื่อนไข โดยจะทำการวัดระยะห่างระหว่างข้อมูลที่ต้องการทำนายกับข้อมูลที่อยู่ใกล้เคียง ในเทคนิคนี้จะใช้วิธีการวัดระยะห่างแบบ Euclidean ดังสมการ

$$D_{Euclidean} = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\ldots+(x_L-y_L)^2} \quad (2.2)$$

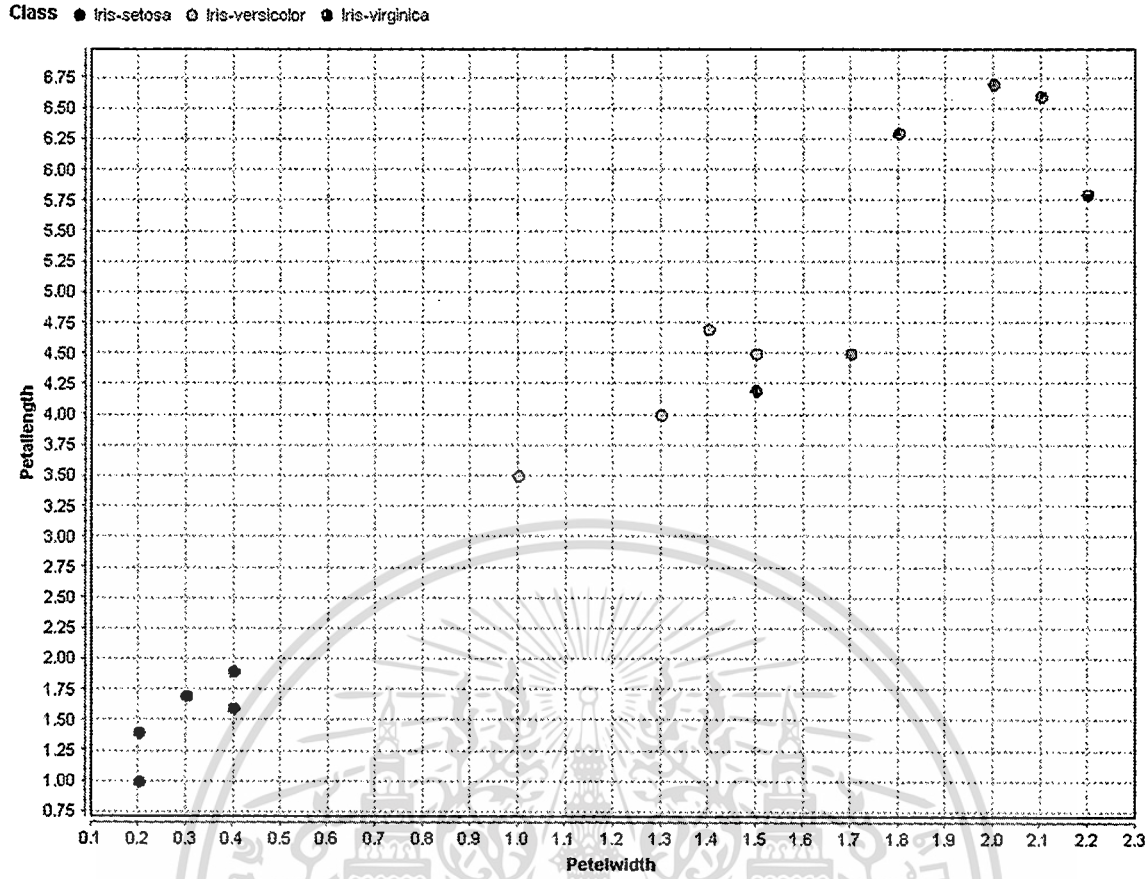
โดยที่ x_1 คือ แอตทริบิวต์ที่ 1 ของข้อมูลจุดที่ 1 และ y_1 คือ แอตทริบิวต์ที่ 2 โดยข้อมูลทั้งสองตัวมีจำนวนแอตทริบิวต์เท่ากับ L

ตัวอย่างด้านล่างจะแสดงวิธีการทำนายข้อมูลโดยใช้วิธี K-NN

ID	Petallength	Petelwidth	Class
1	1.40	0.20	Iris-setosa
2	1.70	0.30	Iris-setosa
3	1.60	0.40	Iris-setosa
4	1.90	0.40	Iris-setosa
5	1.00	0.20	Iris-setosa
6	4.70	1.40	Iris-versicolor
7	4.50	1.50	Iris-versicolor
8	4.00	1.30	Iris-versicolor
9	3.50	1.00	Iris-versicolor
10	4.20	1.50	Iris-virginica
11	5.80	2.20	Iris-virginica
12	6.60	2.10	Iris-virginica
13	4.50	1.70	Iris-virginica
14	6.30	1.80	Iris-virginica
15	6.70	2.00	Iris-virginica

รูปที่ 2-6 ตารางแสดงประเภทของ Iris

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2-7 แสดงการกระจายตัวของข้อมูลจากตารางแสดงประเภท Iris

จากรูป ถ้าสมมติว่ามีข้อมูลใหม่ขึ้นมา แต่ไม่ทราบว่าจะเป็น Iris ประเภทไหน

ID	Petallength	Petelwidth	Class
16	4.25	1.60	?

จะสามารถใช้ K-NN เพื่อทำนายว่าข้อมูลชุดนี้เป็น Iris แบบไหนได้โดย

1. กำหนดว่า K=3 คือการระบุข้อมูลให้พิจารณาเพื่อนบ้านที่ใกล้ที่สุดจำนวน 3 ตัวอย่าง
2. คำนวณระยะห่างระหว่างข้อมูลทั้งหมด กับข้อมูลใหม่ ดังนี้

$$D_{1,16}=\sqrt{(1.40-4.25)^2+(0.20-1.60)^2}=3.18$$

$$D_{2,16}=\sqrt{(1.70-4.25)^2+(0.30-1.60)^2}=2.86$$

$$D_{3,16}=\sqrt{(1.60-4.25)^2+(0.40-1.60)^2}=2.91$$

$$D_{4,16}=\sqrt{(1.90-4.25)^2+(0.40-1.60)^2}=2.64$$

$$D_{5,16}=\sqrt{(1.00-4.25)^2+(0.20-1.60)^2}=3.54$$

$$D_{6,16}=\sqrt{(4.70-4.25)^2+(1.40-1.60)^2}=0.49$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$D_{7,16}=\sqrt{(4.50-4.25)^2+(1.50-1.60)^2}=0.27$$
$$D_{8,16}=\sqrt{(4.00-4.25)^2+(1.30-1.60)^2}=0.39$$
$$D_{9,16}=\sqrt{(3.50-4.25)^2+(1.00-1.60)^2}=0.96$$
$$D_{10,16}=\sqrt{(4.20-4.25)^2+(1.50-1.60)^2}=0.11$$
$$D_{11,16}=\sqrt{(5.80-4.25)^2+(2.20-1.60)^2}=1.66$$
$$D_{12,16}=\sqrt{(6.60-4.25)^2+(2.10-1.60)^2}=2.40$$
$$D_{13,16}=\sqrt{(4.50-4.25)^2+(1.70-1.60)^2}=0.27$$
$$D_{14,16}=\sqrt{(6.30-4.25)^2+(1.80-1.60)^2}=2.06$$
$$D_{15,16}=\sqrt{(6.70-4.25)^2+(2.00-1.60)^2}=0.11$$

จากการคำนวณระยะห่างระหว่างข้อมูลต่างๆกับข้อมูลใหม่ลำดับที่ 16 สามารถเรียงลำดับตามระยะห่างจากค่าที่น้อยที่สุดไปค่าที่มากที่สุด และจะพบว่าข้อมูลที่มีระยะห่างน้อยที่สุด 3อันดับแรก คือ

ข้อมูลลำดับที่ 10 มีระยะห่าง = 0.11 และคลาสเป็น Iris-versicolor

ข้อมูลลำดับที่ 15 มีระยะห่าง = 0.11 และคลาสเป็น Iris-verginica

ข้อมูลลำดับที่ 7 มีระยะห่าง = 0.27 และคลาสเป็น Iris-versicolor

จึงสามารถสรุปได้ว่า ข้อมูลใหม่ลำดับที่ 16 นั้น อยู่ในคลาส Iris-versicolor เนื่องจากข้อมูล 2 ใน 3 เป็นคลาส Iris-versicolor

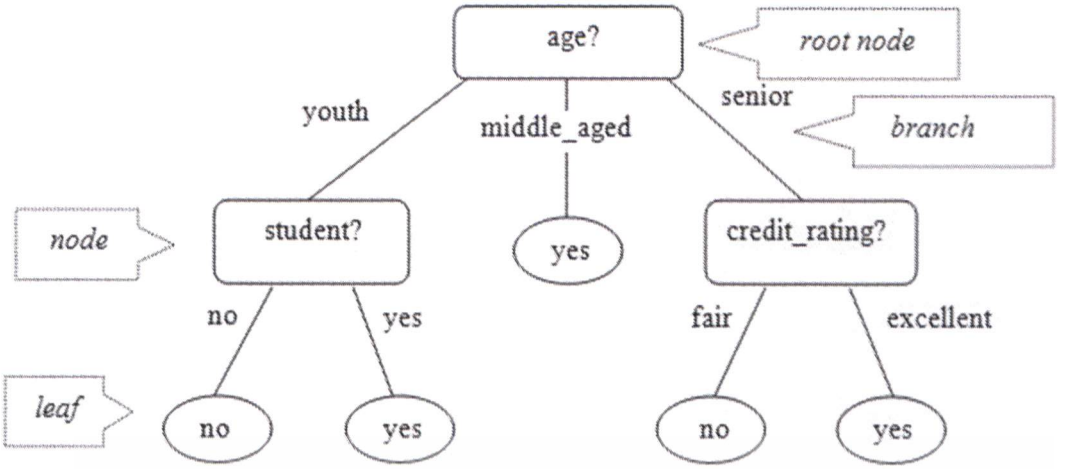
2.4.4 การจำแนกประเภทด้วยวิธี Decision Tree

ต้นไม้ตัดสินใจ (Decision Tree) เป็นโครงสร้างข้อมูลชนิดเป็นลำดับชั้น (hierarchy) ใช้สนับสนุนการตัดสินใจ โดยจะมีลักษณะคล้ายต้นไม้จริงกลับหัวที่มีโหนดรากอยู่ด้านบนสุดและโหนดใบอยู่ล่างสุดของต้นไม้

ภายในต้นไม้จะประกอบไปด้วยโหนด (node) ซึ่งแต่ละโหนดจะมีคุณลักษณะ (attribute) เป็นตัวทดสอบ กิ่งของต้นไม้ (branch) แสดงถึงค่าที่เป็นไปได้ของคุณลักษณะที่ถูกเลือกทดสอบ และใบ (leaf) ซึ่งเป็นสิ่งที่อยู่ล่างสุดของต้นไม้ตัดสินใจแสดงถึงกลุ่มของข้อมูล (class) หรือนั่นก็คือผลลัพธ์ที่ได้จากการทำนาย โหนดที่อยู่บนสุดของต้นไม้เรียกว่าโหนดราก (root node) ดังแสดงโครงสร้างของต้นไม้ตัดสินใจตัดสินใจดังรูปด้านล่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง



รูปที่ 2-8 แสดงโครงสร้างของต้นไม้ตัดสินใจ

จากรูปเป็นต้นไม้ที่ใช้ในการตัดสินใจว่าจะเลือกซื้อคอมพิวเตอร์หรือไม่ (Quinlan, 1986) มีคุณลักษณะที่พิจารณาคืออายุ (age) นักศึกษา (student) และอัตราเครดิต (credit_rating)

โดยที่โหนดสี่เหลี่ยมมุมโค้งจะเป็นการทดสอบคุณลักษณะของข้อมูล ท้ายสุดจะได้ผลลัพธ์ของการทำนายว่าจะซื้อคอมพิวเตอร์ (yes) หรือไม่ซื้อคอมพิวเตอร์ (no) จากการทดสอบตามเส้นทางของต้นไม้ตัดสินใจตั้งแต่โหนดรากไปจนถึงใบ แสดงดังตัวอย่าง

RID	*age*	*income*	*student*	*credit_rating*	Class: *buys_computer*
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	middle_aged	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	middle_aged	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
144195

RID	*age*	*income*	*student*	*credit_rating*	Class: *buys_computer*
12	middle_aged	medium	no	excellent	yes
13	middle_aged	high	yes	fair	yes
14	senior	medium	no	excellent	no

รูปที่ 2-9 ตารางแสดงการตัดสินใจในการซื้อคอมพิวเตอร์

เนื่องจากข้อมูลที่มีนั้นประกอบด้วยข้อมูลหลากหลายกลุ่มปะปนกัน ฉะนั้นจะต้องวัดมาตรฐานเกน (gain) ของแต่ละคุณลักษณะ (attribute) ซึ่งค่ามาตรฐานเกนนี้คำนวณได้โดยใช้ความรู้จากทฤษฎีสารสนเทศ (information gain) ซึ่งค่าสารสนเทศของข้อมูลจะขึ้นอยู่กับความน่าจะเป็นของข้อมูล สามารถเขียนในรูปสมการที่ 1 (Han and Kamber, 2006, p.297) ได้ดังนี้

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i) \qquad (2.3)$$

pi เป็นความน่าจะเป็นที่ข้อมูลในฐานข้อมูล D อยู่ในกลุ่ม Ci ซึ่งมีค่า |Ci,D| / |D|
m เป็นจำนวนกลุ่มทั้งหมดที่ต่างกันของข้อมูลชุดนั้น
Ci เป็นกลุ่มในลำดับที่ i โดยที่ i มีค่าระหว่าง 1 ถึง m
|Ci,D| เป็นจำนวนข้อมูลในฐานข้อมูล D ที่อยู่ในกลุ่ม Ci
|D| เป็นจำนวนข้อมูลในฐานข้อมูล D

ค่า Info(D) ที่ได้สามารถเรียกอีกชื่อหนึ่งว่าค่า Entropy ของฐานข้อมูล D

ค่าความรู้จากทฤษฎีสารสนเทศจะช่วยในการแยกแยะข้อมูลทำให้ลดจำนวนครั้งของการทดสอบได้ อีกทั้งยังรับประกันว่าต้นไม้ตัดสินใจที่ได้จะไม่มีความซับซ้อนมากจนเกินไป

เมื่อทำการพิจารณาเลือกคุณลักษณะเป็นตัวเลือกทดสอบ จะใช้ค่าความรู้จากทฤษฎีสารสนเทศของคุณลักษณะ สามารถเขียนในรูปสมการที่ 2 ได้ดังนี้

$$info_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} \text{ X } info(D_i) \qquad (2.4)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

v เป็นจำนวนค่าที่เป็นไปได้ของคุณลักษณะ

|D| เป็นจำนวนข้อมูลในฐานข้อมูล D

|Dj| เป็นจำนวนข้อมูลในฐานข้อมูล D ที่มีค่าที่ j ของคุณลักษณะ A

ค่ามาตรฐานเกนที่จะใช้พิจารณาเลือกคุณลักษณะ A มาเป็นโหนดของต้นไม้มีค่าเท่ากับผลต่างของความรู้จากทฤษฎีสารสนเทศ กับ ความรู้จากทฤษฎีสารสนเทศของคุณลักษณะ สามารถเขียนในรูปสมการที่ 3 ได้ดังนี้

$$Gain(A) = info(D) - info_A(D) \quad (2.5)$$

เริ่มต้นการสร้างต้นไม้ตัดสินใจเราต้องพิจารณาคุณลักษณะที่มีของข้อมูลเรียนรู้เพื่อเลือกเป็นโหนดราก โดยการคำนวณค่ามาตรฐานเกนของคุณลักษณะที่มีทั้งหมด แล้วจึงตัดสินใจเลือกคุณลักษณะที่มีค่ามาตรฐานเกนสูงที่สุด ในที่นี้จะสังเกตได้ว่าจะมีคุณลักษณะที่สามารถนำมาใช้ตัดสินใจได้คือ age, income, student และ credit_rating ส่วนคุณลักษณะ RID มีลักษณะเป็นค่าไม่ซ้ำ (unique value) จึงไม่เหมาะสมในการนำมาใช้ตัดสินใจ และ Class ก็เป็นกลุ่มของข้อมูล ก็ไม่เหมาะสมเช่นกัน

จากตัวอย่างข้อมูลคุณลักษณะประกอบการตัดสินใจซื้อคอมพิวเตอร์ เซตของข้อมูลเรียนรู้ T ประกอบด้วยข้อมูลจำนวน 14 แถว แบ่งข้อมูลออกเป็น 2 กลุ่มคือ ข้อมูลที่ตัดสินใจซื้อคอมพิวเตอร์ (Class = yes) จำนวน 9 แถว และตัดสินใจไม่ซื้อคอมพิวเตอร์ (Class = no) จำนวน 5 แถว

$$Info(T) = -\left(\frac{9}{14}\right) x \log\left(\frac{9}{14}\right) - \left(\frac{5}{14}\right) x \log_2\left(\frac{5}{14}\right)$$

พิจารณาแต่ละคุณลักษณะโดยการหาค่าความรู้จากทฤษฎีสารสนเทศของคุณลักษณะ แลค่ามาตรฐานเกนออกมาโดยใช้สมการที่ 2 และ 3 ตามลำดับ ดังนี้

$$Info_{age}(T) = \left(\frac{5}{14}\right) x \left(-\frac{2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right)\right)$$
$$+ \left(\frac{4}{14}\right) x \left(-\frac{4}{4}\log_2\left(\frac{4}{4}\right) - \frac{0}{4}\log_2\left(\frac{0}{4}\right)\right)$$
$$+ \left(\frac{5}{14}\right) x \left(-\frac{3}{5}\log_2\left(\frac{3}{5}\right) - \frac{2}{5}\log_2\left(\frac{2}{5}\right)\right)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$= 0.693$$

$$\text{Gain(age)} = \text{Info(T)} - \text{Info}_{\text{age}}(\text{T})$$

$$= 0.940 - 0.693$$

$$= 0.247$$

$$\text{Gain(income)} = \text{Info(T)} - \text{Info}_{\text{income}}(\text{T})$$

$$= 0.940 - 0.911$$

$$= 0.029$$

$$\text{Gain(student)} = \text{Info(T)} - \text{Info}_{\text{student}}(\text{T})$$

$$= 0.940 - 0.788$$

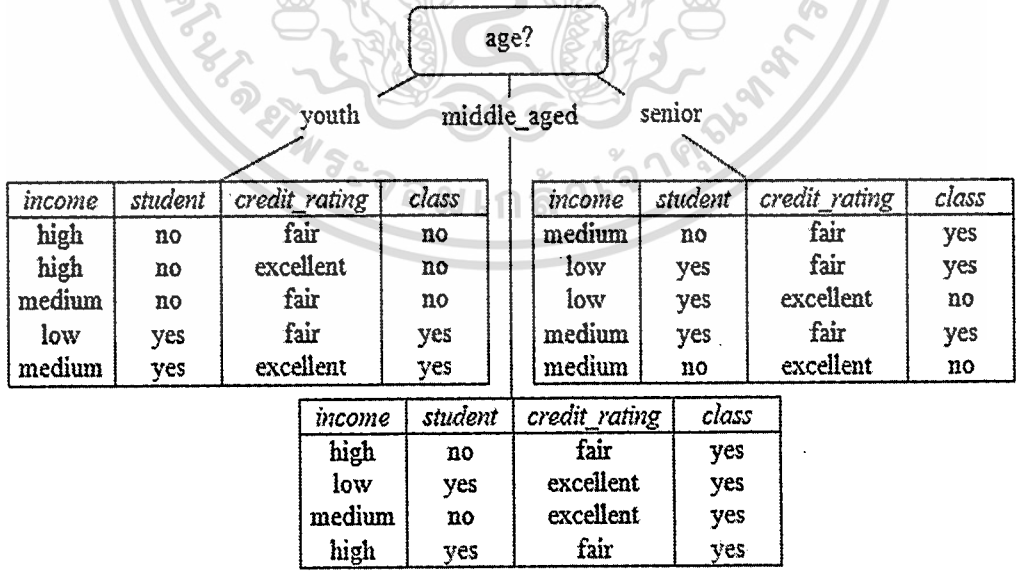
$$= 0.152$$

$$\text{Gain(credit_rating)} = \text{Info(T)} - \text{Info}_{\text{credit_rating}}(\text{T})$$

$$= 0.940 - 0.892$$

$$= 0.048$$

ซึ่งจะเห็นได้ว่าคุณลักษณะที่ให้ค่ามาตรฐานเกนสูงที่สุดคือ age ดังนั้นคุณลักษณะ age จึงถูกเลือกเป็นโหนดรากของต้นไม้ตัดสินใจ แต่เนื่องจากข้อมูลทั้งหมดของแต่ละกิ่งไม่ได้อยู่ในกลุ่มเดียวกันทั้งหมด จึงต้องมีการเลือกโหนดสร้างต้นไม้ตัดสินใจต่อไปบนกิ่งของโหนดราก ยกเว้นกรณี age = middle_aged จะไม่ต้องสร้างต้นไม้ตัดสินใจเพิ่มเติมเนื่องจากสามารถจัดกลุ่มของข้อมูลที่เป็นกลุ่ม yes ได้ทั้งหมดแล้ว ซึ่งจะแสดงได้ดังรูปด้านล่าง



income	*student*	*credit_rating*	*class*
high	no	fair	no
high	no	excellent	no
medium	no	fair	no
low	yes	fair	yes
medium	yes	excellent	yes

income	*student*	*credit_rating*	*class*
medium	no	fair	yes
low	yes	fair	yes
low	yes	excellent	no
medium	yes	fair	yes
medium	no	excellent	no

income	*student*	*credit_rating*	*class*
high	no	fair	yes
low	yes	excellent	yes
medium	no	excellent	yes
high	yes	fair	yes

รูปที่ 2-10 แสดงโหนดของการตัดสินใจในการซื้อคอมพิวเตอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเลือกโหนดระดับที่ 2 จะมีเพียง income, student และ credit_rating เท่านั้นที่สามารถเป็นตัวทดสอบการตัดสินใจได้ ได้ การสร้างต้นไม้ระดับที่ 2 จะแบ่งพิจารณาเป็นทีละส่วนคือage = youth และ age = senior ในที่นี้เราจะพิจารณาต้นไม้ age = youth ก่อนโดยใช้วิธีการหาคุณลักษณะที่เหมาะสมซึ่งจะมีการใช้ค่ามาตรฐานเกนดังเดิม ซึ่งแสดงได้ดังนี้

$$\text{Info}(\text{age} = \text{youth}) = -\left(\frac{2}{5}\right) \times \log_2\left(\frac{2}{5}\right) - \left(\frac{3}{5}\right) \times \log_2\left(\frac{3}{5}\right)$$
$$= 0.971$$

พิจารณาแต่ละคุณลักษณะโดยการหาค่าความรู้จากทฤษฎีสารสนเทศของคุณลักษณะ และค่าเกนออกมาโดยใช้สมการที่ 2 และ 3 ตามลำดับ ดังนี้

$$\text{Info}_{\text{income}}(\text{age} = \text{youth}) = \left(\frac{2}{5}\right) \times \left(-\frac{0}{2}\log_2\left(\frac{0}{2}\right) - \frac{2}{2}\log_2\left(\frac{2}{2}\right)\right)$$
$$+ \left(\frac{2}{5}\right) \times \left(-\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right)\right)$$
$$+ \left(\frac{1}{5}\right) \times \left(-\frac{1}{1}\log_2\left(\frac{1}{1}\right) - \frac{0}{1}\log_2\left(\frac{0}{1}\right)\right)$$

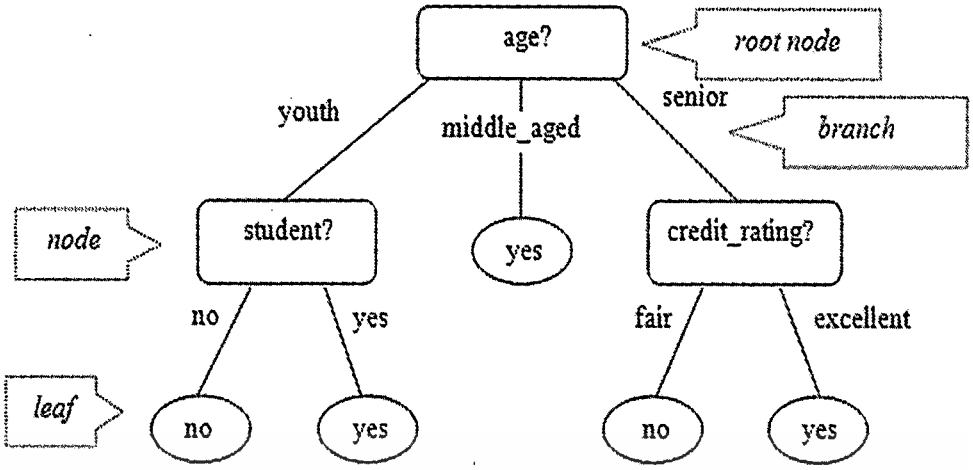
$$\text{Gain}(\text{income}) = \text{Info}(\text{age} = \text{youth}) - \text{Info}_{\text{income}}(\text{age} = \text{youth})$$
$$= 0.971 - 0.4$$
$$= 0.571$$

$$\text{Gain}(\text{student}) = \text{Info}(\text{age} = \text{youth}) - \text{Info}_{\text{student}}(\text{age} = \text{youth})$$
$$= 0.971 - 0$$
$$= 0.971$$

$$\text{Gain}(\text{credit_rating}) = \text{Info}(\text{age} = \text{youth}) - \text{Info}_{\text{credit_rating}}(\text{age} = \text{youth})$$
$$= 0.971 - 0.951$$
$$= 0.020$$

จะเห็นว่าคุณสมบัติที่ให้ค่ามาตรฐานเกนสูงที่สุดคือ student ดังนั้นจึงเลือกคุณสมบัตินี้เป็นโหนดระดับที่ 2 ต่อจาก age = youth และเนื่องจากข้อมูลทั้งหมดของแต่ละกิ่งอยู่ในกลุ่มเดียวกันทั้งหมด ดังนั้นจึงไม่ต้องสร้างต้นไม้ตัดสินใจต่อไป แต่ยังคงเหลือโหนดระดับ 2 ทางขวา (age = senior) ที่ต้องพิจารณาเลือกคุณสมบัติด้วยวิธีการเดียวกับที่ผ่านมา ซึ่งจะได้คุณสมบัติ credit_rating มีค่าเกนสูงที่สุด จึงจะถูกเลือกเป็นโหนดระดับ 2 ต่อจาก age = senior ซึ่งทำให้ข้อมูลทั้งหมดของแต่ละกิ่งอยู่ในกลุ่มเดียวกันทั้งหมด และจะได้โครงสร้างเป็นดังรูป

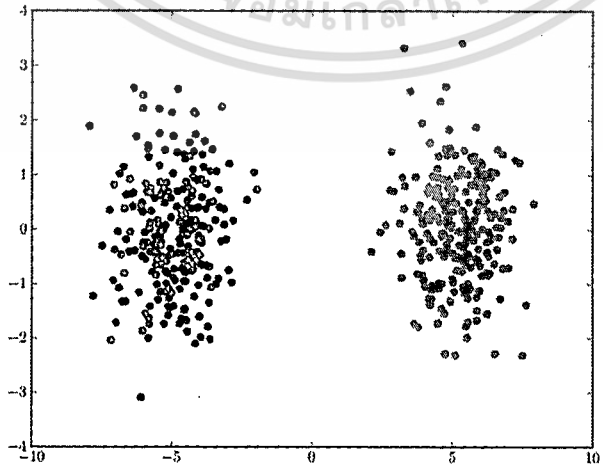
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2-11 แสดงโหนดของการตัดสินใจในการซื้อคอมพิวเตอร์ในรูปแบบโครงสร้าง

2.4.4 การจัดกลุ่ม (Clustering)

การจัดกลุ่ม (Clustering) เป็นวิธีการที่พิจารณาข้อมูลแต่ละแถวเสมือนเป็นวัตถุ (object) ซึ่งจะมีหลักการเหมือนกับการจำแนกประเภทข้อมูล คือจะทำการแบ่งข้อมูลออกเป็นกลุ่ม (คลัสเตอร์) โดยจะจัดให้ข้อมูลที่มีความคล้ายคลึงกันอยู่ในคลัสเตอร์เดียวกัน และข้อมูลที่อยู่ต่างคลัสเตอร์กันจะมีความคล้ายคลึงกันน้อยที่สุด ซึ่งความเหมือนหรือต่างกันสามารถเปรียบเทียบได้กับความใกล้ชิดกันของวัตถุใดๆ โดยใช้ระยะทางเป็นตัวชี้วัด คุณภาพของแต่ละคลัสเตอร์สามารถอธิบายได้จากเส้นผ่านศูนย์กลางของคลัสเตอร์(diameter) ซึ่งแสดงระยะห่างมากสุดของวัตถุสองชิ้นที่อยู่ในคลัสเตอร์เดียวกัน แต่ละคลัสเตอร์จะมีตัวแทนที่สามารถแทนวัตถุทุกชิ้นของคลัสเตอร์นั้นได้ เช่น การใช้จุดศูนย์กลางคลัสเตอร์(centroid) แทนคลัสเตอร์นั้น



รูปที่ 2-12 แสดงภาพจำลองการจัดกลุ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประเภทของขั้นตอนวิธีการจัดแบ่งกลุ่มข้อมูล

1. Partition Method การจัดกลุ่มข้อมูลประเภทนี้จะทำการสร้าง k พาร์ทิชันบนฐานข้อมูลจำนวน n เรคคอร์ด โดยแต่ละพาร์ทิชันจะแสดงถึงข้อมูลที่ถูกแบ่งออกเป็นกลุ่ม ในแต่ละกลุ่มจะประกอบไปด้วยข้อมูลอย่างน้อยที่สุด 1 แถว และข้อมูลแต่ละแถวจะต้องถูกจัดให้อยู่ในกลุ่มข้อมูลเพียงกลุ่มเดียวเท่านั้น (สำหรับบางเทคนิคอาจอนุโลมให้ข้อมูลใดๆสามารถถูกจัดอยู่ในกลุ่มข้อมูลได้มากกว่า 1 กลุ่ม) ตัวอย่างเทคนิคของการจัดกลุ่มแบบนี้ ได้แก่ k-means algorithm, k-medoids algorithm, CLARA (Clustering LARge Application), CLARANS(Clustering LARge Application based upon RANdomized Search)

2. Hierarchical Method การจัดกลุ่มข้อมูลประเภทนี้จะอาศัยหลักการแบ่งข้อมูลออกเป็นลำดับชั้นคล้ายกับต้นไม้ ซึ่งวิธีการแบ่งกลุ่มข้อมูลแบบนี้สามารถแบ่งออกเป็น 2 แนวทางตามลักษณะการสร้างลำดับชั้นคือ Agglomerative approach กับ Divisive approach ตัวอย่างเทคนิคของการจัดกลุ่มแบบนี้ ได้แก่ AGNES (Agglomerative NESting) ซึ่งจะเป็นagglomerative hierarchical clustering, DIANA (Divisive ANAlysis) ซึ่งจะเป็น divisive hierarchical clustering, BIRCH (Balanced Iterative Reducing and Clustering Using Hierarchies), CURE (Clustering Using REpresentatives)

3. Density-Based Method การจัดกลุ่มข้อมูลประเภทนี้จะพิจารณาความหนาแน่นของข้อมูลเป็นเกณฑ์ในการค้นหาคลัสเตอร์ หลักการทั่วไปของเทคนิคนี้คือการแผ่ขยายขอบเขตของคลัสเตอร์ไปเรื่อยๆ ตราบใดที่ความหนาแน่นของข้อมูลยังมีค่าน้อยกว่าหรือเท่ากับค่าที่ผู้ใช้กำหนด นั่นคือแต่ละข้อมูลของคลัสเตอร์ใดๆ จะต้องประกอบด้วยข้อมูลซึ่งอยู่ใกล้กันภายในรัศมีที่กำหนด (neighborhood) ด้วยเทคนิคนี้สามารถใช้ในการกรองข้อมูลรบกวน (noisy) ซึ่งเป็นข้อมูลที่มีความหนาแน่นเบาบางได้ และยังสามารถค้นหาคลัสเตอร์ที่มีรูปทรงซับซ้อนได้อีกด้วย ตัวอย่างเทคนิคของการจัดกลุ่มแบบนี้ ได้แก่ DBSCAN

2.4.5 ระบบลูกค้าสัมพันธ์ (CRM)

CRM ย่อมาจาก Customer Relationship Management คือ การบริหารความสัมพันธ์กับลูกค้า ซึ่งหมายถึงวิธีการที่เราจะบริหารให้ลูกค้ามีความรู้สึกผูกพันธ์กับสินค้า บริการ หรือองค์กรของเรา เมื่อลูกค้าเค้ามีความผูกพันธ์ในทางที่ดีกับเรา แล้วก็ลูกค้านั้นไม่คิดที่จะเปลี่ยนใจไปจากสินค้าหรือบริการของเรา ทำให้เรามีฐานลูกค้าที่มั่นคง และนำมาซึ่งความมั่นคงของบริษัท ดังนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การที่จะรู้ซึ้งถึงสถานะความผูกพันธ์กับลูกค้าได้นั้น เราก็ต้องอาศัยการสังเกตุพฤติกรรมของลูกค้าแล้วนำมาวิเคราะห์หาความเกี่ยวข้องระหว่าง พฤติกรรมของลูกค้ากับกลยุทธ์ทางการตลาดของเรา กระบวนการทำงานของระบบCRMมี4ขั้นตอนดังนี้

1.Identify เก็บข้อมูลว่าลูกค้าของบริษัทเป็นใคร เช่น ชื่อลูกค้า ข้อมูลสำหรับติดต่อกับลูกค้า

2.Differentiate วิเคราะห์พฤติกรรมของลูกค้าแต่ละคน และจัดแบ่งลูกค้าออกเป็นกลุ่มตามคุณสมบัติ

3.Interact มีปฏิสัมพันธ์กับลูกค้าเพื่อเรียนรู้ความต้องการของลูกค้า และเพื่อสร้างความพึงพอใจให้กับลูกค้าในระยะยาว

4.Customize เป็นการนำเสนอสินค้าหรือบริการที่มีความเหมาะสมเฉพาะตัวกับลูกค้าแต่ละคน

ระบบ CRM มีความสัมพันธ์เกี่ยวกับเทคโนโลยีสารสนเทศในด้านการนำซอฟต์แวร์มาใช้ในการพัฒนาเพื่อดึงดูดลูกค้า สร้างความสัมพันธ์กับลูกค้า เช่น Web Site เป็นต้น CRM software ช่วยในการจัดการฐานข้อมูลลูกค้า วิเคราะห์ข้อมูล และเป็นช่องทางในการติดต่อกับลูกค้า ซึ่ง CRM software มักแบ่งออกเป็น3ส่วนดังนี้

1.Operational CRM เป็นซอฟท์แวร์ front office ที่ใช้ช่วยจัดการกระบวนการทางธุรกิจที่เกี่ยวข้องกับลูกค้า ไม่ว่าจะเป็น sales marketing หรือ service เช่น การจัดการข้อมูลติดต่อลูกค้า การเสนอราคา การบริหารฝ่ายขาย การเก็บข้อมูลพฤติกรรมลูกค้า ระบบบริการลูกค้า เป็นต้น

2.Analytical CRM ระบบวิเคราะห์ข้อมูลลูกค้าที่รวบรวมได้จากส่วน Operational CRM หรือจากแหล่งอื่นๆ เพื่อแบ่งลูกค้าออกเป็นกลุ่มและค้นหากลุ่มลูกค้าเป้าหมายที่บริษัทสามารถนำเสนอสินค้าหรือบริการเพิ่มเติมได้

3.Collaborative CRM ระบบช่วยสนับสนุนในการติดต่อปฏิสัมพันธ์กับลูกค้าผ่านทางช่องทางต่างๆ เช่น ติดต่อส่วนตัว จดหมาย แฟกซ์ โทรศัพท์ เว็บไซต์ E-Mail เป็นต้น รวมถึงช่วยจัดการทรัพยากรที่บริษัทมีคือพนักงาน กระบวนการทำงาน และฐานข้อมูลลูกค้า เพื่อนำไปให้บริการแก่ลูกค้าและช่วยรักษาฐานลูกค้าของบริษัทได้ดีขึ้น

<u>ประโยชน์ของCRMต่อธุรกิจ</u>

1. ช่วยเพิ่มความสามารถในการให้บริการลูกค้าได้ดีขึ้น เช่น ใช้เว็บไซต์ในการให้ข้อมูลเกี่ยวกับสินค้า วิธีการใช้สินค้า และให้ความช่วยเหลือแก่ลูกค้าได้ตลอด 24 ชั่วโมง เป็นช่องทางให้ลูกค้าแนะนำติชมต่อบริการของบริษัทได้ง่าย ช่วยให้ลูกค้าสามารถปรับเปลี่ยนความต้องการของตนเองได้ทันที เป็นต้น

2. ช่วยพัฒนาความสัมพันธ์ระหว่างบริษัทกับลูกค้าให้ดีขึ้น ช่วยให้บริษัทรู้ความสนใจ ความต้องการ และพฤติกรรมการซื้อของลูกค้า ทำให้บริษัทสามารถนำเสนอสินค้าที่เหมาะสม กับลูกค้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ได้ และช่วยให้บริษัทสามารถให้บริการหลังการขายแก่ลูกค้าตามที่ลูกค้าต้องการได้ การสร้างความสัมพันธ์กับลูกค้าในระยะยาวจะช่วยเพิ่มจงรักภัคดีที่ลูกค้ามีต่อบริษัท ลดการสูญเสียลูกค้า ลดต้นทุนการตลาด เพิ่มรายได้จากการที่ลูกค้าซื้อซ้ำหรือแนะนำให้คนรู้จักซื้อสินค้าของบริษัท และนั่นหมายถึงกำไรของบริษัทที่เพิ่มมากขึ้น

ในการศึกษาอิสระนี้ จะนำผลการวิเคราะห์การขายที่ได้มาประยุกต์ใช้กับการแนะนำซอฟต์แวร์ที่เหมาะสมให้แก่ลูกค้า โดยผลการวิเคราะห์นั้นเกิดมาจากการเก็บข้อมูลพฤติกรรมลูกค้าในขนาดอุตสาหกรรมต่างๆ

2.4.6ระบบสนับสนุนการตัดสินใจ(DSS)

ระบบสนับสนุนการตัดสินใจ เป็นระบบสารสนเทศที่ช่วยในการตัดสินใจ ซึ่งมีลักษณะมีโครงสร้างไม่ชัดเจน โดยนำข้อมูลมาจากหลายแหล่งช่วยในการนำเสนอและมีลักษณะยืดหยุ่นตามความต้องการ ลักษณะของระบบสนับสนุนการตัดสินใจ (Characteristics of DSS) จะมีการแบ่งแยกการตัดสินใจในลักษณะที่เป็นโครงสร้าง ลักษณะไม่เป็นโครงสร้าง และลักษณะกึ่งโครงสร้าง โดยส่วนใหญ่ปัญหาที่เป็นโครงสร้างเกิดจากการทำงานประจำ ส่วนปัญหาที่ไม่เป็นโครงสร้างเกิดจากความไม่แน่นอน เนื่องจากงานไม่เป็นประจำ ปัญหาที่เกิดขึ้นระหว่างกึ่งโครงสร้าง มักจะเกิดจากปัญหาที่มีโครงสร้างแต่ผิดไปจากงานประจำบ้าง โดยทั่วไประบบสนับสนุนการตัดสินใจ(DSS) มักจะถูกนำมาใช้ในการออกแบบเพื่อสนับสนุนการตัดสินใจ วิเคราะห์ปัญหาในลักษณะกึ่งโครงสร้างและไม่เป็นโครงสร้าง ซึ่งแตกต่างจากระบบการประมวลผลรายการ (TPS) และระบบข้อมูลเพื่อการบริหาร(MIS) ซึ่งมีการจัดการกับปัญหาที่เป็นโครงสร้าง และในขณะเดียวกันองค์การมักจะใช้การวิจัยเชิงปฏิบัติ (Operations Research (OR)) ในการจัดเตรียมตัวแบบเพื่อใช้ในการสร้างทางเลือก แต่ในปัจจุบันระบบสนับสนุนการตัดสินใจ(DSS) ได้ถูกออกแบบเพื่อรวมความสามารถของระบบการประมวลผลรายการ (TPS) ระบบข้อมูลเพื่อการบริหาร(MIS) และการสร้างตัวแบบของระบบวิจัยเชิงปฏิบัติการ (OR) เข้าด้วยกัน เพื่อช่วยในการวิเคราะห์ปัญหาในรูปแบบต่าง ๆ และสร้างเครื่องมือในการวิเคราะห์ให้กับผู้ใช้

การออกแบบระบบสนับสนุนการตัดสินใจ(DSS) ที่ดีจะต้องสามารถตอบสนองต่อการนำไปใช้ของผู้บริหารระดับสูง สามารถช่วยในเรื่องการพยากรณ์ การวางแผนเชิงกลยุทธ์ และผู้บริหารระดับกลางสามารถที่จะนำไปปฏิบัติได้ ในอดีตการตัดสินใจที่ผิดพลาดมักจะเกิดจากตัวบุคคล แต่ในปัจจุบันการตัดสินใจมักจะขึ้นกับกลุ่มบุคคลมากกว่าตัวบุคคล จึงมีการนำระบบสนับสนุนการตัดสินใจ(DSS) มาช่วยในการตัดสินใจเป็นกลุ่ม (Group Decision Support System (GDSS)) ระบบสนับสนุนการตัดสินใจ(DSS) ที่ดีจะสามารถช่วยในการหาข้อมูล เลือกข้อมูลและ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใช้ตัวแบบที่เหมาะสมในการวิเคราะห์ข้อมูล รวมถึงการควบคุมการทำงานด้วยตนเองปราศจากการแทรกแซงของผู้ชำนาญ ผู้ชำนาญจะทำหน้าที่ในการให้คำปรึกษา ฝึกอบรม แนะนำ และสนับสนุน ส่วนการปฏิบัติเป็นหน้าที่ของผู้ใช้

<u>เป้าหมายของระบบสนับสนุนการตัดสินใจ</u>

เป้าหมายของระบบสนับสนุนการติดสินใจ (DSS Goals) และการประยุกต์ใช้ในองค์การส่วนใหญ่ มักจะใช้ระบบสนับสนุนการตัดสินใจ(DSS) โดยมีเป้าหมายหรือจุดมุ่งหมายหลัก 3 ประการดังนี้

1. การตัดสินใจแบบกึ่งโครงสร้างและแบบไม่มีโครงสร้าง (Semistructured and Unstructured decisions) ในความเป็นจริงข้อมูลที่เกิดขึ้นภายในองค์การมีข้อมูลที่หลากหลาย ข้อมูลบางประเภทเกิดจากกิจกรรมของธุรกิจที่เกิดขึ้นประจำวัน (Routine) ซึ่งข้อมูลเหล่านั้นย่อมมีมาตรฐานหรือมีโครงสร้างที่แน่นอน ซึ่งองค์การสามารถใช้ระบบข้อมูลเพื่อการบริหาร(MIS) ในส่วนระบบการประมวลผลรายการ (TPS) ได้ แต่ในขณะเดียวกันมีข้อมูลบางประเภทที่มีความจำเป็นต้องใช้ในการวิเคราะห์ และมีการเปลี่ยนแปลงอยู่ตลอดเวลา เช่น ต้นทุนของวัตถุดิบหรือค่าใช้จ่ายต่าง ๆ ที่อยู่นอกเหนือการควบคุมขององค์การ ในลักษณะนี้จะเกิดข้อมูลที่ไม่มีโครงสร้างขึ้นหรือกึ่งโครงสร้าง ซึ่งระบบการประมวลผลรายการ (TPS) จะไม่เหมาะสำหรับการประมวลผลประเภทนี้ ผู้บริหารจึงจำเป็นต้องใช้ระบบสนับสนุนการตัดสินใจ(DSS) เพื่อช่วยในการตัดสินใจ

2. ความสามารถในการปรับปรุงความต้องการที่เปลี่ยนไป (Ability to adapt changing needs) ผู้บริหารระดับสูงมักจะมีความต้องการสารสนเทศที่หลากหลาย เพื่อช่วยในการตัดสินใจ แต่ระบบการประมวลผลรายการ (TPS) มักจะให้สารสนเทศในเชิงบริหาร เช่น งบดุล งบกำไรขาดทุน แต่ไม่มีงบการเงินเพื่อช่วยในการตัดสินใจของผู้บริหาร ดังนั้นระบบสนับสนุนการตัดสินใจ (DSS) จึงต้องมีความสามารถในการปรับปรุงข้อมูล ให้ออกมาในรูปที่ช่วยในการตัดสินใจได้

3. ง่ายต่อการเรียนรู้และนำมาใช้ (Ease of learning and use) ระบบ DSS เป็นระบบที่สร้างขึ้นมาโดยคาดหวังว่าผู้ใช้โดยทั่วไปสามารถเรียนรู้ได้ และนำไปปฏิบัติได้ โดยไม่จำเป็นต้องอาศัยผู้เชี่ยวชาญทางด้านคอมพิวเตอร์ เช่น โปรแกรม โลตัส 1-2-3 (Lotus 1-2-3) MS-EXCEL เป็นต้น นอกจากนั้นยังสามารถแสดงผลในรูปของข้อความ (Text) รูปภาพ (Graphics) และตัวเลข (Numeric) ได้

<u>ส่วนประกอบของระบบสนับสนุนการตัดสินใจ</u>

ส่วนประกอบของระบบสนับสนุนการตัดสินใจ (Component of DSS) ส่วนประกอบเบื้องต้นของระบบสนับสนุนการตัดสินใจ(DSS) เป็นการเชื่อมระหว่างฮาร์ดแวร์ ซอฟต์แวร์ และข้อมูลที่จำเป็นที่จะต้องใช้ในการตัดสินใจของผู้จัดการหรือผู้ใช้ส่วนประกอบของระบบสนับสนุนการตัดสินใจ (DSS) สามารถจำแนกออกเป็น 4 ส่วน ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. อุปกรณ์ เป็นส่วนประกอบแรกและเป็นโครงสร้างพื้นฐานของ DSS โดยอุปกรณ์ที่เกี่ยวข้องกับระบบสนับสนุนการตัดสินใจจะสามารถแบ่งออกเป็น 3 กลุ่ม ด้วยกันคือ

1.1. อุปกรณ์ประมวลผล ประกอบด้วยระบบคอมพิวเตอร์ ซึ่งในสมัยเริ่มแรกจะใช้คอมพิวเตอร์ขนาดใหญ่ (Mainframe) หรือมินิคอมพิวเตอร์ (Mini Computer) ในสำนักงานเป็นหลัก แต่ในปัจจุบันองค์การส่วนมากหันมาใช้ระบบเครือข่ายของคอมพิวเตอร์ส่วนบุคคล (Personal Computer) แทนเนื่องจากมีราคาถูก มีประสิทธิภาพดี และสะดวกต่อการใช้งาน ตลอดจนผู้ใช้มีความรู้ ความเข้าใจ และทักษะในงานสารสนเทศสูงขึ้น โดยเฉพาะผู้บริหารรุ่นใหม่ที่มีความรู้เกี่ยวกับเทคโนโลยีคอมพิวเตอร์และสารสนเทศ สามารถที่จะพัฒนาระบบสนับสนุนการตัดสินใจ (DSS) ขึ้นบนคอมพิวเตอร์ส่วนบุคคลโดยใช้ชุดคำสั่งประเภทฐานข้อมูล และSpread Sheet ประกอบ

1.2. อุปกรณ์สื่อสาร ประกอบด้วยระบบสื่อสารต่างๆ เช่น ระบบเครือข่ายเฉพาะพื้นที่ (LAN) ได้ถูกนำเข้ามาประยุกต์ เพื่อทำการสื่อสารข้อมูลและสารสนเทศของDSS โดยในบางครั้งอาจจะใช้การประชุมโดยอาศัยสื่อวีดีโอ (Video Conference) หรือการประชุมทางไกล (Teleconference)
ประกอบ เนื่องจากผู้มีหน้าที่ตัดสินใจอาจอยู่กันคนละพื้นที่

1.3. อุปกรณ์แสดงผล ระบบสนับสนุนการตัดสินใจ (DSS) ที่มีประสิทธิภาพจำเป็นต้องมีอุปกรณ์แสดงผล เช่น จอภาพที่มีความละเอียดสูง เครื่องพิมพ์อย่างดี และอุปกรณ์ประกอบอื่นๆ เพื่อช่วยถ่ายทอดข้อมูลสารสนเทศ ตลอดจนสร้างความเข้าใจในสารสนเทศให้แก่ผู้ใช้ และช่วยให้การดำเนินงานมีประสิทธิภาพ

2. ระบบการทำงาน มีนักวิชาการหลายท่านให้ความเห็นว่า ระบบการทำงานเป็นส่วนประกอบหลักของ DSS เพราะถือว่าเป็นส่วนประกอบสำคัญในการที่จะทำให้ DSSทำงานได้ตามวัตถุประสงค์และความต้องการของผู้ใช้ ซึ่งระบบการทำงานจะประกอบด้วยส่วนประกอบสำคัญ 3 ส่วนคือ

2.1. ฐานข้อมูล (Database) ระบบสนับสนุนการตัดสินใจ (DSS) จะไม่มีหน้าที่สร้าง ค้นหาหรือปรับปรุงข้อมูลในฐานข้อมูลขององค์การ เนื่องจากระบบข้อมูลขององค์การเป็นระบบขนาดใหญ่มีข้อมูลหลากหลายและเกี่ยวข้องกับข้อมูลหลายประเภท แต่ระบบสนับสนุนการตัดสินใจ (DSS) จะมีฐานข้อมูลของตัวเอง ซึ่งจะมีหน้าที่รวบรวมข้อมูลที่สำคัญจากอดีตถึงปัจจุบันและนำมาจัดเก็บ เพื่อให้ง่ายต่อการค้นหา ซึ่งข้อมูลเหล่านี้จะถูกเก็บไว้อย่างสมบูรณ์ ครบถ้วน และแน่นอนเพื่อรอการนำไปประมวลผลประกอบการตัดสินใจ ขณะเดียวกันระบบสนับสนุนการตัดสินใจ (DSS) อาจจะต่อเชื่อมกับระบบฐานข้อมูลขององค์การ เพื่อดึงข้อมูลสำคัญบางประเภทมาใช้งาน

2.2. ฐานแบบจำลอง (Model Base) มีหน้าที่รวบรวมแบบจำลองทางคณิตศาสตร์ และแบบจำลองในการวิเคราะห์ปัญหาที่สำคัญ เพื่ออำนวยความสะดวกต่อผู้ใช้ปกติ DSS จะถูก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พัฒนาขึ้นมาตามจุดประสงค์เฉพาะอย่าง ดังนั้นระบบสนับสนุนการตัดสินใจ (DSS) จะประกอบด้วยแบบจำลองที่ต่างกันตามวัตถุประสงค์ในการนำไปใช้

2.3. ระบบชุดคำสั่งของ DSS (DSS Software System) เป็นส่วนประกอบสำคัญที่ช่วยอำนวยความสะดวกในการโต้ตอบระหว่างผู้ใช้กับฐานข้อมูลและฐานแบบจำลอง โดยระบบชุดคำสั่งของระบบสนับสนุนการตัดสินใจ (DSS) จะมีหน้าที่จัดการ ควบคุมการพัฒนา จัดเก็บ และเรียกใช้แบบจำลองต่างๆ เพื่อนำมาประมวลผลกับข้อมูลขากฐานข้อมูล นอกจากนี้ระบบชุดคำสั่งยังมีหน้าที่ให้ความช่วยเหลือผู้ใช้ในการโต้ตอบกับ DSS โดยที่สามารถแสดงความสัมพันธ์ของส่วนประกอบทั้ง 3 ส่วนคือ ผู้ใช้ ฐานแบบจำลอง ฐานข้อมูล

3. ข้อมูล เป็นองค์ประกอบที่สำคัญอีกส่วนของระบบสนับสนุนการตัดสินใจ (DSS) ไม่ว่า ระบบสนับสนุนการตัดสินใจ (DSS) จะประกอบด้วยอุปกรณ์ที่ทันสมัย และได้รับการออกแบบการทำงานให้สอดคล้องกันและเหมาะสมกับการใช้งานมากเพียงใด ถ้าข้อมูลที่นำมาใช้ในการประมวลผลไม่มีคุณภาพเพียงพอแล้วก็จะไม่สามารถช่วยสนับสนุนการตัดสินใจของผู้ใช้ได้อย่างเหมาะสม ซึ่งยังอาจจะสร้างปัญหา หรือความผิดพลาดในการตัดสินใจขึ้นได้ ข้อมูลที่จะนำมาใช้กับระบบสนับสนุนการตัดสินใจ (DSS) จะแตกต่างจากข้อมูลในระบบสารสนเทศอื่น โดยที่ข้อมูลระบบสนับสนุนการตัดสินใจ (DSS) ที่เหมาะสม สมควรที่จะมีลักษณะ ดังต่อไปนี้

3.1. มีปริมาณพอเหมาะแก่การนำไปใช้งาน

3.2. มีความถูกต้องและทันสมัยในระดับที่เหมาะสมกับความต้องการ

3.3. สามารถนำมาใช้ได้สะดวก รวดเร็ว และครบถ้วน

3.4. มีความยืดหยุ่นและสามารถนำมาจัดรูปแบบ เพื่อการวิเคราะห์ได้อย่างเหมาะสม

4. บุคลากร เป็นส่วนประกอบที่สำคัญอีกส่วนหนึ่งของระบบสนับสนุนการตัดสินใจ เนื่องจากบุคคลจะเกี่ยวข้องกับระบบสนับสนุนการตัดสินใจ (DSS) ตั้งแต่ การกำหนดเป้าหมายและความต้องการ การพัฒนา ออกแบบ และการใช้ระบบสนับสนุนการตัดสินใจ (DSS) ซึ่งสามารถแบ่งบุคลากรที่เกี่ยวข้องกับระบบสนับสนุนการตัดสินใจ (DSS) ออกเป็น 2 กลุ่มดังนี้

4.1. ผู้ใช้ (End-user) เป็นผู้ใช้งานโดยตรงของระบบสนับสนุนการตัดสินใจ (DSS) ได้แก่ ผู้บริหารในระดับต่างๆตลอดจนนักวิเคราะห์และผู้เชี่ยวชาญทางด้านธุรกิจที่ต้องการข้อมูลสำหรับประกอบการตัดสินใจในปัญหาที่เกิดขึ้น

4.2. ผู้สนับสนุน DSS (DSS Supports) ได้แก่ ผู้ควบคุมดูแลรักษาอุปกรณ์ต่างๆ ผู้พัฒนาระบบสนับสนุนการตัดสินใจ ผู้จัดการข้อมูล และที่ปรึกษาเกี่ยวกับระบบ เพื่อให้ระบบสนับสนุนการตัดสินใจ (DSS) มีความสมบูรณ์ และสามารถดำเนินงานอย่างเต็มประสิทธิภาพและตรงตามความต้องการของผู้ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เราจะเห็นว่าหัวใจสำคัญของระบบสนับสนุนการตัดสินใจ (DSS) ที่ดีจำเป็นที่จะต้องมีบุคลากรที่มีความสามารถเหมาะสมที่สามารถปรับตัวให้เข้ากับระบบ มีความคิดสร้างสรรค์ และสามารถทำงานได้อย่างสอดคล้องกัน เพื่อให้บรรลุเป้าหมายตามความต้องการขององค์การได้อย่างมีประสิทธิภาพและประสิทธิผล

<u>หน้าที่หลักของระบบสนับสนุนการตัดสินใจ(DSS)</u>

1.ระบบสารสนเทศที่ใช้สำหรับการสนับสนุนผู้ตัดสินใจทางการบริหารทั้งที่เป็นตัวบุคคลหรือกลุ่ม โดยการตัดสินใจนั้นจะเกี่ยวข้องกับสถานการณ์ที่มีลักษณะเป็นแบบ ไม่มีโครงสร้าง (unstructured situations) โดยจะมีการนำวิจารณญาณของมนุษย์กับข้อมูล จากคอมพิวเตอร์มาใช้ประกอบในการตัดสินใจ

2.ระบบสนับสนุนการตัดสินใจ(DSS) ช่วยในการตอบสนองความต้องการที่ไม่ได้คาดการณ์มาก่อน โดยผู้ใช้สามารถปรับข้อมูลในระบบสนับสนุนการตัดสินใจ(DSS) ได้ตลอดเวลาเพื่อจัดการกับเงื่อนไขต่างๆ ที่เปลี่ยนแปลงไป โดยใช้การวิเคราะห์ที่เรียกว่า SensitivityAnalysis

3.ช่วยในการตัดสินใจที่ต้องการความรวดเร็วสูง เพื่อใช้ประกอบในการกำหนดกลยุทธ์ในการแข่งขัน ดังนั้น ระบบสนับสนุนการตัดสินใจ (DSS) จึงมีลักษณะการโต้ตอบได้ (interactive)

4.เสนอทางวิเคราะห์ในทางเลือกต่างๆในสถานการณ์ที่มีความซับซ้อน

5.จัดการเก็บข้อมูลซึ่งมาจากหลายแหล่งได้ทั้งภายในและภายนอกหน่วยงาน

6.นำเสนอได้ทั้งรายงานที่เป็นข้อความและกราฟฟิค

2.5 เทคโนโลยีที่เกี่ยวข้องกับการพัฒนาระบบ

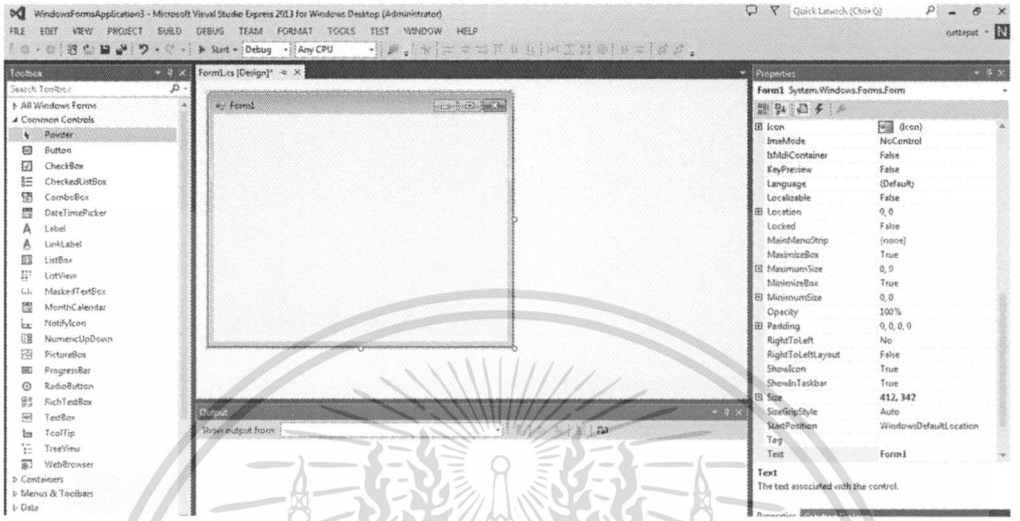
2.5.1 ภาษา C# (C# Programming Language)

ภาษาซีชาร์ป (C# Programming Language) เป็นภาษาโปรแกรมแบบหลายโมเดล ที่ใช้ระบบชนิดข้อมูลแบบรัดกุม (strong typing) และสนับสนุนการเขียนโปรแกรมเชิงคำสั่ง การเขียนโปรแกรมเชิงประกาศ การเขียนโปรแกรมเชิงฟังก์ชัน การเขียนโปรแกรมเชิงกระบวนการ การเขียนโปรแกรมเชิงวัตถุ (แบบคลาส) และการเขียนโปรแกรมเชิงส่วนประกอบ พัฒนาเริ่มแรกโดยบริษัทไมโครซอฟท์เพื่อทำงานบนดอตเน็ตเฟรมเวิร์ก โดยมีแอนเดอร์ เฮลส์เบิร์ก (Anders Hejlsberg) เป็นหัวหน้าโครงการ และมีรากฐานมาจากภาษาซีพลัสพลัสและภาษาอื่นๆ (โดยเฉพาะภาษาเดลไฟและจาวา) โดยมีจุดมุ่งหมายให้เป็นภาษาสมัยใหม่ที่ไม่ซับซ้อน ใช้งานได้ทั่วไป (general-purpose) และเป็นเชิงวัตถุเป็นหลัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.2 Visual Studio Express

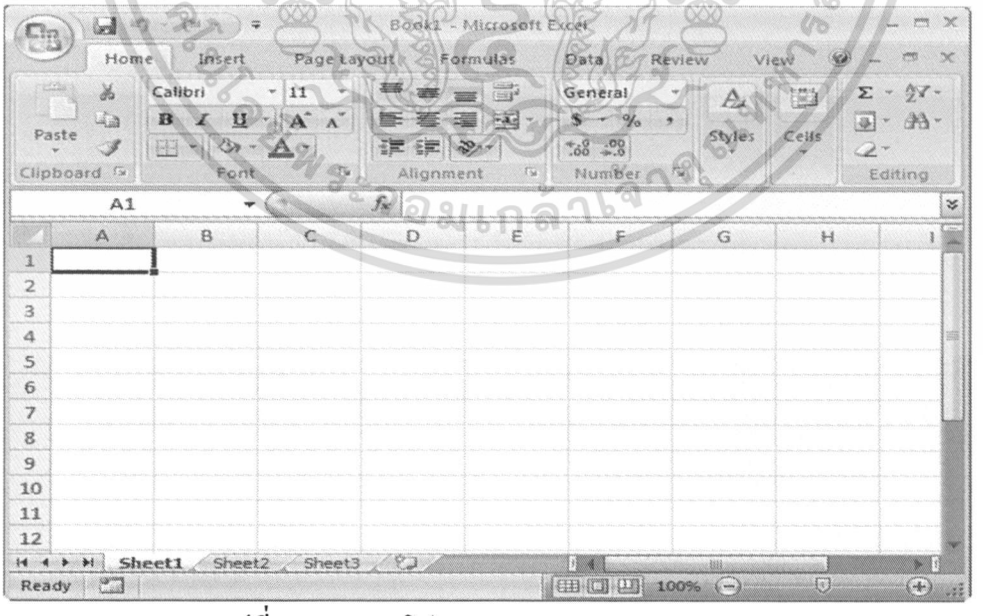
เป็นโปรแกรมประเภท IDE (Integrated Development Environment) ที่เป็นเครื่องมือที่สร้างสภาพแวดล้อม (Environment) ให้เหมาะแก่การเขียนโปรแกรม ซึ่งใช้ในการพัฒนาแอปพลิเคชัน



รูปที่ 2-13 แสดงโปรแกรม Microsoft Visual Studio Express

2.5.3 Microsoft Excel

เป็นเครื่องมือในการจัดเก็บข้อมูลที่อยู่ในรูปไฟล์ .csv

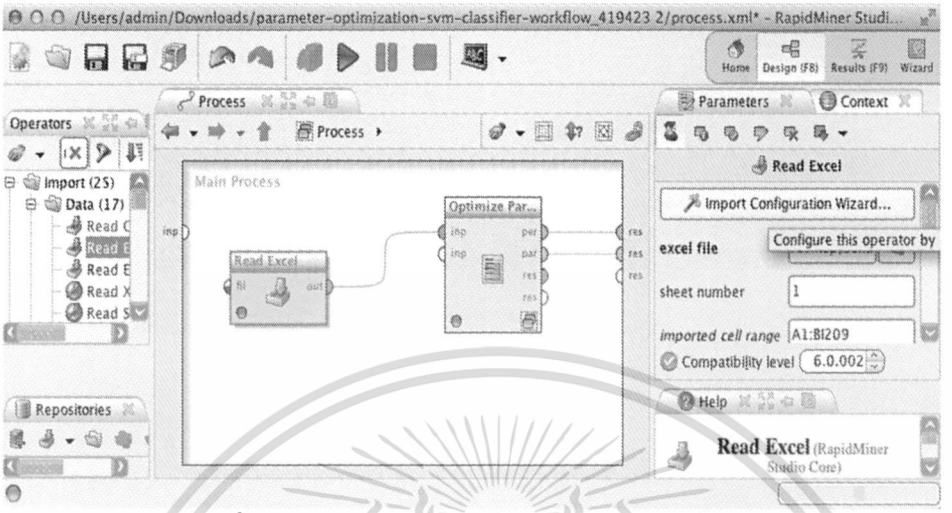


รูปที่ 2-14 แสดงโปรแกรม Microsoft Excel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.4 Rapidminer Studio 6

เป็นเครื่องมือที่ใช้สำหรับการสร้างรูปแบบ วิเคราะห์ข้อมูล จัดกลุ่มข้อมูล



รูปที่ 2-15 แสดงโปรแกรม Rapidminer Studio 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การวิเคราะห์ระบบปัจจุบัน

3.1 การทำงานของระบบงานปัจจุบัน

การทำงานของบริษัทออริซอฟท์ในปัจจุบัน ซึ่งเป็นบริษัทขนาดเล็ก การจัดเก็บข้อมูลต่างๆ จะถูกเก็บในรูปแบบของเอกสารเป็นส่วนใหญ่ และมีขั้นตอนการดำเนินงานดังนี้

- การหาลูกค้า

ขั้นตอนการหาลูกค้าของบริษัทนั้น เริ่มต้นจากพนักงานขายโทรไปสอบถามยังบริษัทต่างๆ ถึงระบบซอฟต์แวร์ด้านบุคคลที่ใช้อยู่ในปัจจุบัน และชักชวนบริษัทนั้นๆให้เปลี่ยนมาใช้ระบบซอฟต์แวร์ของทางออริซอฟท์

- การเก็บข้อมูลลูกค้าและข้อมูลการขาย

เมื่อลูกค้าตกลงที่จะซื้อซอฟต์แวร์กับทางบริษัทแล้ว พนักงานขายจะทำเอกสารการซื้อขายให้แก่ลูกค้า และรวบรวมข้อมูลการซื้อขาย ข้อมูลลูกค้า จัดทำเป็นเอกสารเพื่อเก็บไว้เป็นหลักฐาน

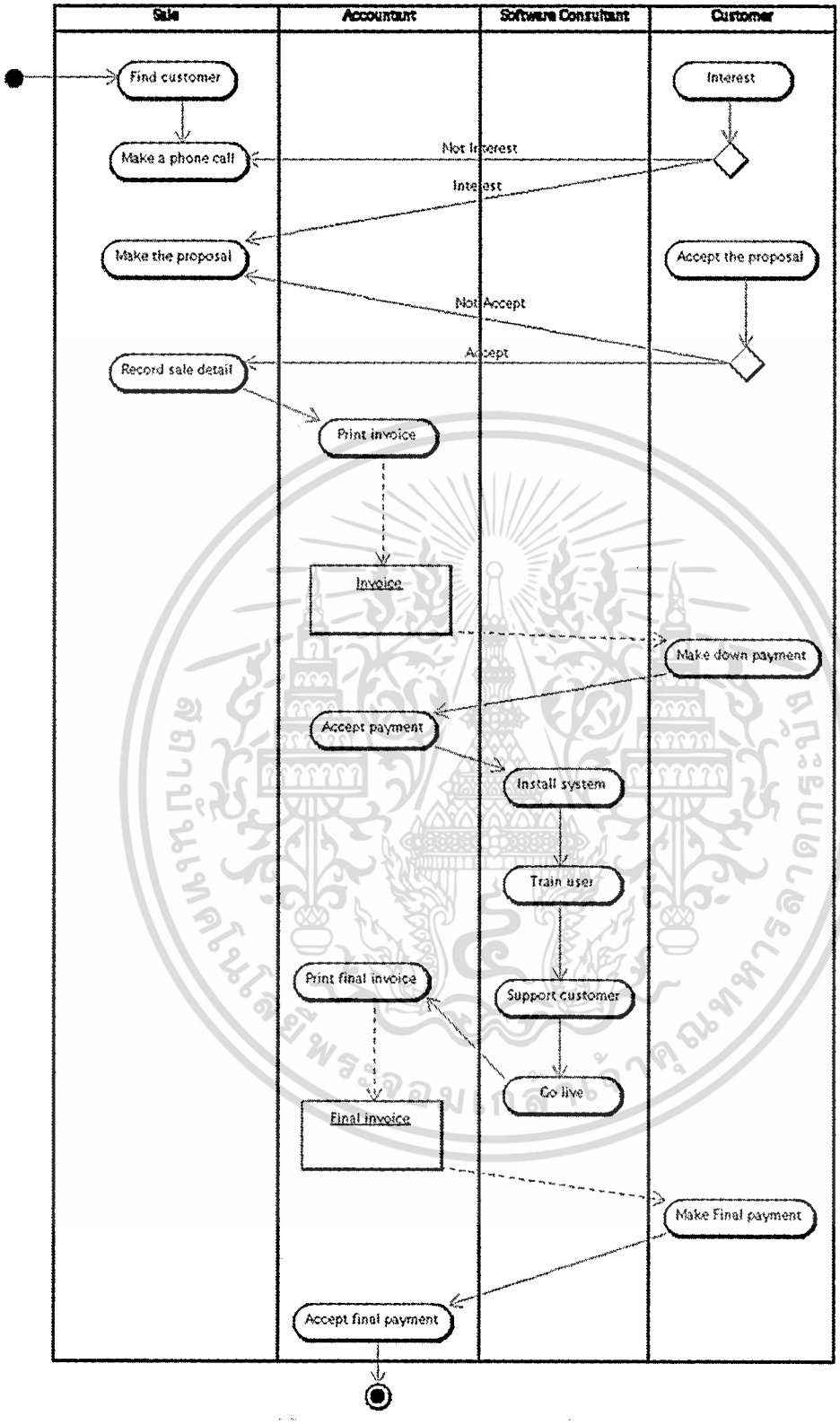
- การออกใบเสนอราคา

หลักจากที่ทางพนักงานขายและลูกค้าตกลงที่จะทำการซื้อขายเป็นที่เรียบร้อยแล้ว ฝ่ายบัญชีจะมีหน้าที่ออกใบเสนอราคาโดยดูจากเอกสารข้อมูลลูกค้าและข้อมูลการขายที่พนักงานขายทำไว้ จากนั้นจะทำใบเสนอราคาไปให้ลูกค้า และทำสัญญาจะซื้อจะขายกันในลำดับต่อไป

- การติดตั้งซอฟต์แวร์

เมื่อทำสัญญาจะซื้อจะขายเรียบร้อย ทางทีมงานให้คำปรึกษาจะเข้าไปยังบริษัทลูกค้าเพื่อติดตั้งซอฟต์แวร์ สอนการใช้งานระบบ และทำงานขึ้นระบบเป็นลำดับต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3-1 แสดงขั้นตอนการทำงานภายในองค์กร

3.2 แนวทางการแก้ปัญ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการดำเนินงานในปัจจุบัน การทำงานเกี่ยวกับการจัดเก็บข้อมูลการซื้อขายและข้อมูลลูกค้านั้น พบปัญหาว่าไม่สะดวกในการใช้งาน ทางบิษัทยังใช้ระบบเอกสารในการจัดเก็บ ซึ่งไม่มีความเป็นศูนย์กลางและเสี่ยงต่อการสูญหาย

ในด้านขั้นตอนของการหาลูกค้านั้น พบปัญหาว่าบางครั้งเสนอการขายได้ไม่ต้องกับความต้องการของลูกค้า มีการโทรกระจายไปเรื่อยไม่มีแบบแผน ยิ่งไปกว่านั้นปัจจุบันได้มีซอฟต์แวร์จัดการฝ่ายบุคคลถูกสร้างขึ้นมามากขึ้น จึงทำให้เกิดส่วนแบ่งทางการตลาดมากขึ้นไปด้วย

จากการสัมภาษณ์ผู้บริหารของบริษัทนั้น ผู้บริหารได้ตระหนักถึงปัญหาที่จะเกิดขึ้นในอนาคต เนื่องจากเอกสารการซื้อขายและข้อมูลลูกค้านั้นเพิ่มขึ้นเรื่อยๆ จึงต้องการที่จะให้มีระบบจัดเก็บข้อมูลการซื้อขายและข้อมูลลูกค้าที่ใช้ง่าย เก็บอยู่ในพื้นที่ส่วนกลางที่สามารถใช้ชุดคำสั่งเรียกดูได้ และต้องการระบบวิเคราะห์ข้อมูลการซื้อขายเพื่อนำมาใช้ประกอบการตัดสินใจ รวมทั้งสามารถนำข้อมูลไปใช้แนะนำลูกค้าในลักษณะของลูกค้าสัมพันธ์ได้ด้วย

โครงงานการศึกษาวิชาอิสระนี้ถูกพัฒนาเพื่อรองรับความต้องการของผู้บริการโดยตรง โดยจะแบ่งการทำงานของระบบออกเป็น 2 ส่วนหลัก คือ ส่วนของการจัดเก็บข้อมูลการซื้อขายและข้อมูลลูกค้า และส่วนของการวิเคราะห์ข้อมูลซึ่งประมวลผลออกมาในรูปผลการทำนายเพื่อนำไปใช้ในการประกอบการตัดสินใจ สามารถดูรายละเอียดของแต่ละส่วนได้ในบทถัดไป

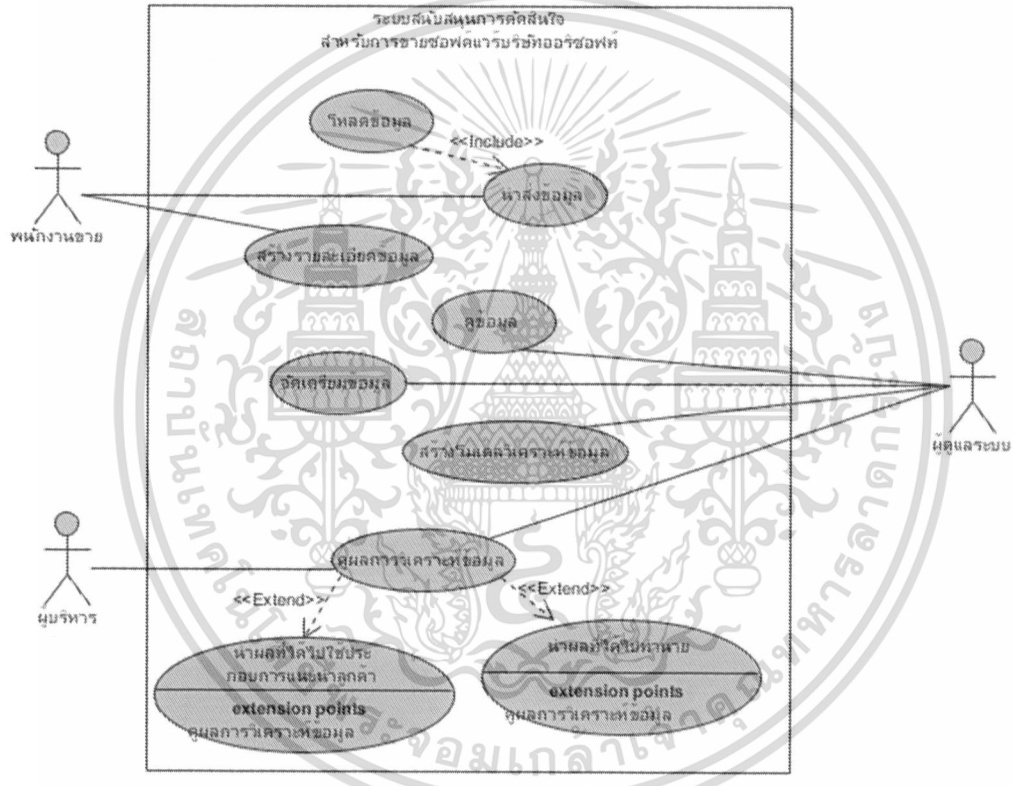
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การวิเคราะห์และออกแบบระบบใหม่

การศึกษาและวิเคราะห์การทำงานของระบบปัจจุบันได้ถูกนำมาวิเคราะห์และออกแบบระบบให้รองรับกับการทำงานรูปแบบใหม่ มีรายละเอียดดังนี้

4.1 ยูสเคสไดอะแกรม (Use Case Diagram)

4.1.1 แผนภาพยูสเคสไดอะแกรม (Use Case Diagram)



รูปที่ 4-1 แสดงยูสเคสไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.2 รายละเอียดยูสเคส

ตารางที่ 4.1 ยูสเคส โหลดข้อมูล

Use Case Name :	โหลดข้อมูล	ID : 01
Actors :	-	
Triggering Event :	เมื่อข้อมูลถูกส่งไปยังฐานข้อมูล	
Brief Description :	เมื่อข้อมูลทั้งหมดถูกส่งไปยังฐานข้อมูล หลังจากที่พนักงานขายทำการกดส่งผ่านข้อมูลแล้ว	
Related Use Cases :	นำส่งข้อมูล	
Pre-Conditions :	ต้องผ่านขั้นตอนของการส่งผ่านข้อมูลของพนักงานขาย	
Post-Conditions :	-	
Flow of Events :	Actor	System
	-	ระบบทำการโหลดข้อมูลไปยังฐานข้อมูล
Exception Conditions :	-	

ตารางที่ 4.2 ยูสเคส ส่งผ่านข้อมูล

Use Case Name :	นำส่งข้อมูล	ID : 02
Actors :	พนักงานขาย	
Triggering Event :	เมื่อพนักงานขายต้องการที่จะส่งข้อมูลไปยังฐานข้อมูล	
Brief Description :	เมื่อพนักงานขายต้องการที่จะส่งข้อมูลรายละเอียดการซื้อขายและรายละเอียดลูกค้าที่กรอกไปแล้วให้ไปอยู่ในยังฐานข้อมูล	
Related Use Cases :	โหลดข้อมูล, สร้างรายละเอียดข้อมูล	
Pre-Conditions :	พนักงานขายมีสิทธิ์การใช้งาน	
Post-Conditions :	-	
Flow of Events :	Actor	System

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 (ต่อ)

	พนักงานขายเลือกกดปุ่มบันทึกข้อมูล	ระบบทำการแปลงข้อมูลจากหน้าแอพพลิเคชั่นให้อยู่ในรูปของตารางฐานข้อมูล
Exception Conditions :	-	

ตารางที่ 4.3 ยูสเคส สร้างข้อมูล

Use Case Name :	สร้างรายละเอียดข้อมูล	ID : 03
Actors :	พนักงานขาย	
Triggering Event :	เมื่อพนักงานขายต้องการกรอกรายละเอียดข้อมูล	
Brief Description :	เมื่อพนักงานขายต้องการกรอกรายละเอียดข้อมูลการซื้อขายและข้อมูลลูกค้า รวมถึงข้อมูลอื่นๆเพื่อเก็บไว้เป็นหลักฐานในฐานข้อมูล	
Related Use Cases :	-	
Pre-Conditions :	พนักงานขายมีสิทธิ์การใช้งาน	
Post-Conditions :	-	
Flow of Events :	Actor	System
	พนักงานขายกรอกรายละเอียดข้อมูลการซื้อขายและข้อมูลลูกค้า	-
Exception Conditions :	-	

ตารางที่ 4.4 ยูสเคส แสดงข้อมูล

Use Case Name :	ดูข้อมูล	ID : 04
Actors :	ผู้ดูแลระบบ	
Triggering Event :	เมื่อต้องการดูข้อมูลต่างๆที่พนักงานขายใส่รายละเอียดเข้าไป	
Brief Description :	เมื่อต้องการดูข้อมูลต่างๆที่พนักงานขายใส่รายละเอียดเข้าไป อาจดูข้อมูลด้วยชุดคำสั่งข้อมูล	
Related Use Cases :	นำส่งข้อมูล	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.4 (ต่อ)

Pre-Conditions :	ผู้ดูแลระบบมีสิทธิ์การใช้งาน	
Post-Conditions :	ระบบแสดงรายละเอียดการซื้อขายและข้อมูลลูกค้า	
Flow of Events :	Actor	System
	ผู้ดูแลระบบเลือกไฟล์ฐานข้อมูลที่ถูกบันทึกไว้โดยพนักงานขาย	ระบบแสดงรายละเอียดการซื้อขายและข้อมูลลูกค้าขึ้นมา
Exception Conditions :	-	

ตารางที่ 4.5 ยูสเคส จัดเตรียมข้อมูล

Use Case Name :	จัดเตรียมข้อมูล	ID : 05
Actors :	ผู้ดูแลระบบ	
Triggering Event :	เมื่อต้องการที่จะเตรียมข้อมูลเพื่อใช้ในการวิเคราะห์	
Brief Description :	เมื่อต้องการที่จะเตรียมข้อมูลเพื่อใช้ในการวิเคราะห์ โดยผู้ดูแลระบบจะทำการแก้ไขข้อมูลที่ผิดพลาด ซ้ำซ้อน	
Related Use Cases :	-	
Pre-Conditions :	ผู้ดูแลระบบมีสิทธิ์การใช้งาน	
Post-Conditions :	-	
Flow of Events :	Actor	System
	1.ผู้ดูแลระบบสร้างรูปแบบของข้อมูลที่ไว้ใช้สำหรับการจัดเตรียมข้อมูล 3.ผู้ดูแลระบบแก้ไขข้อมูลที่ผิดพลาด ซ้ำซ้อน	2.ระบบช่วยตรวจจับข้อมูลที่ผิดพลาด ซ้ำซ้อน
Exception Conditions :	-	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 ยูสเคส สร้างตัวแบบข้อมูล

Use Case Name :	สร้างตัวแบบข้อมูล	ID : 06
Actors :	ผู้ดูแลระบบ	
Triggering Event :	เมื่อผู้ดูแลระบบต้องการที่จะสร้างตัวแบบข้อมูลเพื่อใช้ในการวิเคราะห์	
Brief Description :	เมื่อผู้ดูแลระบบต้องการที่จะสร้างตัวแบบข้อมูลเพื่อใช้ในการวิเคราะห์ โดยจะใช้อัลกอริที่มของ Neural Network มาช่วยวิเคราะห์ข้อมูล	
Related Use Cases :	จัดเตรียมข้อมูล	
Pre-Conditions :	ผู้ดูแลระบบสิทธิ์การใช้งาน	
Post-Conditions :	-	
Flow of Events :	Actor	System
	ผู้ดูแลระบบสร้างตัวแบบของข้อมูลที่ไว้ใช้สำหรับการวิเคราะห์ข้อมูล	ระบบประมวลผลข้อมูลเหล่านั้นและแสดงออกมาในรูปแบบของผลการทำนาย
Exception Conditions :	-	

ตารางที่ 4.7 ยูสเคส ดูผลการวิเคราะห์ข้อมูล

Use Case Name :	ดูผลการวิเคราะห์ข้อมูล	ID : 07
Actors :	ผู้ดูแลระบบ, ผู้บริหาร	
Triggering Event :	เมื่อผู้บริหารต้องการดูผลการวิเคราะห์ข้อมูล	
Brief Description :	เมื่อผู้บริหารต้องการดูการวิเคราะห์ข้อมูลในรูปแบบของผลการทำนายเพื่อไปประกอบการตัดสินใจ	
Related Use Cases :	สร้างโมเดล	
Pre-Conditions :	ผู้ดูแลระบบและผู้บริหารมีสิทธิ์การใช้งาน	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.7 (ต่อ)

Post-Conditions :	-	
Flow of Events :	Actor	System
	ผู้บริหารเลือกที่เมนูวิเคราะห์ผล	ระบบแสดงข้อมูลที่วิเคราะห์ออกมาอยู่ในรูปแบบของผลการทำนาย
Exception Conditions :	-	

ตารางที่ 4.8 ยูสเคส นำผลที่ได้ไปใช้ประกอบการแนะนำลูกค้า

Use Case Name :	นำผลที่ได้ไปใช้ประกอบการแนะนำลูกค้า	ID : 08
Actors :	ผู้บริหาร	
Triggering Event :	ผู้บริหารใช้ผลของการวิเคราะห์ข้อมูลไปใช้ในการช่วยหาลูกค้า	
Brief Description :	ผู้บริหารใช้ผลของการวิเคราะห์ข้อมูลไปใช้ในการหาลูกค้า แนะนำซอฟต์แวร์ที่เหมาะสมแก่องค์กรลูกค้า	
Related Use Cases :	ดูผลการวิเคราะห์ข้อมูล	
Pre-Conditions :	ผู้บริหารมีสิทธิ์การใช้งาน	
Post-Conditions :	-	
Flow of Events :	Actor	System
	1.ผู้บริหารเลือกที่เมนูวิเคราะห์ผล	2.ระบบแสดงข้อมูลที่วิเคราะห์ออกมาอยู่ในรูปแบบของผลการทำนาย
	3.ผู้บริหารนำข้อมูลนั้นไปใช้ในการแนะนำลูกค้า	
Exception Conditions :	-	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.9 ยูสเคส นำผลที่ได้ไปใช้ในการทำนาย

Use Case Name :	นำผลที่ได้ไปใช้ในการทำนาย	ID : 09
Actors :	ผู้บริหาร	
Triggering Event :	ผู้บริหารใช้ผลของการวิเคราะห์ข้อมูลไปใช้ในการทำนายยอดขาย	
Brief Description :	ผู้บริหารใช้ผลของการวิเคราะห์ข้อมูลไปใช้ในการทำนายยอดขายของซอฟต์แวร์ประเภทต่างๆ	
Related Use Cases :	ดูผลการวิเคราะห์ข้อมูล	
Pre-Conditions :	ผู้บริหารมีสิทธิ์การใช้งาน	
Post-Conditions :	-	
Flow of Events :	Actor	System
	1.ผู้บริหารเลือกที่เมนูวิเคราะห์ผล 3.ผู้บริหารนำข้อมูลนั้นไปใช้ในการทำนายการซื้อขาย	2.ระบบแสดงข้อมูลที่วิเคราะห์ออกมาอยู่ในนำผลที่ได้ไปใช้ในการทำนาย
Exception Conditions :	-	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 แอคติวิตี้ไดอะแกรม (Activity Diagram)

การวิเคราะห์และออกแบบระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์ ผู้พัฒนาได้ทำแอคติวิตี้ไดอะแกรม (Activity Diagram) มาใช้ในการอธิบายขั้นตอนการดำเนินงาน ลักษณะภาพรวมของระบบการทำงานหลักแบ่งออกเป็น 5 ส่วนคือ การสร้างฐานข้อมูลการซื้อขายและข้อมูลลูกค้า ขั้นตอนการจัดเตรียมข้อมูลเพื่อนำไปใช้ในการวิเคราะห์ ขั้นตอนการสร้างรูปแบบสำหรับการวิเคราะห์ ขั้นตอนการแนะนำลูกค้าใหม่ ขั้นตอนการทำนายผลการซื้อขาย ซึ่งการทำงานของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์จะมีขั้นตอนการทำงานดังนี้

1.พนักงานขายกรอกข้อมูลการซื้อขายและข้อมูลลูกค้าลงในแอพพลิเคชั่น

2.กดปุ่มบันทึกข้อมูลที่อยู่ในหน้าแอพลลิเคชั่น

3.เมื่อทำการบันทึกข้อมูลแล้ว ข้อมูลทั้งหมดจะถูกส่งไปยังฐานข้อมูล

4.ผู้ดูแลระบบจะเขียนโปรแกรมเพื่อดึงข้อมูลออกมาในรูปไฟล์ .csv

5.ผู้ดูแลระบบจะทำการจัดเตรียมข้อมูลเพื่อนำไปใช้ในการวิเคราะห์ในขั้นตอนนี้จะประกอบด้วยการแก้ไขข้อมูลที่ผิดพลาดการแก้ไขข้อมูลที่ซ้ำซ้อนให้ถูกต้องพร้อมนำไปวิเคราะห์

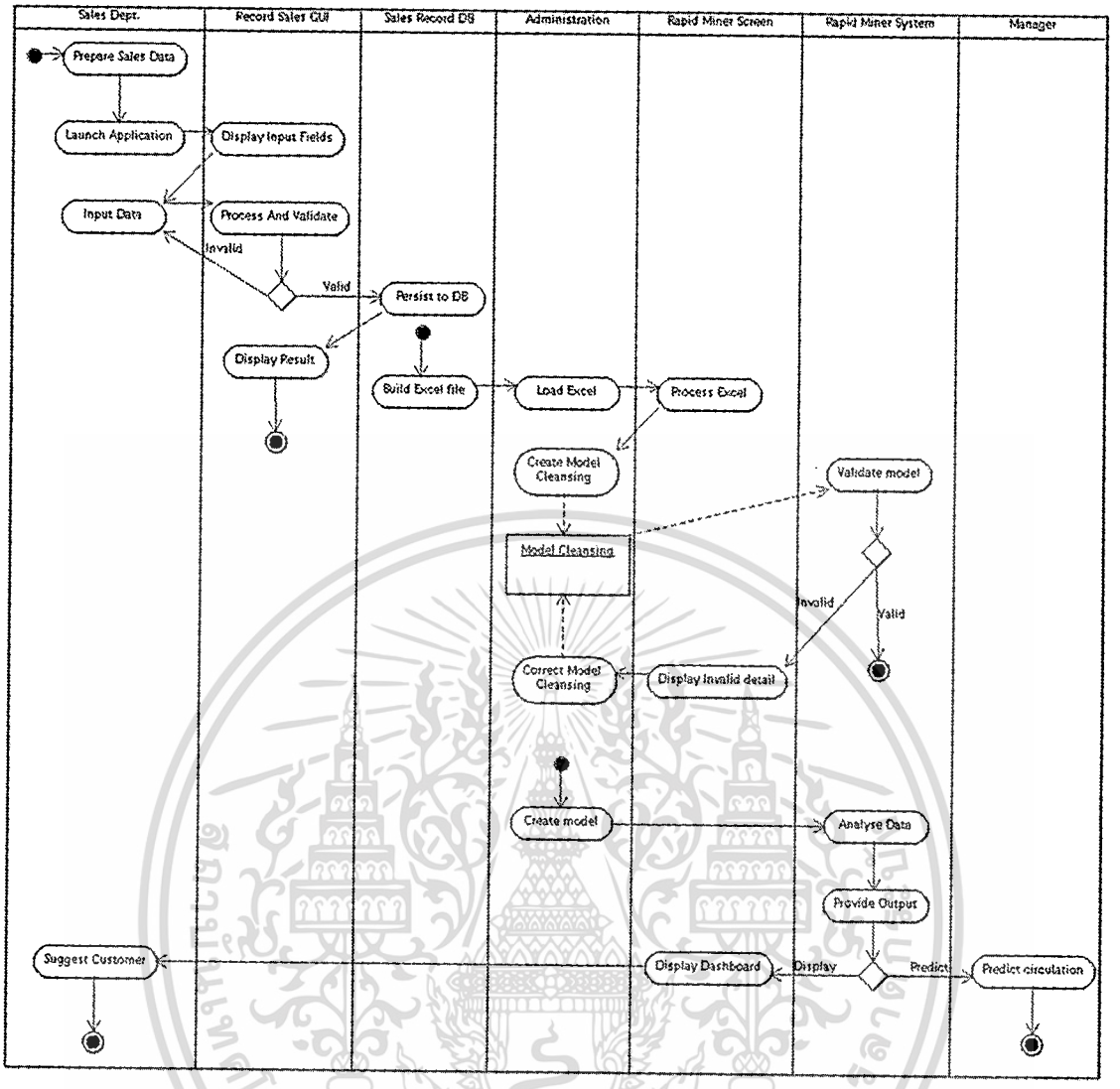
6.ผู้ดูแลระบบจะทำการสร้างรูปแบบข้อมูลเพื่อใช้ในการวิเคราะห์ด้วยโปรแกรม Rapidminer โดยในการศึกษาอิสระนี้จะ ใช้อัลกอริทึ่ม Neural Network มาใช้ในการวิเคราะห์

7.โหลดข้อมูลการซื้อขายและข้อมูลลูกค้าซึ่งเป็นไฟล์ .csv เข้ามายังโปรแกรม Rapidminer

8.โปรแกรม Rapidminer ทำการคำนวณผลการวิเคราะห์

9.ผู้บริหารนำผลที่ทำนายที่ได้นั้นไปใช้ในการประกอบการตัดสินใจและแนะนำลูกค้า

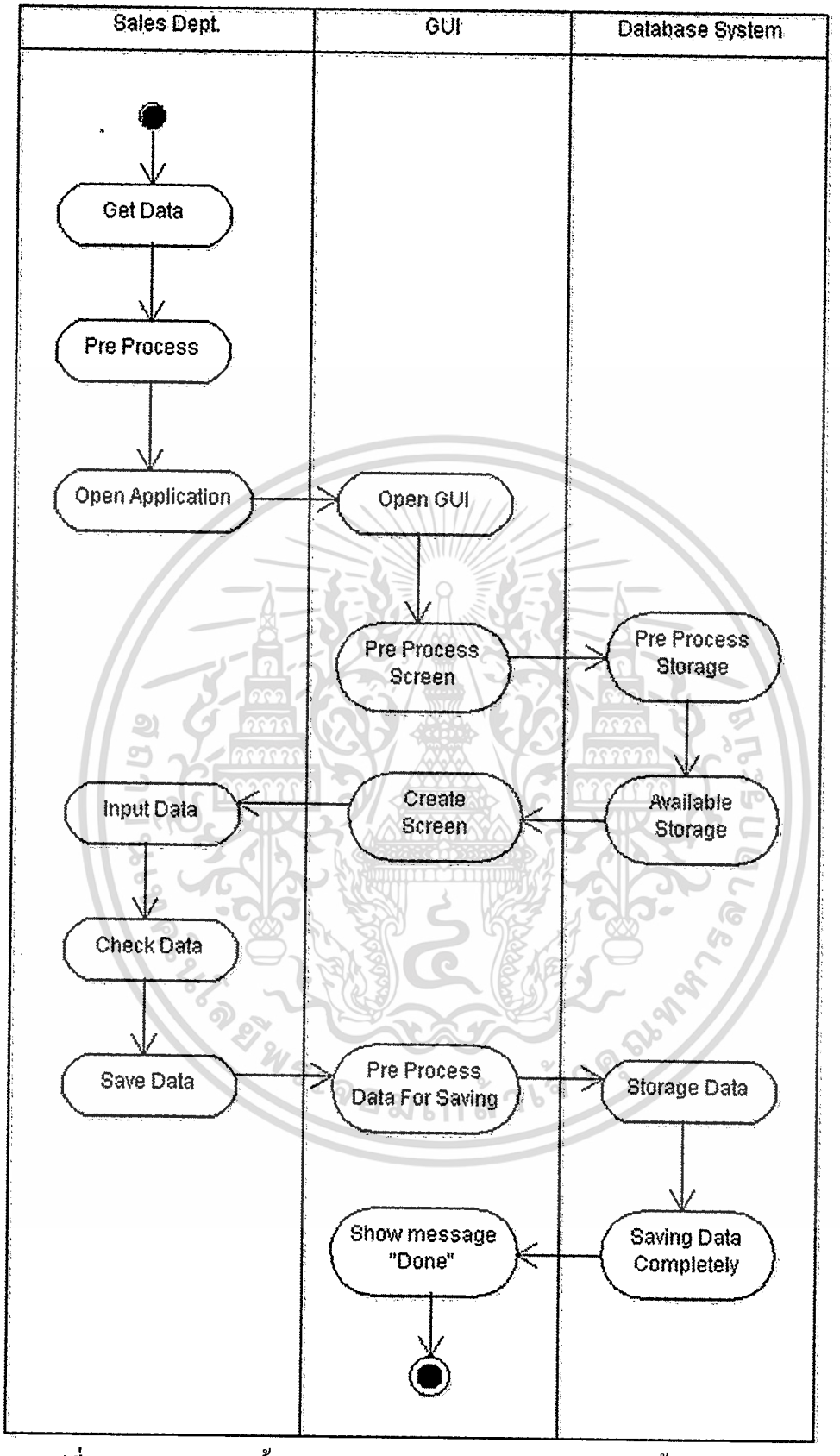
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-2 แสดงแอคติวิตี้ไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

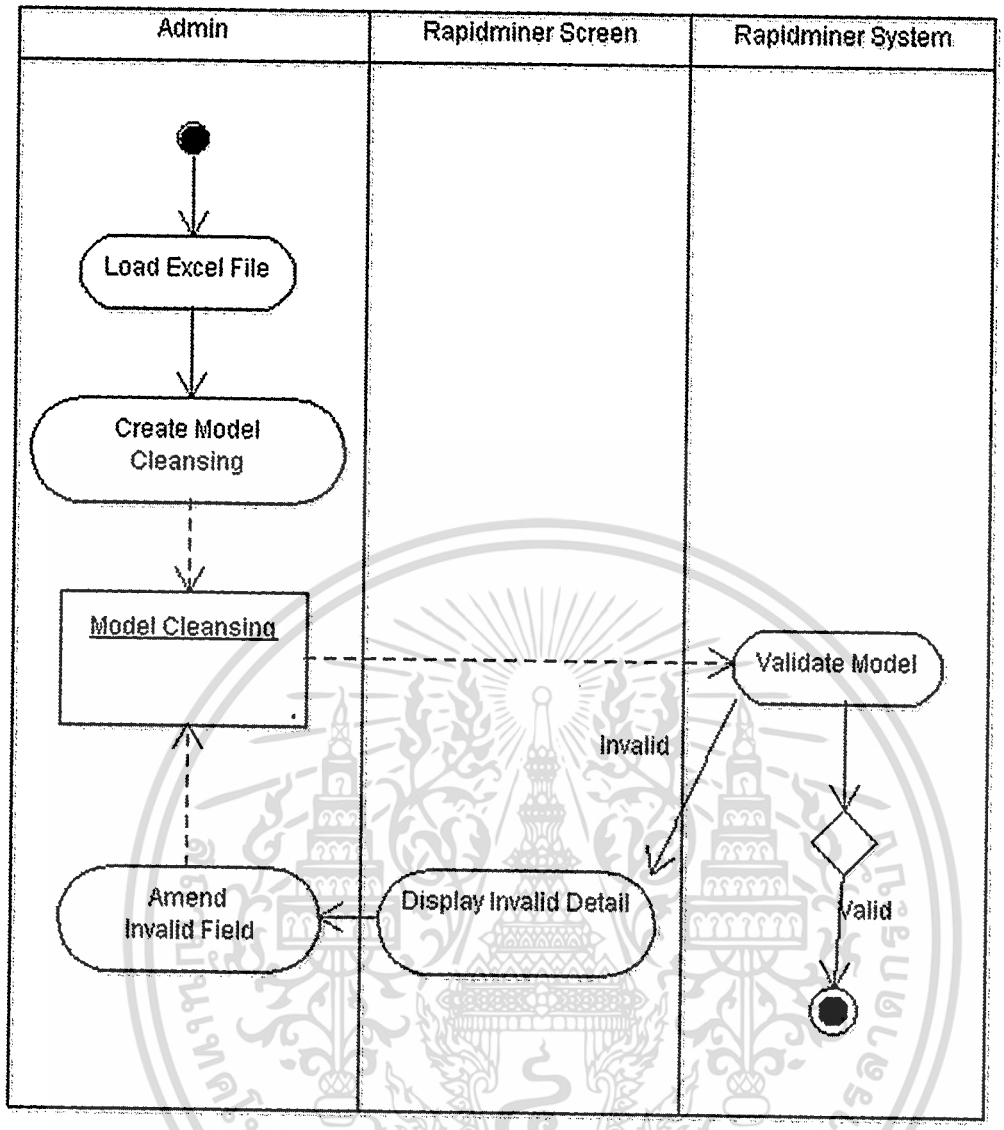
ทางผู้พัฒนาระบบ ได้ทำการแยกส่วนประกอบทั้ง 5 ส่วนของระบบออกมาในลักษณะของแอคติวิตี้ไดอะแกรม ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



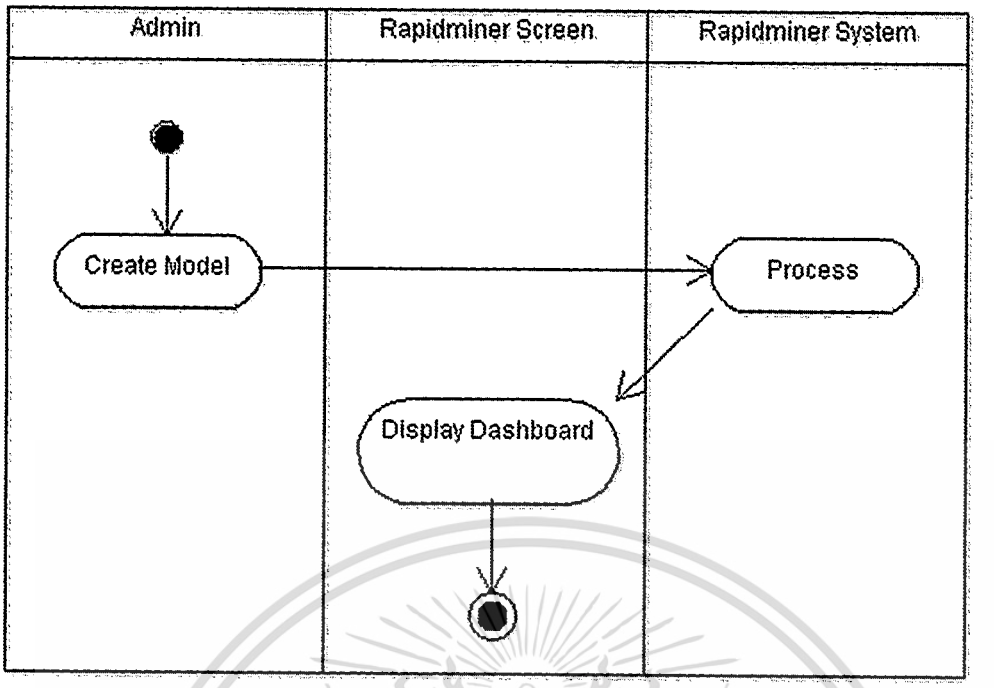
รูปที่ 4-3 แสดงแอคติวิตี้ไดอะแกรมของการสร้างฐานข้อมูลการซื้อขายและข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



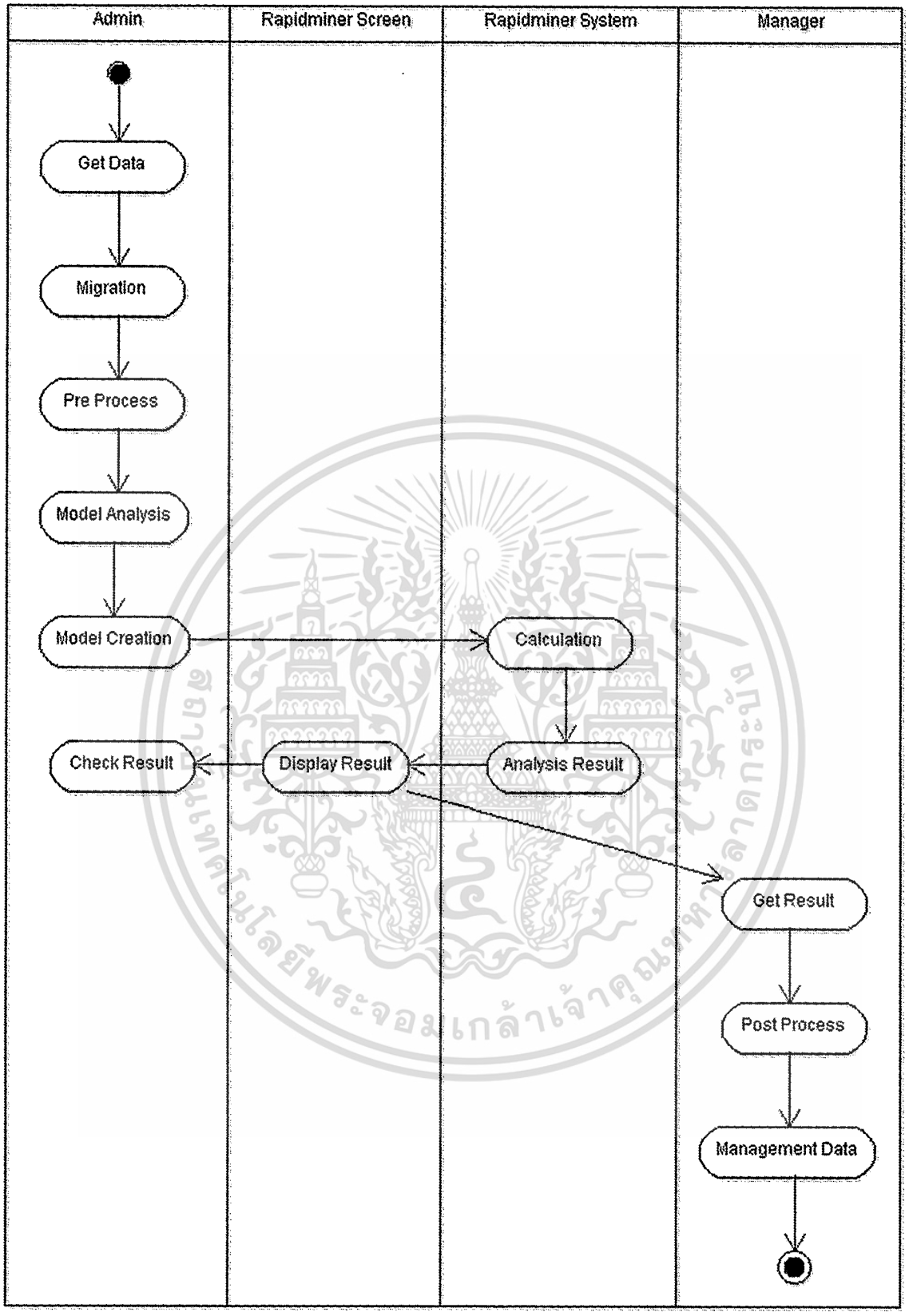
รูปที่ 4-4 แสดงแอคติวิตี้ไดอะแกรมของการจัดเตรียมข้อมูลเพื่อนำไปใช้ในการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



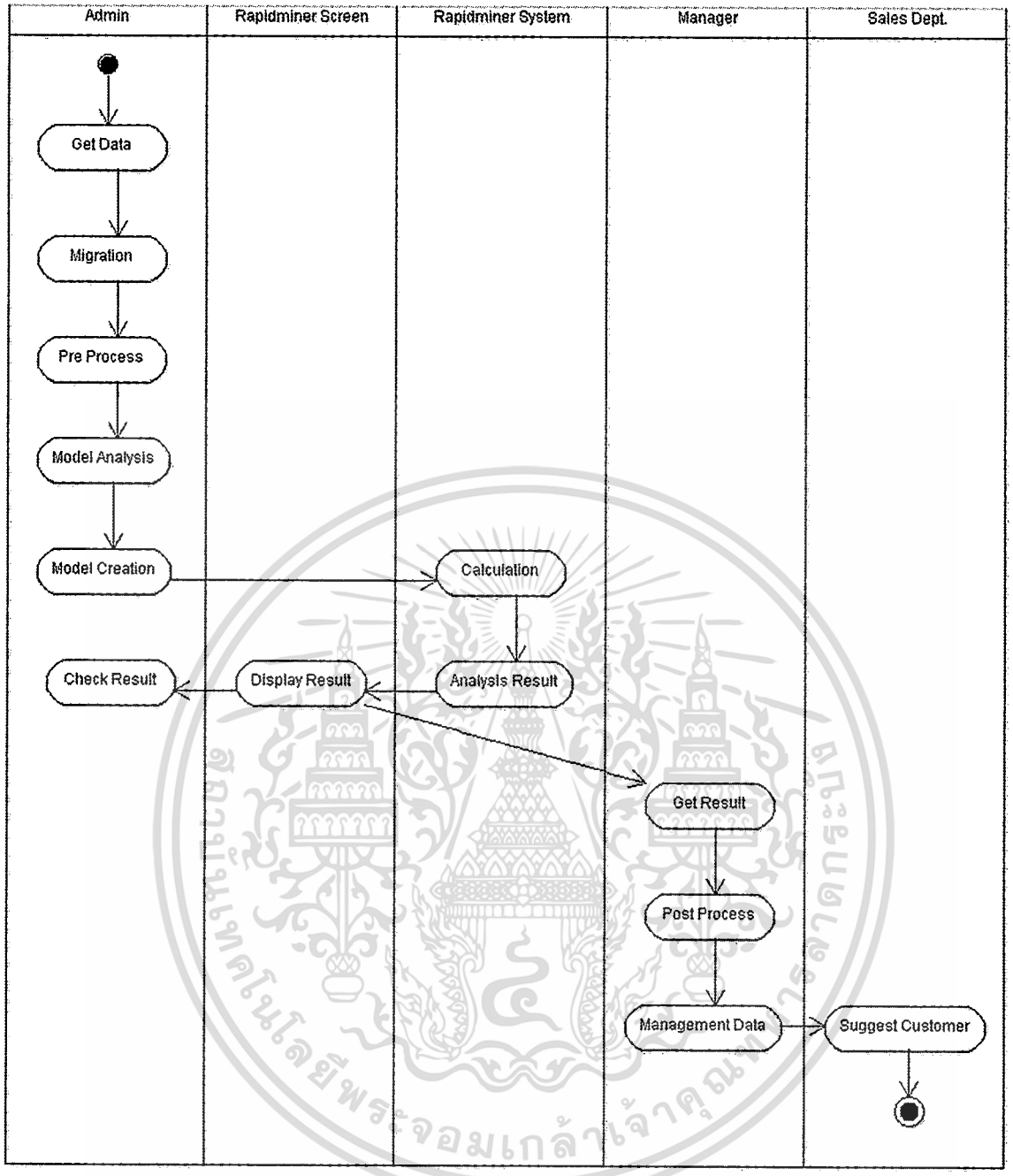
รูปที่ 4-5 แสดงแอคติวิตี้ไดอะแกรมของการทำนายและแสดงผลสัดส่วนด้วยแผนภาพวงกลม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-6 แสดงแอคติวิตี้ไดอะแกรมของการทำนายผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

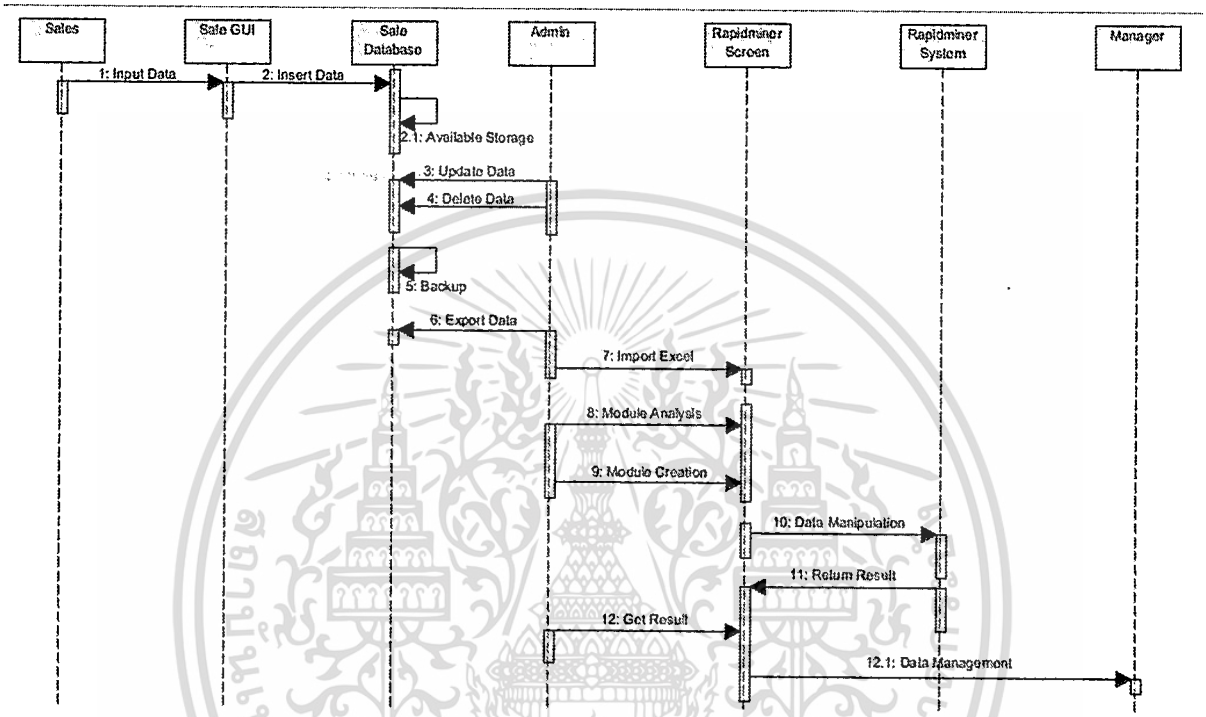


รูปที่ 4-7 แสดงแอคติวิตี้ไดอะแกรมของการแนะนำลูกค้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

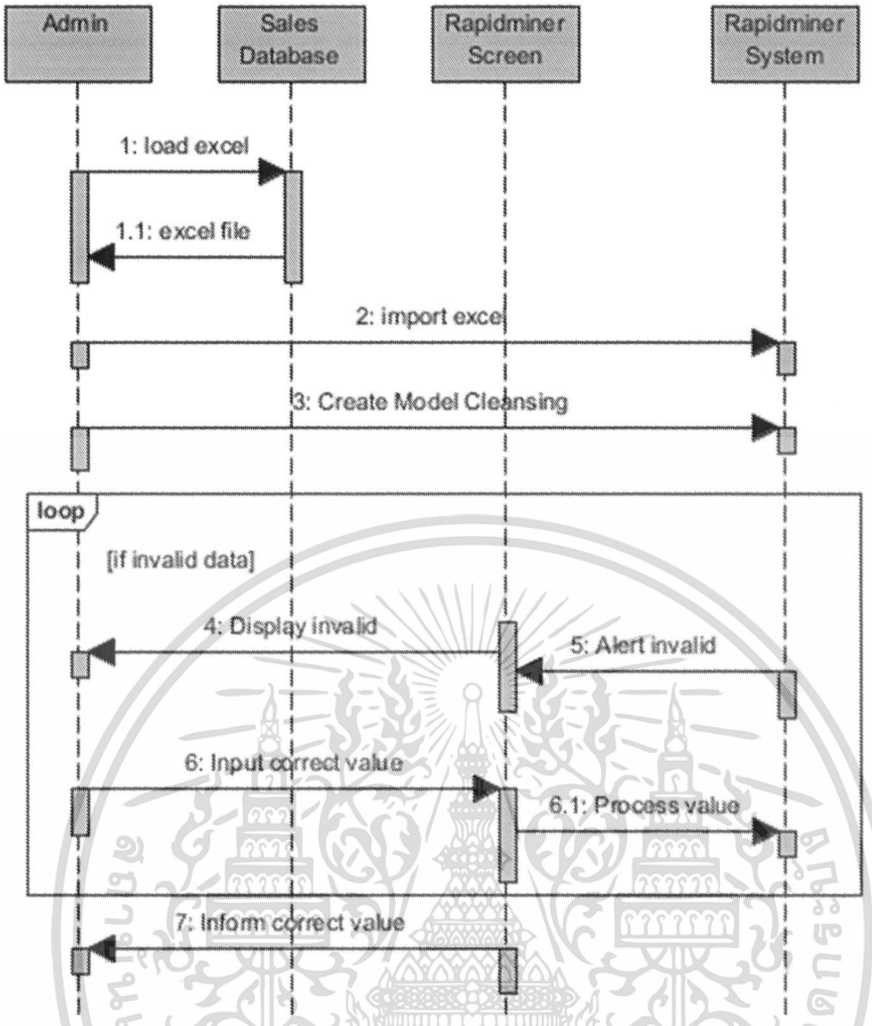
4.3 ซีเควนไดอะแกรม (Sequence Diagram)

การวิเคราะห์และออกแบบระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์ ผู้พัฒนาได้ทำซีเควนไดอะแกรม (Sequence Diagram) มาใช้ในการอธิบายขั้นตอนการดำเนินงาน โดยมีขั้นตอนการทำงานเป็นดังรูป



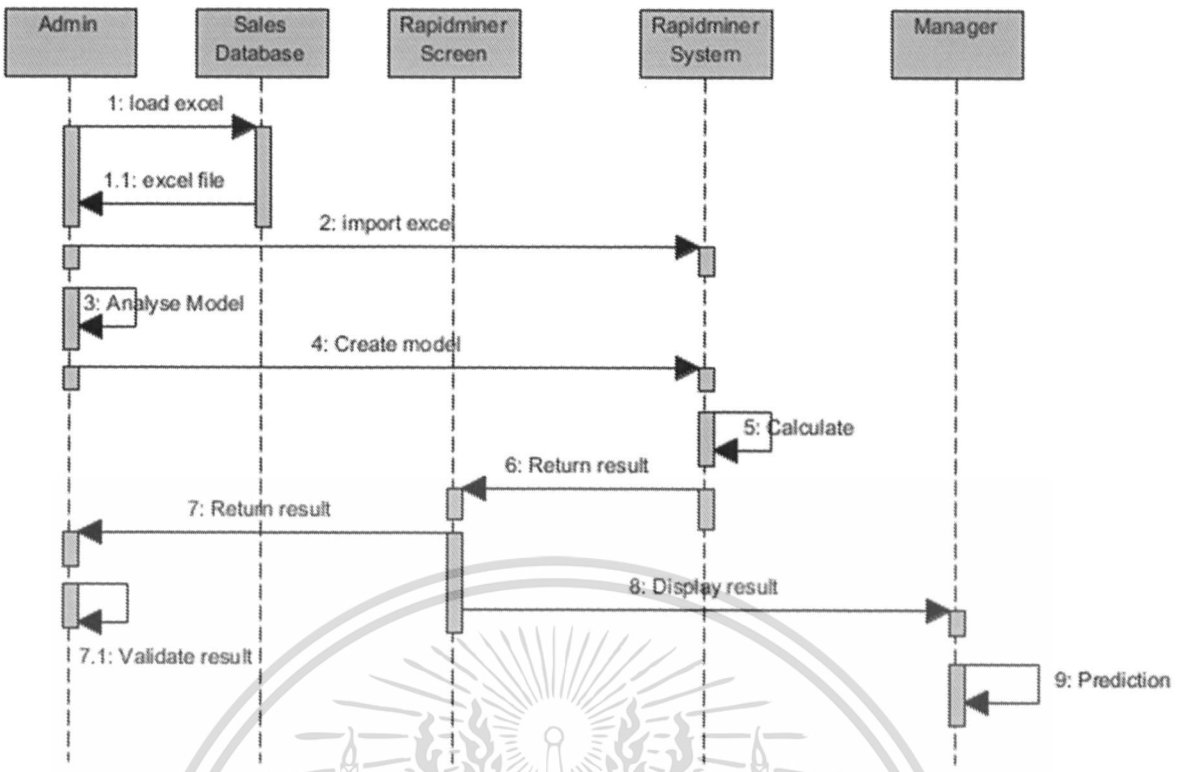
รูปที่ 4-8 แสดงซีเควนไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-9 แสดงซีเควนไดอะแกรมของการจัดเตรียมข้อมูลเพื่อนำไปใช้ในการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

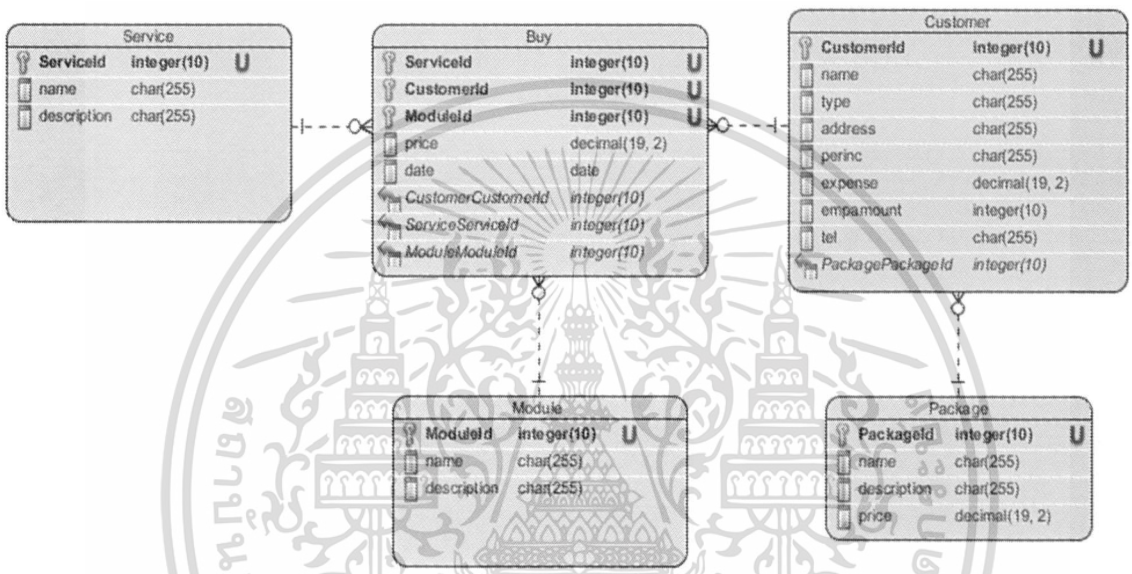


รูปที่ 4-10 แสดงซีเควนไดอะแกรมของการนำข้อมูลที่วิเคราะห์ได้ไปทำนายผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
การออกแบบฐานข้อมูล

5.1 อีอาร์ไดอะแกรม (ER Diagram)



รูปที่ 5-1 แสดงอีอาร์ไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

ตารางที่ 5.1 ความหมายของเอนทิตีระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

ชื่อเอนทิตี	ความหมาย
Customer	เอนทิตีของลูกค้า
Buy	เอนทิตีของการขาย
Service	เอนทิตีของรูปแบบบริการ
Module	เอนทิตีของโมดูล
Package	เอนทิตีของแพคเกจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 พจนานุกรมข้อมูล

ตารางที่ 5.2 พจนานุกรมข้อมูลของเอนทิตีของลูกค้า (Customer)

ชื่อแอตทริบิวต์	ความหมาย	ชนิดข้อมูล	จำเป็น	คีย์	ตาราง ที่เกี่ยวข้อง
CustomerId	รหัสลูกค้า	integer(10)	Y	PK	Buy
name	ชื่อบริษัทลูกค้า	char(255)	Y		
type	ประเภทอุตสาหกรรม	char(255)	Y		
address	ที่อยู่บริษัท	char(255)	Y		
perinc	ชื่อบุคคลที่ติดต่อ	char(255)	Y		
expense	ทุนจดทะเบียน โดยประมาณ	decimal(19,2)	Y		
empamount	จำนวนพนักงาน	integer(10)	Y		
tel	เบอร์โทรศัพท์	char(255)	Y		
PackageId	รหัสแพคเกจ	integer(10)	Y	FK	Package

ตารางที่ 5.3 พจนานุกรมข้อมูลของเอนทิตีของการขาย (Buy)

ชื่อแอตทริบิวต์	ความหมาย	ชนิดข้อมูล	จำเป็น	คีย์	ตาราง ที่เกี่ยวข้อง
ServiceId	รหัสรูปแบบบริการ	integer(10)	Y	PK	
CustomerId	รหัสลูกค้า	integer(10)	Y		
ModuleId	รหัสโมดูล	integer(10)	Y		
price	ราคา	decimal(19,2)	Y		
date	วันที่ทำการซื้อขาย	date	Y		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.3 พจนานุกรมข้อมูลของเอนทิตีของการขาย (Buy) (ต่อ)

date	วันที่ทำการซื้อขาย	date	Y		
ชื่อแอตทริบิวต์	**ความหมาย**	**ชนิดข้อมูล**	**จำเป็น**	**คีย์**	**ตาราง ที่เกี่ยวข้อง**
ServiceId	รหัสรูปแบบบริการ	integer(10)	Y	FK	Service
CustomerId	รหัสลูกค้า	integer(10)	Y	FK	Customer
ModuleId	รหัสโมดูล	integer(10)	Y	FK	Module

ตารางที่ 5.4 พจนานุกรมข้อมูลของเอนทิตีของรูปแบบบริการ (Service)

ชื่อแอตทริบิวต์	**ความหมาย**	**ชนิดข้อมูล**	**จำเป็น**	**คีย์**	**ตาราง ที่เกี่ยวข้อง**
ServiceId	รหัสรูปแบบบริการ	integer(10)	Y	PK	Buy
name	ชื่อรูปแบบบริการ	char(255)	Y		
description	คำอธิบายรูปแบบบริการ	char(255)	Y		

ตารางที่ 5.5 พจนานุกรมข้อมูลของเอนทิตีของโมดูล (Module)

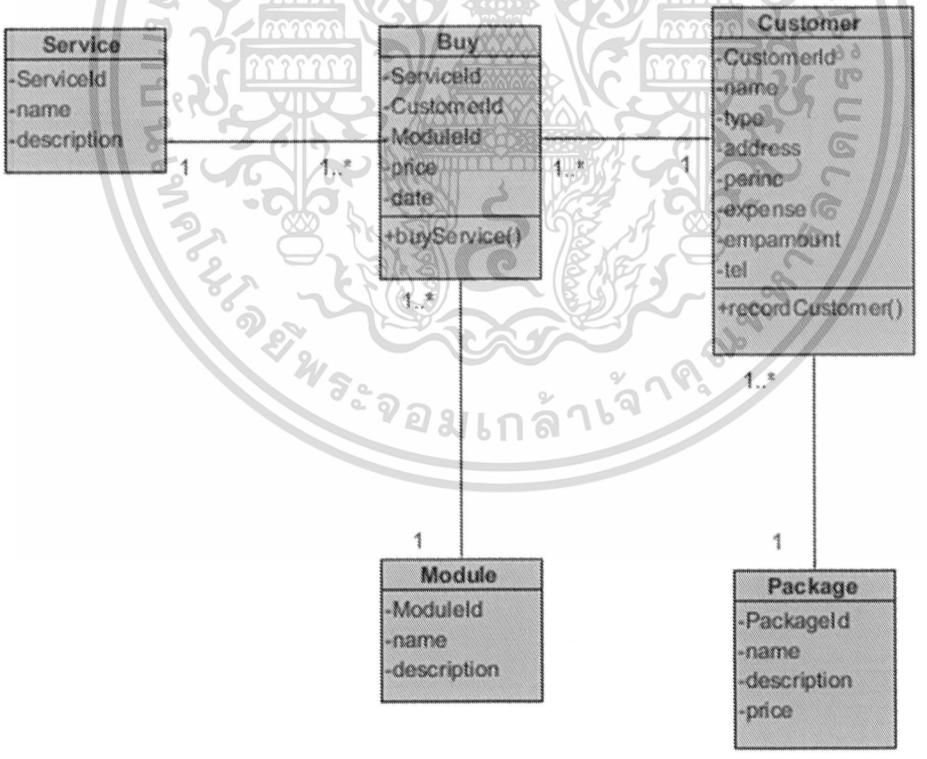
ชื่อแอตทริบิวต์	**ความหมาย**	**ชนิดข้อมูล**	**จำเป็น**	**คีย์**	**ตาราง ที่เกี่ยวข้อง**
ModuleId	รหัสโมดูล	integer(10)	Y	PK	Buy
name	ชื่อโมดูล	char(255)	Y		
description	คำอธิบายโมดูล	char(255)	Y		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.6 พจนานุกรมข้อมูลของเอนทิตีของแพคเกจ (Package)

ชื่อแอตทริบิวต์	ความหมาย	ชนิดข้อมูล	จำเป็น	คีย์	ตาราง ที่เกี่ยวข้อง
PackageId	รหัสแพคเกจ	integer(10)	Y	PK	Customer
name	ชื่อแพคเกจ	char(255)	Y		
description	คำอธิบายแพคเกจ	char(255)	Y		
price	ราคาของแพคเกจ	decimal(19,2)	Y		

5.3 คลาสไดอะแกรม (Class Diagram)



รูปที่ 5-1 แสดงคลาสไดอะแกรมของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

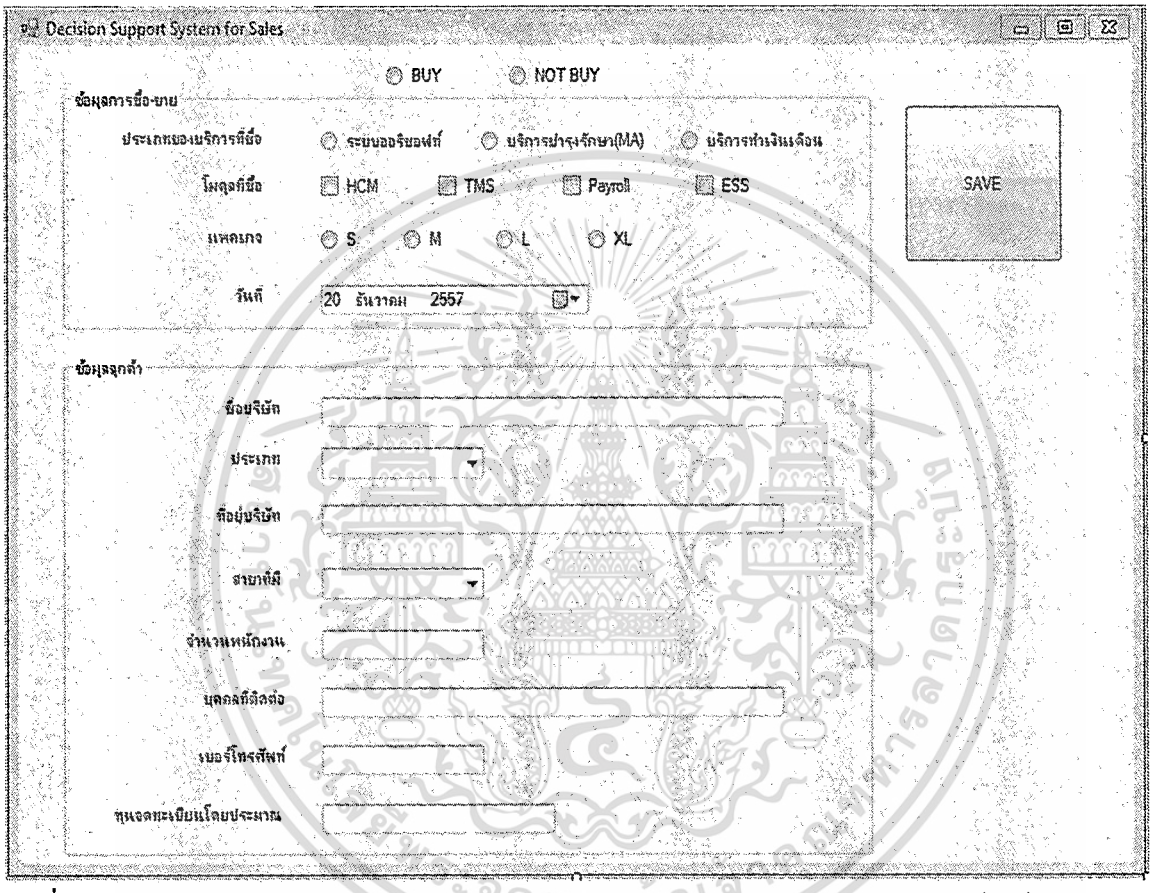
ตารางที่ 5.7 ความหมายคลาสของระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์

ชื่อคลาส	ความหมาย
Customer	คลาสของลูกค้า
Buy	คลาสของการขาย
Service	คลาสของรูปแบบบริการ
Module	คลาสของโมดูล
Package	คลาสของแพคเกจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6
การออกแบบส่วนต่อประสานผู้ใช้

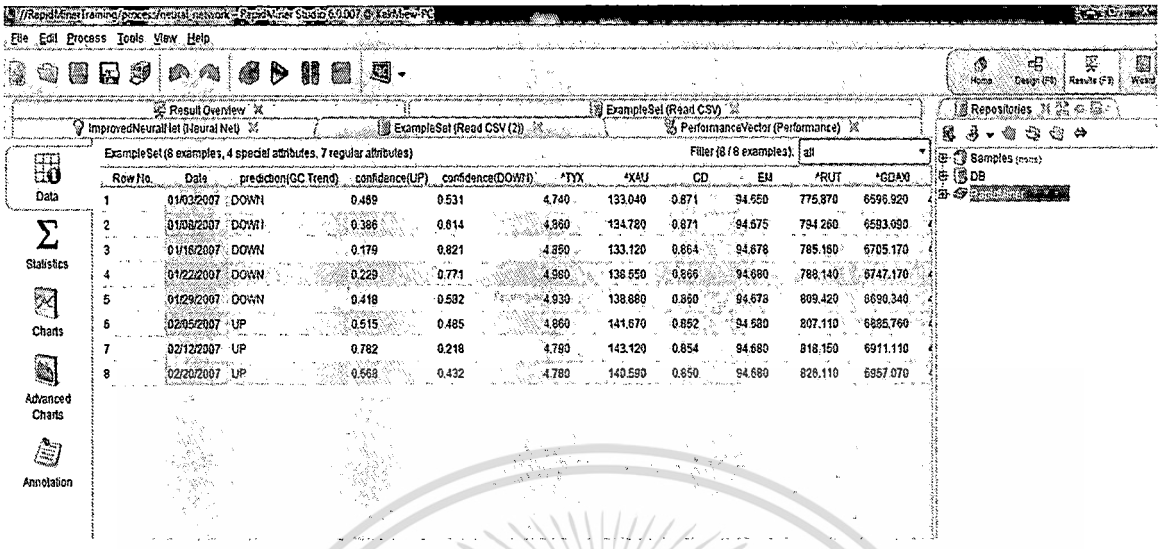
6.1 ส่วนต่อประสานผู้ใช้สำหรับใส่รายละเอียดข้อมูล



รูปที่ 6-1 แสดงส่วนต่อประสานผู้ใช้ระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์ในส่วนของการใส่ข้อมูลการซื้อขายและข้อมูลลูกค้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.2 ส่วนต่อประสานผู้ใช้สำหรับการวิเคราะห์ข้อมูล



Row No.	Date	prediction(GC Trend)	confidence(UP)	confidence(DOWN)	^TYX	^X4U	GD	EM	^RUT	^GDAXI
1	01/03/2007	DOWN	0.469	0.531	4.740	133.040	0.871	94.650	775.870	6596.920
2	01/08/2007	DOWN	0.386	0.614	4.860	134.780	0.871	94.675	794.260	6593.090
3	01/16/2007	DOWN	0.179	0.821	4.850	133.120	0.864	94.678	785.160	6705.170
4	01/22/2007	DOWN	0.229	0.771	4.980	138.550	0.866	94.680	788.140	6747.170
5	01/29/2007	DOWN	0.418	0.582	4.930	138.880	0.860	94.678	809.420	6690.340
6	02/05/2007	UP	0.515	0.485	4.860	141.670	0.852	94.680	807.110	6885.760
7	02/12/2007	UP	0.782	0.218	4.790	143.120	0.854	94.680	818.150	6911.110
8	02/20/2007	UP	0.568	0.432	4.780	140.590	0.850	94.680	826.110	6957.070

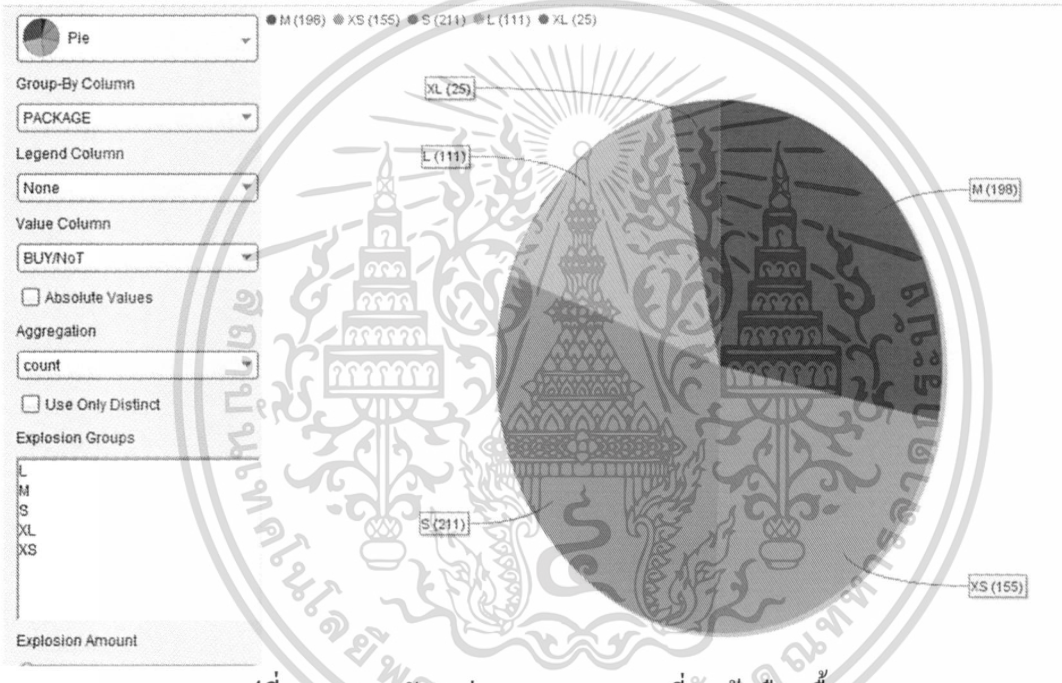
รูปที่ 6-2 แสดงส่วนต่อประสานผู้ใช้ระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์ในส่วนของการวิเคราะห์ข้อมูลและแสดงผลลัพธ์ด้วยการแสดงผลการทำนาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7
สรุปผลและข้อเสนอแนะ

7.1 สรุปโครงการ

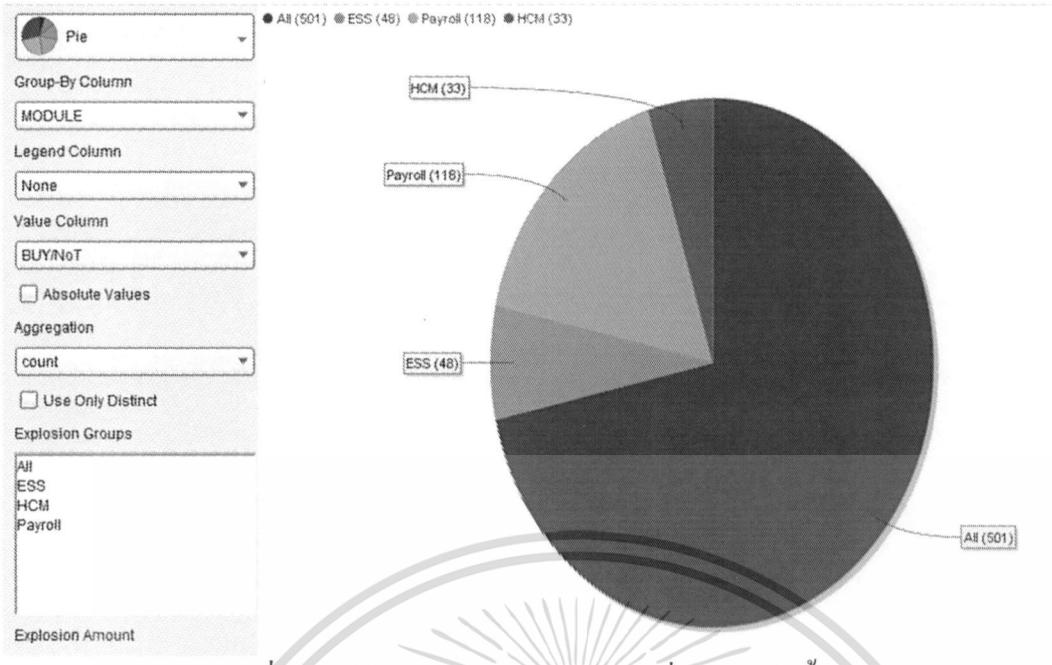
โครงงานศึกษาวิชาอิสระนี้ เป็นการนำข้อมูลการซื้อขายซอฟต์แวร์ของบริษัทออริซอฟท์มา เพื่อวิเคราะห์ โดยใช้เทคนิคต่างๆด้วยเครื่องมือ Rapidminer Studio โดยได้ผลออกมาดังนี้



รูปที่ 7-1 แสดงอัตราส่วนตาม Package ที่ลูกค้าเลือกซื้อ

รูปกราฟวงกลมแสดงอัตราส่วนการซื้อขนาดของ package ซึ่งจะเห็นได้ว่า package ที่ขายดีที่สุดจะเป็น package ขนาดเล็ก ที่ใช้กับองค์กรที่มีขนาดจำนวนพนักงาน 300-500 คน ดังนั้น ในส่วนของเรื่องการทำตลาด จึงควรสนใจองค์กรที่มีจำนวนพนักงานที่ไม่มากนัก

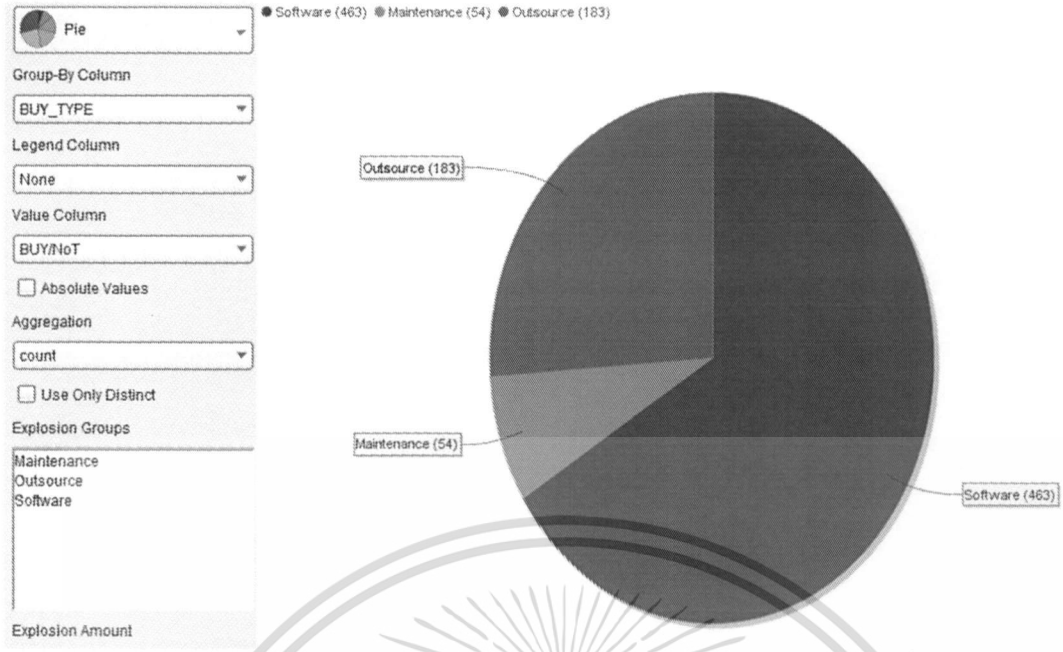
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7-2 แสดงอัตราส่วนตาม Module ที่ลูกค้าเลือกซื้อ

รูปกราฟวงกลมแสดงอัตราส่วน module ที่ซื้อ ซึ่งจะเห็นได้ว่า module ที่ขายดีที่สุดคือการซื้อทุก module พร้อมกัน

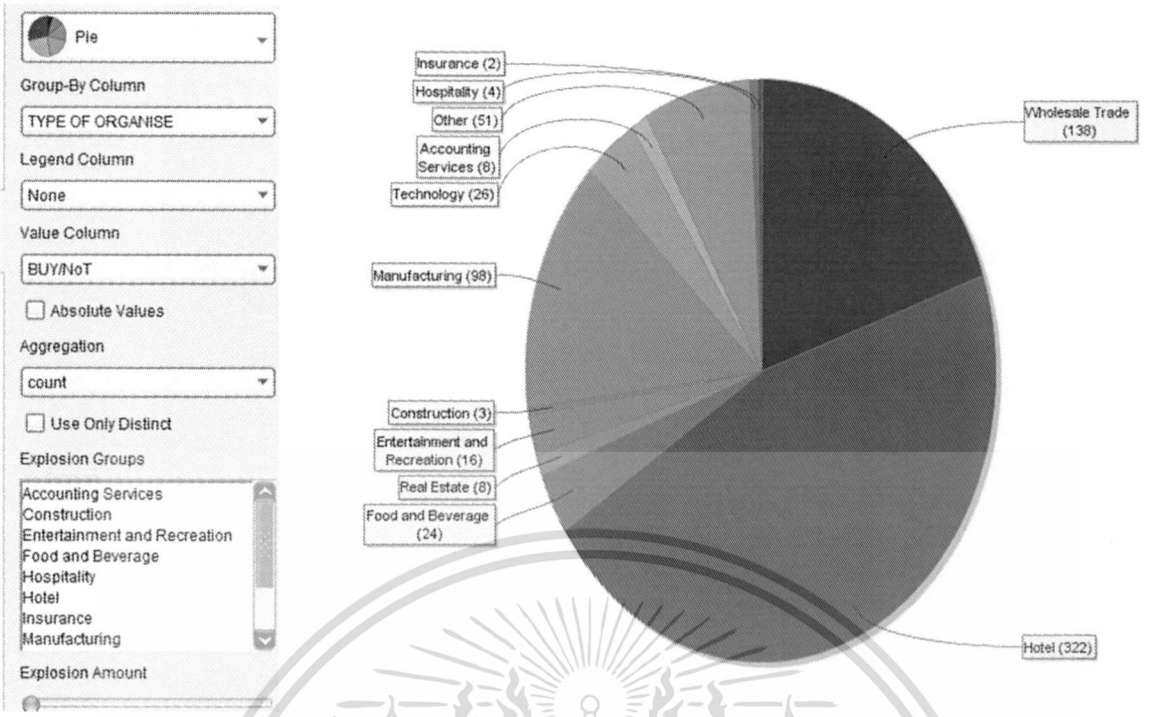
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7-3 แสดงอัตราส่วนตามประเภทของซอฟต์แวร์ที่ลูกค้าเลือกซื้อ

รูปกราฟวงกลมแสดงอัตราส่วน ประเภทของซอฟต์แวร์ที่ลูกค้าเลือกซื้อ ซึ่งจะเห็นได้ว่าลูกค้าส่วนใหญ่มักจะนิยมซื้อซอฟต์แวร์ที่ใช้สำหรับติดตั้งที่สถานที่ทำงานลูกค้าเอง แต่ลูกค้ามักจะไม่ซื้อการ maintenence จากบริษัทเท่าไหร่นัก อาจเป็นเพราะลูกค้ามีความมั่นใจว่าซอฟต์แวร์ออริซอฟท์สามารถทำงานได้อย่างต่อเนื่อง ไม่เกิดปัญหามากนัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7-4 แสดงอัตราส่วนตามประเภทของลูกค้า

รูปกราฟวงกลมแสดงอัตราส่วนตามประเภทธุรกิจของลูกค้า ซึ่งจะเห็นได้ว่า ซอฟต์แวร์ของออริซอฟท์นั้น เป็นที่นิยมกับกลุ่มลูกค้าธุรกิจโรงแรมมากที่สุด อาจด้วยตัวซอฟต์แวร์เองสามารถซัพพอร์ตตารางการทำงานแบบกะของพนักงานในโรงแรมได้ มีความยืดหยุ่น พนักงานฝ่ายบุคคลสามารถกำหนดกะการทำงานของพนักงานในโรงแรมได้เองตลอดเวลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Growth) | AssociationRules (Create Association Rules)

No.	Premises	Conclusion	Su...	Confid...	LaPl...	Gain	p-s	Lift	Con...
85	Pay	HCM	0.880	0.898	0.949	-1.08	-0.00	0.998	0.980
96	HCM	Pay	0.880	0.978	0.989	-0.92	-0.00	0.998	0.900
71	Pay	TMS	0.800	0.816	0.909	-1.16	0.016	1.020	1.089
97	TMS	Pay	0.800	1	1	-0.80	0.016	1.020	∞
38	Pay	HCM, TMS	0.740	0.755	0.879	-1.22	0.015	1.020	1.062
72	HCM	TMS	0.740	0.822	0.916	-1.06	0.020	1.028	1.125
73	HCM	Pay, TMS	0.740	0.822	0.916	-1.06	0.020	1.028	1.125
77	Pay, HCM	TMS	0.740	0.841	0.926	-1.02	0.036	1.051	1.257
90	TMS	HCM	0.740	0.925	0.967	-0.86	0.020	1.028	1.333
91	TMS	Pay, HCM	0.740	0.925	0.967	-0.86	0.036	1.051	1.600
92	Pay, TMS	HCM	0.740	0.925	0.967	-0.86	0.020	1.028	1.333
102	HCM, TMS	Pay	0.740	1	1	-0.74	0.015	1.020	∞
98	ESS	Pay	0.280	1	1	-0.28	0.006	1.020	∞
93	ESS	HCM	0.260	0.929	0.984	-0.30	0.008	1.032	1.400
94	ESS	Pay, HCM	0.260	0.929	0.984	-0.30	0.014	1.055	1.680
95	Pay, ESS	HCM	0.260	0.929	0.984	-0.30	0.008	1.032	1.400
103	HCM, ESS	Pay	0.260	1	1	-0.26	0.005	1.020	∞
81	ESS	TMS	0.240	0.857	0.969	-0.32	0.016	1.071	1.400
82	ESS	Pay, TMS	0.240	0.857	0.969	-0.32	0.016	1.071	1.400
83	Pay, ESS	TMS	0.240	0.857	0.969	-0.32	0.016	1.071	1.400
107	TMS, ESS	Pay	0.240	1	1	-0.24	0.005	1.020	∞
67	ESS	HCM, TMS	0.220	0.786	0.953	-0.34	0.013	1.062	1.213
68	ESS	Pay, HCM, TMS	0.220	0.786	0.953	-0.34	0.013	1.062	1.213
69	Pay, ESS	HCM, TMS	0.220	0.786	0.953	-0.34	0.013	1.062	1.213
78	HCM, ESS	TMS	0.220	0.846	0.966	-0.30	0.012	1.058	1.300
79	HCM, ESS	Pay, TMS	0.220	0.846	0.966	-0.30	0.012	1.058	1.300
80	Pay, HCM, ESS	TMS	0.220	0.846	0.966	-0.30	0.012	1.058	1.300
87	TMS, ESS	HCM	0.220	0.917	0.984	-0.26	0.004	1.019	1.200

รูปที่ 7-5 แสดงความสัมพันธ์ทั้งหมดของซอฟต์แวร์ที่ลูกค้ามักเลือกซื้อพร้อมกัน

เมื่อนำข้อมูลการซื้อขายซอฟต์แวร์มาวิเคราะห์โดยใช้กฎคสามสัมพันธ์ (Association Rule) แล้ว เราจะทำการดูผลจากค่า Support และค่า Confident ประกอบกัน ซึ่งถ้า 2 ตัวนี้มีค่าสูง นั่นหมายความว่าสินค้าหรือซอฟต์แวร์โมดูลดังกล่าวมักจะถูกซื้อพร้อมกัน

จากรูปผลลัพธ์ด้านบน จะเห็นได้ว่าบรรทัดที่ 2 นั้น จะมีค่า Support และค่า Confident มากที่สุด จึงสามารถสรุปได้ว่าลูกค้ามักซื้อโมดูล HCM และ Pay พร้อมกันมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ExampleSet (26 examples, 3 special attributes, 15 regular attributes)

Row No.	CUS_ID	Manufacturi...	cluster	TMS	Pay	HCM	ESS	
1	10001	1	cluster_1	1	1	1	1	0
2	10008	1	cluster_1	0	1	1	1	0
3	10009	1	cluster_1	0	1	1	1	0
4	10011	1	cluster_1	1	1	1	1	0
5	10012	0	cluster_1	0	1	1	0	0
6	10013	0	cluster_1	0	1	1	0	0
7	10016	1	cluster_1	1	1	1	1	0
8	10017	0	cluster_1	1	1	1	0	0
9	10018	0	cluster_2	1	1	1	0	0
10	10019	0	cluster_0	0	1	0	0	0
11	10023	1	cluster_1	1	1	1	1	0
12	10024	0	cluster_1	1	1	1	0	0
13	10026	1	cluster_1	1	1	1	1	0
14	10027	1	cluster_1	1	1	1	1	0
15	10031	0	cluster_1	1	1	0	0	0
16	10035	0	cluster_1	1	1	1	0	0
17	10036	0	cluster_1	1	1	1	0	0
18	10037	0	cluster_1	1	1	1	0	0
19	10038	0	cluster_1	1	1	1	0	0
20	10039	0	cluster_2	1	1	1	0	0
21	10040	0	cluster_1	1	1	1	0	0

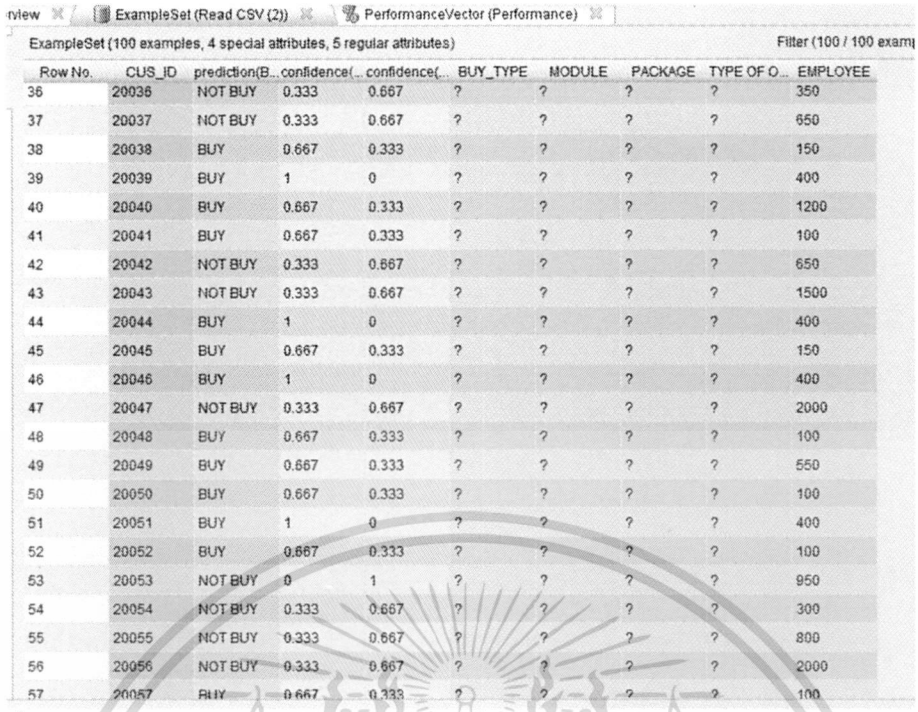
รูปที่ 7-6 แสดงผลการแบ่งกลุ่มของลูกค้าประเภทโรงงาน

เมื่อนำข้อมูลการซื้อขายซอฟต์แวร์มาวิเคราะห์โดยผลการทดลองด้านบนเป็นการจัดกลุ่มลูกค้าประเภทโรงงาน การทดลองนี้ได้จัดกลุ่มลูกค้าประเภทโรงงานออกเป็น 3 กลุ่ม คือ

1. Cluster0 คือกลุ่มของลูกค้าที่ซื้อซอฟต์แวร์เฉพาะโมดูล Pay
2. Cluster1 คือกลุ่มของลูกค้าที่ซื้อซอฟต์แวร์โมดูล Pay เป็นหลัก และมีโมดูลอื่นพร้อมกันด้วย
3. Cluster2 คือกลุ่มของลูกค้าที่ซื้อซอฟต์แวร์ 3 โมดูล คือ TMS Pay HCM

ทั้งนี้ การจัดกลุ่มลูกค้านั้นจะสามารถปรับเปลี่ยนจำนวนกลุ่มเองได้ตามความเหมาะสม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

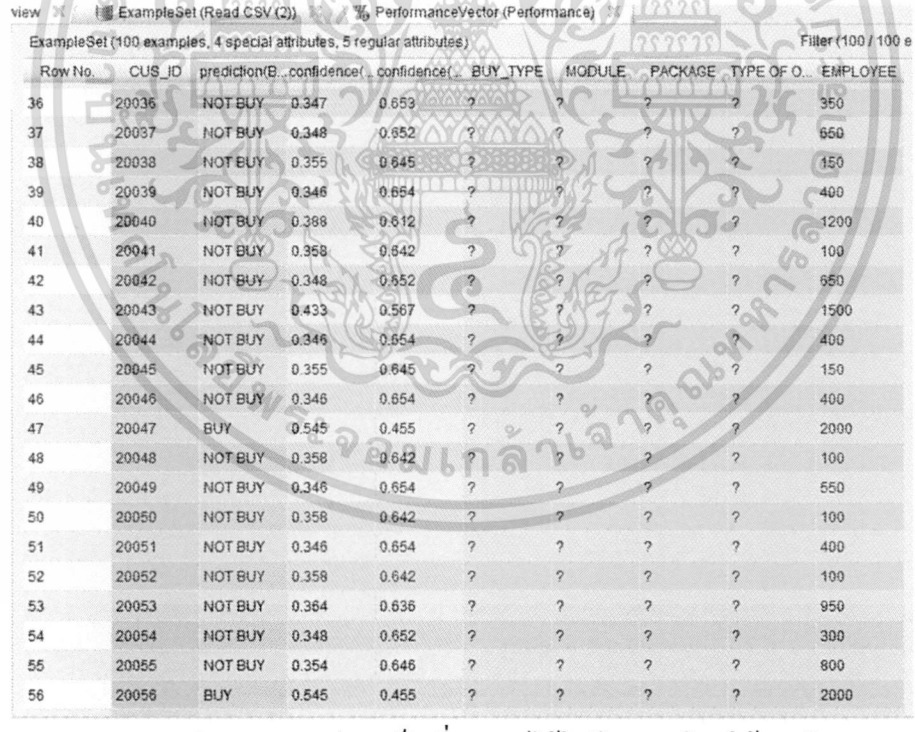


rview | ExampleSet (Read CSV (2)) | PerformanceVector (Performance)

ExampleSet (100 examples, 4 special attributes, 5 regular attributes) Filter (100 / 100 exam

Row No.	CUS_ID	prediction(B...	confidence(...	confidence(...	BUY_TYPE	MODULE	PACKAGE	TYPE OF O...	EMPLOYEE
36	20036	NOT BUY	0.333	0.667	?	?	?	?	350
37	20037	NOT BUY	0.333	0.667	?	?	?	?	650
38	20038	BUY	0.667	0.333	?	?	?	?	150
39	20039	BUY	1	0	?	?	?	?	400
40	20040	BUY	0.667	0.333	?	?	?	?	1200
41	20041	BUY	0.667	0.333	?	?	?	?	100
42	20042	NOT BUY	0.333	0.667	?	?	?	?	650
43	20043	NOT BUY	0.333	0.667	?	?	?	?	1500
44	20044	BUY	1	0	?	?	?	?	400
45	20045	BUY	0.667	0.333	?	?	?	?	150
46	20046	BUY	1	0	?	?	?	?	400
47	20047	NOT BUY	0.333	0.667	?	?	?	?	2000
48	20048	BUY	0.667	0.333	?	?	?	?	100
49	20049	BUY	0.667	0.333	?	?	?	?	550
50	20050	BUY	0.667	0.333	?	?	?	?	100
51	20051	BUY	1	0	?	?	?	?	400
52	20052	BUY	0.667	0.333	?	?	?	?	100
53	20053	NOT BUY	0	1	?	?	?	?	950
54	20054	NOT BUY	0.333	0.667	?	?	?	?	300
55	20055	NOT BUY	0.333	0.667	?	?	?	?	800
56	20056	NOT BUY	0.333	0.667	?	?	?	?	2000
57	20057	BUY	0.667	0.333	?	?	?	?	100

รูปที่ 7-7 แสดงผลการทำนายความน่าจะเป็นที่จะขายได้ในปี 2015 โดยใช้เทคนิค K-NN



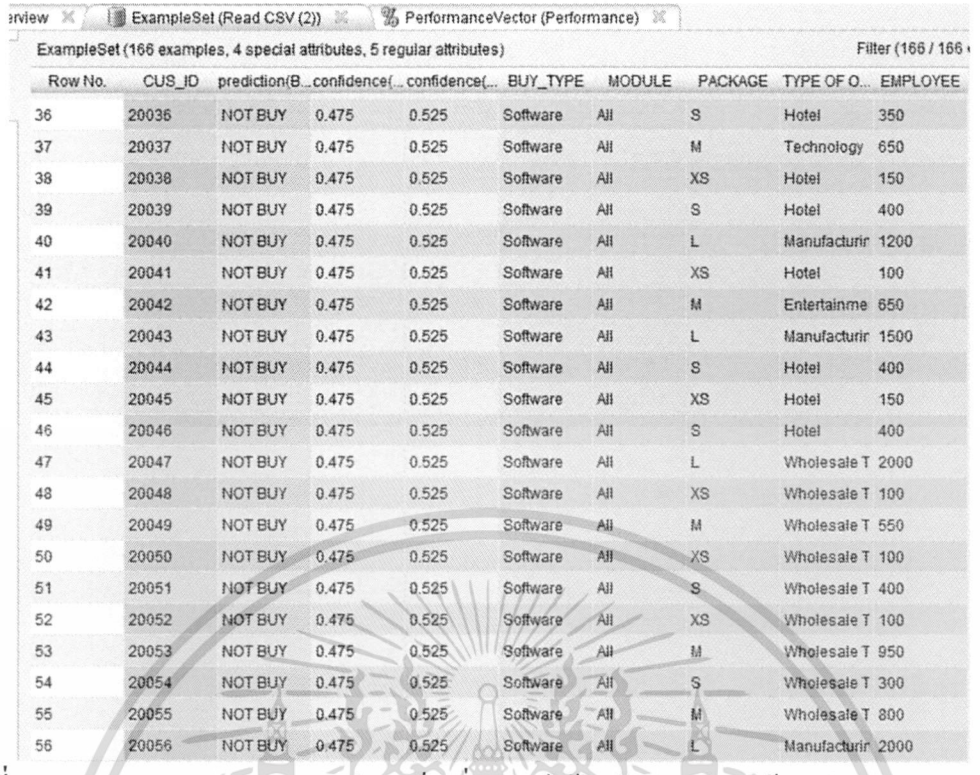
view | ExampleSet (Read CSV (2)) | PerformanceVector (Performance)

ExampleSet (100 examples, 4 special attributes, 5 regular attributes) Filter (100 / 100 e

Row No.	CUS_ID	prediction(B...	confidence(...	confidence(...	BUY_TYPE	MODULE	PACKAGE	TYPE OF O...	EMPLOYEE
36	20036	NOT BUY	0.347	0.653	?	?	?	?	350
37	20037	NOT BUY	0.348	0.652	?	?	?	?	650
38	20038	NOT BUY	0.355	0.645	?	?	?	?	150
39	20039	NOT BUY	0.346	0.654	?	?	?	?	400
40	20040	NOT BUY	0.388	0.612	?	?	?	?	1200
41	20041	NOT BUY	0.358	0.642	?	?	?	?	100
42	20042	NOT BUY	0.348	0.652	?	?	?	?	650
43	20043	NOT BUY	0.433	0.567	?	?	?	?	1500
44	20044	NOT BUY	0.346	0.654	?	?	?	?	400
45	20045	NOT BUY	0.355	0.645	?	?	?	?	150
46	20046	NOT BUY	0.346	0.654	?	?	?	?	400
47	20047	BUY	0.545	0.455	?	?	?	?	2000
48	20048	NOT BUY	0.358	0.642	?	?	?	?	100
49	20049	NOT BUY	0.346	0.654	?	?	?	?	550
50	20050	NOT BUY	0.358	0.642	?	?	?	?	100
51	20051	NOT BUY	0.346	0.654	?	?	?	?	400
52	20052	NOT BUY	0.358	0.642	?	?	?	?	100
53	20053	NOT BUY	0.364	0.636	?	?	?	?	950
54	20054	NOT BUY	0.348	0.652	?	?	?	?	300
55	20055	NOT BUY	0.354	0.646	?	?	?	?	800
56	20056	BUY	0.545	0.455	?	?	?	?	2000

รูปที่ 7-8 แสดงผลการทำนายความน่าจะเป็นที่จะขายได้ในปี 2015 โดยใช้เทคนิค Naïve Bayes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ExampleSet (Read CSV (2)) | PerformanceVector (Performance)

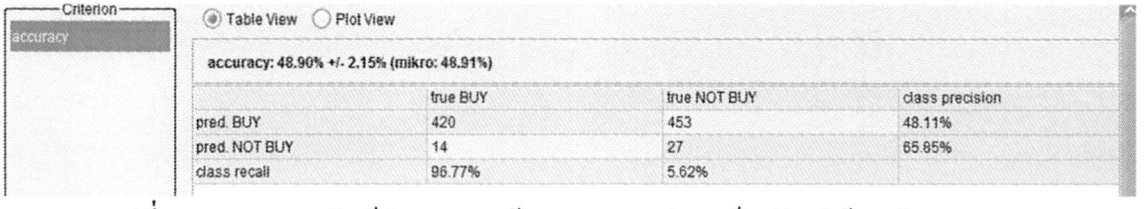
ExampleSet (166 examples, 4 special attributes, 5 regular attributes) Filter (166 / 166

Row No.	CUS_ID	prediction(B...	confidence(...	confidence(...	BUY_TYPE	MODULE	PACKAGE	TYPE OF O...	EMPLOYEE
36	20036	NOT BUY	0.475	0.525	Software	All	S	Hotel	350
37	20037	NOT BUY	0.475	0.525	Software	All	M	Technology	650
38	20038	NOT BUY	0.475	0.525	Software	All	XS	Hotel	150
39	20039	NOT BUY	0.475	0.525	Software	All	S	Hotel	400
40	20040	NOT BUY	0.475	0.525	Software	All	L	Manufacturir	1200
41	20041	NOT BUY	0.475	0.525	Software	All	XS	Hotel	100
42	20042	NOT BUY	0.475	0.525	Software	All	M	Entertainme	650
43	20043	NOT BUY	0.475	0.525	Software	All	L	Manufacturir	1500
44	20044	NOT BUY	0.475	0.525	Software	All	S	Hotel	400
45	20045	NOT BUY	0.475	0.525	Software	All	XS	Hotel	150
46	20046	NOT BUY	0.475	0.525	Software	All	S	Hotel	400
47	20047	NOT BUY	0.475	0.525	Software	All	L	Wholesale T	2000
48	20048	NOT BUY	0.475	0.525	Software	All	XS	Wholesale T	100
49	20049	NOT BUY	0.475	0.525	Software	All	M	Wholesale T	550
50	20050	NOT BUY	0.475	0.525	Software	All	XS	Wholesale T	100
51	20051	NOT BUY	0.475	0.525	Software	All	S	Wholesale T	400
52	20052	NOT BUY	0.475	0.525	Software	All	XS	Wholesale T	100
53	20053	NOT BUY	0.475	0.525	Software	All	M	Wholesale T	950
54	20054	NOT BUY	0.475	0.525	Software	All	S	Wholesale T	300
55	20055	NOT BUY	0.475	0.525	Software	All	M	Wholesale T	800
56	20056	NOT BUY	0.475	0.525	Software	All	L	Manufacturir	2000

รูปที่ 7-9 แสดงผลการทำนายความน่าจะเป็นที่จะขายได้ในปี 2015 โดยใช้เทคนิค Decision Tree

สำหรับส่วนการทำนายการขายซอฟต์แวร์ที่จะเกิดขึ้นในอนาคต จะใช้การขายจากอดีตมาเป็นข้อมูลประกอบการทำนายในปีนี้ ซึ่งด้านบนจะเป็นรายการของลูกค้าที่คาดว่าจะมีการติดต่อกันเกิดขึ้น โดยระบบจะทำการทำนายว่าลูกค้าเหล่านั้นมีแนวโน้มที่จะทำการซื้อขายกับทางบริษัทหรือไม่ ทั้งนี้เพื่อเป็นการใช้ข้อมูลมาเพื่อการประกอบการตัดสินใจ และทำการตลาดได้อย่างถูกต้องต่อไป

ถ้าสังเกตจากผลลัพธ์การวิเคราะห์ด้านบน อาจจะสังเกตเห็นได้ว่า ข้อมูลการทำนายจาก 3 เทคนิคนั้นมีบางส่วนที่ยังขัดแย้งกันอยู่ ดังนั้นเพื่อความถูกต้องที่สุด จึงมีการนำเทคนิคจึงมีการนำโอเปอร์เรเตอร์ Performance มาใช้เป็นตัวชี้วัดค่าความถูกต้องของทั้ง 3 อัลกอริทึม ดังรูป



Criterion: accuracy

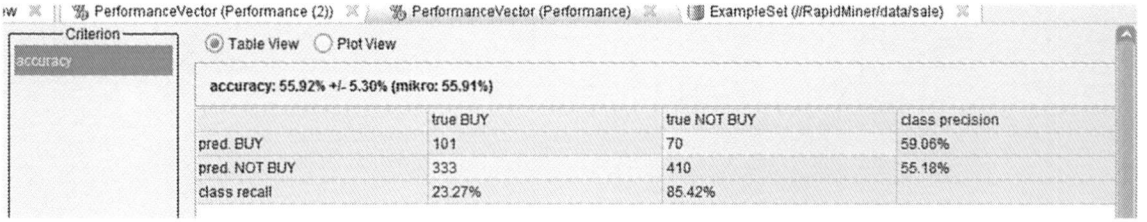
Table View | Plot View

accuracy: 48.90% +/- 2.15% (mikro: 48.91%)

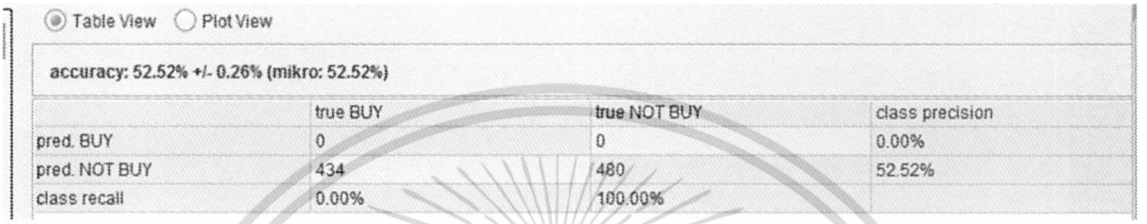
	true BUY	true NOT BUY	class precision
pred. BUY	420	453	48.11%
pred. NOT BUY	14	27	65.85%
class recall	96.77%	5.62%	

รูปที่ 7-10 แสดงผลลัพธ์ค่าความถูกต้องของความน่าจะเป็น โดยใช้เทคนิค K-NN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 7-11 แสดงผลลัพธ์ค่าความถูกต้องของความน่าจะเป็น โดยใช้เทคนิค Naïve Bayes



รูปที่ 7-12 แสดงผลลัพธ์ค่าความถูกต้องของความน่าจะเป็น โดยใช้เทคนิค Decision Tree

สำหรับข้อมูลการซื้อขายซอฟต์แวร์บริษัทออริซอฟท์นั้น จะเห็นได้ว่าค่าความถูกต้องของความน่าจะเป็นของอัลกอริทึม Naïve Bayes นั้นมีค่ามากที่สุด คือ 55.92% ดังนั้นเทคนิคที่เหมาะสมกับข้อมูลการขายมากที่สุดและจะใช้เทคนิคนี้ในการทำนาย นั่นก็คือ เทคนิค Naïve Bayes

7.2 ประโยชน์ที่ได้รับจากการออกแบบและพัฒนาระบบ

โครงงานการศึกษาอิสระนี้ ถูกพัฒนาขึ้นมาเพื่อใช้สำหรับเป็นที่เก็บข้อมูลการซื้อขาย และนำรายการเหล่านั้นมาวิเคราะห์เพื่อใช้เป็นแนวทางประกอบการตัดสินใจ ประโยชน์ที่ได้รับมีดังนี้

7.2.1 การจัดเก็บข้อมูลมีความครบถ้วน รวมอยู่ในที่เดียวกัน มีมาตรฐานเดียวกัน

7.2.2 ระบบวิเคราะห์ยอดขายที่สามารถนำไปใช้ประกอบการตัดสินใจ เพื่อเพิ่มยอดขาย หรือใช้เป็นระบบลูกค้าสัมพันธ์

7.3.3 สามารถนำผลลัพธ์รายงานในรูปแบบแผนภูมินำเสนอกับผู้บริหารระดับสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.3 ข้อเสนอแนะและแนวทางในการพัฒนาระบบเพิ่มเติม

แม้ว่าระบบสนับสนุนการตัดสินใจสำหรับการขายซอฟต์แวร์บริษัทออริซอฟท์ จะได้รับการออกแบบและพัฒนาระบบตามความต้องการของผู้ใช้แล้ว แต่ยังมีบางส่วนที่ควรต้องพัฒนาเพิ่มเติม ดังนี้

7.3.1 พัฒนาระบบให้มีความปลอดภัยมากยิ่งขึ้น ด้วยการใช้ User Login ก่อนการใช้หน้าใส่ข้อมูลการซื้อขาย

7.3.2 ด้วยเทคโนโลยีที่ก้าวกระโดดในสมัยนี้ อาจมีเครื่องมือวิเคราะห์ข้อมูลที่สามารถนำมาประยุกต์ใช้เพื่อเป็นระบบสนับสนุนการตัดสินใจได้เยอะขึ้น ผู้ที่ต้องการจะศึกษาด้านนี้ควรหาความรู้และเครื่องมือใหม่ๆอยู่เสมอ

7.3.3 ศึกษาและพัฒนาในส่วนของการเก็บข้อมูลการซื้อขาย ให้อยู่ในรูปแบบของโครงสร้างข้อมูล เพื่อที่จะสามารถนำมาประยุกต์ใช้กับระบบอื่นๆได้ง่ายขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

ธารินี แขวงเมือง. 15 ตุลาคม 2557. **ระบบสารสนเทศในองค์กร**. [Online]. เข้าถึงได้จาก: http://tharinee55.blogspot.com/2012/11/esstpsmrsdsskwsoas.html

ดร. เอกสิทธิ์ พัชรวงศ์ศักดา. 2557. **An Introduction to Data Mining Technique**

ดร. เอกสิทธิ์ พัชรวงศ์ศักดา. **โมเดล Naive Bayes และการแปลความหมาย.** [Online] เข้าถึงได้จาก: http://dataminingtrend.com/2014/naive-bayes/

Petia Wohed, Wil M.P. van der Aalst, Marlon Dumas, Arthur H.M. ter Hofstede and Nick Russell., 2014. **Pattern-based Analysis of UML Activity Diagrams**. [Online] Available: http://cms.ieis.tue.nl/Beta/Files/WorkingPapers/Beta_wp129.pdf

Satzinger, J. W. Jackson, R. B. and Burd, S. D. 2012. **Introduction to Systems Analysis and Design**: An Agile, Iterative Approach. 6th ed. Cengage Learning.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-นามสกุล	นางสาวณัฏฐพัชร กิตติชัยศรี
วัน เดือน ปีเกิด	23 พฤษภาคม 2531
สถานที่เกิด	กรุงเทพมหานคร
ที่อยู่	95/46 ซอยลาดพร้าว 20 ถนนลาดพร้าว แขวงจอมพล เขตจตุจักร กรุงเทพมหานคร 10900

การศึกษา

พ.ศ. 2550	จบการศึกษาระดับมัธยมศึกษา โรงเรียนพระปฐมวิทยาลัย จังหวัดนครปฐม
พ.ศ. 2554	จบการศึกษาระดับปริญญาตรี คณะวิศวกรรมศาสตร์ สาขาวิศวกรรมสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2556	เข้าศึกษาระดับปริญญาโท คณะเทคโนโลยีสารสนเทศ สาขาเทคโนโลยีระบบสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ประสบการณ์ทำงาน

พ.ศ. 2554 – พ.ศ. 2556	Programmer ที่ International Financial Data Services
พ.ศ. 2557 – ปัจจุบัน	Software Consultant ที่ Orisoft (Thailand) Co., Ltd.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้