

การใช้เทคนิค Data Mining กับโปรแกรมส่งเสริมการขาย

Data Mining Technique for Effective Sales Promotion



วัน เดือน ปี.....	0 1 ส.ค. 2550
เลขทะเบียน.....	02517
เลขเรียกหนังสือ.....	วทศ.ช 893ก 2541
"ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล."	

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการศึกษาระดับปริญญาตรี
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
ภาคเรียนที่ 2 ปีการศึกษา 2541
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง



ชื่อหัวข้อ	การใช้เทคนิค Data Mining กับโปรแกรมส่งเสริมการขาย
นักศึกษา	นายเอนก กองอมรภิญโญ
อาจารย์ที่ปรึกษา	ดร. วรพจน์ กริสุระเดช
ระดับการศึกษา	วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
แขนงวิชา	การจัดการเทคโนโลยีสารสนเทศ
ปีการศึกษา	2541

บทคัดย่อ

การจัดโปรแกรมส่งเสริมการขาย (Sales Promotion) โดยเฉพาะจดหมายเวียนเป็นเครื่องมือหนึ่งซึ่งบริษัท 3 เอ็มได้ใช้ช่วยกระตุ้นยอดขายในสถานการณ์ที่บริษัทต้องการเพิ่มยอดขาย เพื่อให้บรรลุจุดประสงค์ของโปรแกรมจะเห็นได้ว่าทางองค์กรจะต้องจัดสรรงบประมาณจำนวนมาก เพราะเนื่องมาจากต้นทุนในการจัดส่ง ต้นทุนในการทำใบปลิวหรือแค็ตตาล็อกเพื่อแนบไปกับจดหมายเวียนนั้นมีมูลค่าสูง ยิ่งจำนวนที่ส่งมากเท่าใด ก็ยิ่งทำให้ค่าใช้จ่ายสูงตามไปเท่านั้น จากเดิมในการทำโปรแกรมส่งเสริมการขาย การเลือกกลุ่มเป้าหมายที่องค์กรจะส่งไปให้นั้นองค์กรจะเลือกรายชื่อของลูกค้าจากฐานข้อมูลในองค์กร หรือเลือกมาจากสิ่งพิมพ์ต่าง ๆ โดยการเลือกนั้นมีพื้นฐานมาจากความเชื่อหรือสมมติฐานขององค์กรเอง โดยใช้ประสบการณ์เป็นตัวกำหนดว่าจดหมายเวียนดังกล่าวจะส่งไปถึงลูกค้าคนใดหรือกลุ่มใด แต่ปัญหาที่พบจากโปรแกรมคือว่า จำนวนการตอบรับต่อโปรแกรมนั้นต่ำมากและไม่คุ้มค่าต่อการลงทุน

จากปัญหาดังกล่าวข้างต้น ทำให้เทคโนโลยีสารสนเทศเริ่มที่จะเข้ามามีบทบาทในการแก้ปัญหา โดยองค์กรจะเลือกใช้เทคนิค Data Mining ที่มีลักษณะการทำงานแบบ Decision Tree โดยใช้ซอฟต์แวร์ “Knowledge Seeker” ของบริษัท Angoss เข้ามากำหนดลักษณะกลุ่มของลูกค้าที่แท้จริงว่าควรจะเป็นลูกค้ากลุ่มใดหรือมีลักษณะร่วมกันอย่างไร จากการศึกษาจะพบว่าคำตอบที่ได้จากการทำ Data Mining นั้น อาจจะแตกต่างจากสมมติฐานเดิมขององค์กรอย่างสิ้นเชิง ประโยชน์ที่ได้จากการทำงานของ Data Mining คือทำให้องค์กรพบกลุ่มที่แท้จริงของลูกค้าเป้าหมายที่จะมุ่งนำเสนอรายการส่งเสริมการขาย ทำให้โอกาสของความสำเร็จมีมากขึ้น

Title	Data Mining Technique for Effective Sales Promotion
Student	Mr. Anek Kongamornpinyo
Advisor	Worapoj Kreesuradej, Ph. D.
Level of Study	Master of Science in Information Technology
Major	Information Technology Management
Academic Year	1998

ABSTRACT

Direct Mailing is one of the most popular sales promotion techniques that 3M are using in order to increase sales substantially. To achieve the sales target, the company need to allocate a lot of budget to distribute the mail to the right customers. They also have to spend a huge budget for leaflet, catalog and mailing. For the past program, the mailing list is selected from company's database and industrial directory by their own experience. However it is found that the response rate is very low and the program does not worth investing.

From the above problem, the data mining technique by using Knowledge Seeker, one of the popular softwares from Angoss Corporation, is chosen to help define the right target customers to send the mail. This software is working under the classification and regression tree (CART) algorithm. They will find that the answer they have got from the algorithm is absolutely different from the preivous assumption which was made earlier. The benefit they would expect from the software is that they could target the program to the right customer and the software would help increase the success of the program at a reasonable budget.

กิตติกรรมประกาศ

ทางผู้จัดทำใคร่ขอขอบคุณอย่างสูงต่อท่านผู้ให้ความช่วยเหลือดังต่อไปนี้ ที่ช่วยให้โครงการศึกษาระณีพิเศษฉบับนี้ดำเนินเรื่องลุล่วงไปได้ด้วยดี

- อาจารย์วรพจน์ ที่กรุณาสละเวลาให้คำปรึกษา หนังสือและเอกสารประกอบการทำโครงการศึกษา
- บริษัท 3 เอ็ม ประเทศไทย จำกัด ที่เอื้อเพื่อให้ข้อมูลเกี่ยวกับตลาด และทีมฝ่ายขายที่ช่วยเหลือในเรื่องของการเก็บข้อมูล
- บริษัท Angoss ที่สนับสนุนทางด้าน Software “Knowledge Seeker”



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญภาพ.....	VI
สารบัญตาราง.....	VII
บทที่	
1. บทนำ.....	1
1.1 ความสำคัญและเหตุผลในการศึกษา.....	1
1.2 วัตถุประสงค์.....	1
1.3 ขอบเขตของการศึกษาและปัญหา.....	1
1.4 ขั้นตอนการศึกษา.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	2
2. ปัญหาทางธุรกิจและ Data Mining.....	3
2.1 ความเป็นมาและความสำคัญของปัญหา.....	3
2.2 ปัญหาที่พบได้บ่อยในการจัดโปรแกรมส่งเสริมการขาย.....	3
2.3 ทำไมจึงต้องมี Data Mining.....	4
3. Data Mining และทฤษฎีที่เกี่ยวข้อง.....	6
3.1 ความหมายของ Data Mining.....	6
3.2 Data Mining แตกต่างจาก Tool อื่น ๆ อย่างไร.....	7
3.3 ชนิดของ Data Mining.....	8
3.4 Data Warehousing, Data Mining และ Data Mart.....	11
3.5 Data Mining Tasks and Techniques.....	14
3.6 ขั้นตอนการทำงานของ Data Mining (Data Mining Process).....	20
3.7 การใช้งานทางด้าน Data Mining (Data Mining Application).....	26

สารบัญ (ต่อ)

หน้า

4. Decision Tree และขั้นตอนการดำเนินงานในระบบงาน	28
4.1 ขั้นตอนการศึกษาและตัวแปรที่เกี่ยวข้อง	28
4.2 สรุปผลการศึกษา.....	32
5. บทสรุป.....	33
6. บรรณานุกรม.....	35
7. ภาคผนวก.....	36
8. ประวัติผู้เขียน	42



สารบัญภาพ

รูปที่	หน้า
3.1 การสืบค้นข้อมูลและ Tool ชนิดต่าง ๆ	7
3.2 การแบ่งกลุ่มของลูกค้า (Clustering)	10
3.3 ตำแหน่งของ Data Warehousing	11
3.4 Data Mining และเครื่องมือทางธุรกิจต่าง ๆ	13
3.5 รูปแบบการตัดสินใจของบริษัทประกันภัย	16
3.6 การแบ่งกลุ่ม Target Group ในรูปแบบของ Clustering	17
3.7 รูปแบบการทำงานของ Neural Network	18
3.8 รูปแบบการทำงานของ Decision Tree	19
3.9 เวลาในแต่ละขั้นตอนของการทำ Data Mining	21
3.10 เครื่องมือทางกราฟฟิค	23
4.1 การเลือก Dependent Variable สำหรับการศึกษา	29
4.2 รูปการทำงานของ Decision Tree โดยให้ Algorithm สร้าง Tree ขึ้นมาเอง	30
4.3 รูปการทำงานของ Decision Tree ที่ศึกษาต่ออิงกลุ่มชนิดของอุตสาหกรรม	31
4.4 รูปการทำงานของ Decision Tree ที่ศึกษาถึงกลุ่มตลาดรองรับของอุตสาหกรรม	31

สารบัญตาราง

	หน้า
ตารางที่	
3.1 Data Element ของข้อมูล.....	9
4.1 ตารางการวิเคราะห์แนวโน้มการซื้อขายสินค้าในรายการส่งเสริมการขาย.....	28



บทที่ 1

บทนำ

1.1 ความสำคัญและเหตุผลในการศึกษา

การดำเนินธุรกิจท่ามกลางการแข่งขันอย่างรุนแรงในปัจจุบันไม่ว่าจะเป็นการแข่งขันทั้งภายในและภายนอกประเทศ ทำให้องค์กรต่าง ๆ ต่างปรับตัวเพื่อให้มีความได้เปรียบในเชิงการแข่งขันมากที่สุด เพื่อที่องค์กรจะสามารถครองส่วนแบ่งตลาดและทำกำไรให้ได้สูงสุด ซึ่งในขณะเดียวกันก็ยังทำให้ผู้บริโภคมีความพึงพอใจสูงสุดด้วย เทคโนโลยีสารสนเทศได้เข้ามามีบทบาทในองค์กรต่าง ๆ มากมาย ตัวอย่างเช่น การจัดการคลังข้อมูล (Data Warehousing) ซึ่งเป็นตัวผลักดันตัวหนึ่งที่ทำให้เกิดเทคนิคที่เรียกว่า “Data Mining”

การที่ในตลาดมีการแข่งขันอย่างรุนแรงและรูปแบบพฤติกรรมผู้บริโภคใช้จ่ายสินค้าของถูกค้ำที่เปลี่ยนไปมีความซับซ้อนมากกว่าในอดีต ปัจจัยดังที่ได้กล่าวมานี้เป็นตัวผลักดันที่ทำให้กิจกรรมทางการตลาดเปลี่ยนโฉมไป เป็นหน้าที่ที่องค์กรจะต้องปรับกลยุทธ์เพื่อสร้างความพึงพอใจให้กับลูกค้าในแต่ละกลุ่ม เราจะได้ยิน Concept ใหม่ ๆ เกิดขึ้น ไม่ว่าจะเป็น One-To-One Marketing, Database Marketing เป็นต้น สถานะการณ์เหล่านี้ต้องการการวิเคราะห์ที่แตกต่างไปจากเดิม จากสาเหตุดังกล่าวจึงได้มีการนำเอาเทคนิค Data Mining เข้ามาช่วยในการออกแบบและพัฒนาโปรแกรมส่งเสริมการขายโดยเฉพาะ โปรแกรมที่เป็น Direct Marketing ซึ่งจะกล่าวถึงรายละเอียดในบทต่อ ๆ ไป

1.2 วัตถุประสงค์

เพื่อนำเอาเทคนิค Data Mining มาใช้ในการออกแบบโปรแกรมส่งเสริมการขาย วัตถุประสงค์เพื่อให้โปรแกรมตรงกับลูกค้ากลุ่มเป้าหมายที่สุดเพื่อเพิ่มยอดขายสินค้าขององค์กร โดยที่ใช้อองค์กรจะต้องลงทุนออกค่าใช้จ่ายในการทำโปรแกรมที่เหมาะสม และลดค่าใช้จ่ายในการทำโปรแกรมส่งเสริมการขายอีกด้วย

1.3 ขอบเขตของการศึกษาและปัญหา

ชนิดของการส่งเสริมการขายผ่านการ สินค้าโปรแกรมส่งเสริมการขาย คือ เมื่อลูกค้าซื้อสินค้ากระดุมทราย 20 ม้วนทางองค์กรจะแถมเครื่องจักรกระดุมทรายให้ฟรีจำนวน 1 เครื่อง

โครงการนี้จะทำการศึกษาดึงการนำเอาซอฟต์แวร์สำเร็จรูป (Data Mining Technique) มาปรับปรุงเพื่อใช้ในงานส่วนหนึ่งของระบบ การออกแบบโปรแกรมส่งเสริมการขายเพื่อกลุ่มลูกค้าที่เป็นโรงงานไม้ โรงงานโลหะ และโรงงานอุตสาหกรรมรถยนต์ ผ่านการใช้ฐานข้อมูลจากภายในบริษัท เพื่อหากลุ่มเป้าหมายที่จะตอบสนองต่อโปรแกรมมากที่สุด โดยมองไปที่พฤติกรรม การซื้อ, ลักษณะของธุรกิจ, ประวัติส่วนตัวของลูกค้า และปัจจัยอื่น ๆ

ปัญหาของโปรแกรมโปร โมชั่นที่พบหลังจาก ได้มีการเสนอโปรแกรมให้ลูกค้า

- จำนวนของจดหมายเวียนที่เหมาะสม
- จำนวนของลูกค้าและลักษณะของลูกค้าที่ควรส่ง ไปให้เพื่อให้ได้ผลมากที่สุด

1.4 ขั้นตอนการศึกษา

เพื่อให้การศึกษามรรดตามวัตถุประสงค์ที่กำหนดไว้ ภายใต้ขอบเขตของการศึกษาจึงได้กำหนดขั้นตอนในการศึกษาไว้ดังนี้

ขั้นที่1 เก็บรวบรวมข้อมูล

ขั้นที่2 ศึกษาทฤษฎีที่เกี่ยวข้อง

ขั้นที่3 ศึกษาวิธีการทำงานและแนวทางในการนำเอาเทคนิคใหม่มาใช้

ขั้นที่4 สรุปผลการศึกษา

1.5 ประโยชน์ที่คาดว่าจะได้รับจากโครงการ

- 1) เพื่อเพิ่มยอดขายขององค์กร
- 2) เพื่อลดค่าใช้จ่ายในการจัดทำโปรแกรมส่งเสริมการขาย
- 3) เพิ่มความเร็วในการเข้าสู่ตลาดของสินค้าใหม่
- 4) สร้างความพึงพอใจให้กับลูกค้า

ในบทต่อ ๆ ไปจะกล่าวถึงปัญหาที่เกิดขึ้นจริงในเชิงธุรกิจ ทฤษฎีที่เกี่ยวข้องและเทคนิคในการแก้ไขปัญหาต่าง ๆ จะพบว่าไม่ใช่ว่าเพียงเทคนิค Data Mining ด้วยตัวของมันเองที่จะแก้ปัญหาได้ ยังต้องอาศัยเทคนิคอื่น ๆ ที่เกี่ยวข้องมาประกอบกัน เช่น เทคนิคของการทำ Data Warehousing เป็นต้น

บทที่ 2

ปัญหาทางธุรกิจและ Data Mining

2.1 ความเป็นมาและความสำคัญของปัญหา

การจัดโปรแกรมส่งเสริมการขาย (Sales Promotion) เป็นกิจกรรมทางการตลาดอย่างหนึ่งซึ่งมีความสำคัญเป็นอย่างมาก โปรแกรมส่งเสริมการขายจะช่วยเพิ่มและกระตุ้นยอดขายให้กับองค์กรหรือบริษัทต่าง ๆ ได้ในสถานการณ์ที่องค์กรต้องการยอดขายเพื่อมาหล่อเลี้ยงธุรกิจขององค์กรเอง ตัวอย่างของการจัดโปรแกรมส่งเสริมการขายได้แก่ การจัดโปรโมชั่น การลดราคาสินค้า การแอมสสินค้า การทำจดหมายเวียนถึงลูกค้า เป็นต้น

แต่เราจะมั่นใจได้อย่างไรว่าโปรแกรมที่เราจัดแต่ละโปรแกรมนั้นจะเป็นโปรแกรมที่ถูกต้อง ตรงกลุ่มเป้าหมาย และใช้งบประมาณอย่างคุ้มค่าที่สุด จะเห็นได้ว่าการจัดโปรแกรมแต่ละครั้งจะมีการจัดสรรงบประมาณจำนวนมาก การวางแผนต้องรัดกุมรอบคอบ ประกอบกับสภาวะปัจจุบันซึ่งสภาพเศรษฐกิจซบเซา เพื่อให้บรรลุจุดประสงค์ในการทำโปรแกรมทำให้องค์กรต่าง ๆ พยายามจะใช้อำนาจงบประมาณในการทำโปรแกรมแต่ละโปรแกรมให้คุ้มค่ากับเงินที่ดองลงทุนไปมากที่สุด โดยพิจารณาจากสิ่งแวดล้อมทั้งภายในและภายนอกองค์กรเองโดยใช้ทรัพยากรที่มีอยู่ภายในองค์กรเพื่อพิจารณาจัดโปรแกรม

2.2 ปัญหาที่พบได้บ่อยในการจัดโปรแกรมส่งเสริมการขาย

- การจัดโปรแกรมการขายแต่ละครั้งใช้ค่าใช้จ่ายจำนวนมาก
- องค์กร ไม่ทราบว่าจำนวนงบประมาณที่แน่นอนที่เหมาะสมในแต่ละโปรแกรม
- โปรแกรมไม่สามารถบรรลุจุดประสงค์ที่ตั้งไว้
- งบประมาณมีจำกัด
- ภาวะกดดันทางการตลาด เช่น คู่แข่ง
- โปรแกรมไม่ตรงตามความต้องการของลูกค้า

2.3 ทำไมจึงต้องมี Data Mining

องค์กรเริ่มค้นพบว่าสินทรัพย์ที่มีค่าอย่างมากขององค์กรก็คือ Information ที่มีในรูปแบบของฐานข้อมูลลูกค้า และ พฤติกรรมการซื้อสินค้า ความสามารถในการแข่งขันขององค์กรขึ้นอยู่กับความสามารถในการตัดสินใจขององค์กรเอง โดยสามารถพิจารณาจากการตัดสินใจในอดีตและพิจารณารายการ Transaction ที่เกิดขึ้น ความสามารถที่จะรู้จักลูกค้าและตลาดจะทำให้องค์กรสามารถวางเป้าหมายสินค้าไปสู่กลุ่มเป้าหมายที่ถูกต้องได้ ทำให้องค์กรสามารถวางเป้าหมายของการทำโปรแกรมส่งเสริมการขายและการจัดการที่ถูกต้อง

เนื่องจากลักษณะของธุรกิจเปลี่ยนไปและปรับตัวไปตามการแข่งขันและพฤติกรรมของผู้บริโภคที่เปลี่ยนไป ทำให้องค์กรต้องปรับตัวตามสภาพแวดล้อมที่เกิดขึ้น

- Mass Marketing and Database Marketing กลุ่มของผู้บริโภคมีมากขึ้นและยังมีความหลากหลายมากยิ่งขึ้น ดังนั้นกิจกรรมทางการตลาดเดิม ๆ ที่ใช้อยู่จึงจะไม่ได้ดีเท่าที่ควร จึงทำให้เกิดกิจกรรมทางการตลาดใหม่ ๆ เกิดขึ้น โดยกิจกรรมดังกล่าวตั้งอยู่บนพื้นฐานของการใช้ฐานข้อมูลขนาดใหญ่ของลูกค้า เพื่อเพิ่มความได้เปรียบต่อคู่แข่ง องค์กรจะต้องพัฒนากิจกรรมที่เข้าถึงได้โดยตรงไปที่ตัวของผู้บริโภคซึ่งมีฐานของลูกค้ากว้างมาก ซึ่งเราเรียกกิจกรรมดังกล่าวว่า “Direct Marketing” ตัวอย่างเช่น การใช้จดหมายเวียน
- พฤติกรรมของผู้บริโภคเปลี่ยนไปซึ่งพฤติกรรมดังกล่าวอาจปรับเปลี่ยนไปตามสภาพของเศรษฐกิจหรือสภาวะแวดล้อม เช่น พฤติกรรมของผู้บริโภคปัจจุบันมักจะนิยมซื้อสินค้าอุปโภคบริโภคในห้างสรรพสินค้าเพราะมีที่จอดรถสะดวก
- ตลาดอิมิตัว เนื่องจากปริมาณการเสนอขายสินค้ามีมากเกินไปความต้องการการบริโภคของผู้บริโภค ทำให้คู่แข่งแต่ละรายจะต้องปรับตัวเพื่อตอบสนองความต้องการของลูกค้าให้มากที่สุด
- ตลาดกลุ่มย่อยใหม่ ๆ (Niche) เกิดขึ้นตลอดเวลา เนื่องจากความซับซ้อนของพฤติกรรมของผู้บริโภค ความชอบที่ต่างกันของกลุ่มคน ทำให้เกิดพฤติกรรมการบริโภคแบบใหม่ซึ่งหลบซ่อนอยู่
- วัฏจักรอายุสินค้า เนื่องจากพฤติกรรมของผู้บริโภคเปลี่ยนเร็วมาก ทำให้อายุของสินค้าสั้นลงกว่าในอดีต ความซับซ้อนในการบริโภคสินค้ามีมากขึ้น
- Time to Market คู่แข่งขันในตลาดมีมากขึ้น องค์กรจะต้องหาวิธีที่จะทำให้สินค้าไปสู่ผู้บริโภคเร็วที่สุด
- คู่แข่งและความเสี่ยงทางธุรกิจมีมากขึ้น

หลังจากที่ได้รับทราบปัญหาในเชิงธุรกิจไปเรียบร้อยแล้ว ต่อไปเราจะได้ศึกษาถึงเรื่องในเชิงเทคนิคของ Data Mining เพื่อจะได้นำเทคนิคต่าง ๆ ที่จะได้อีกต่อไป มาเพื่อประกอบการตัดสินใจในแต่ละปัญหา และจะพบว่าปัญหาที่เกิดขึ้นเดียวกันอาจใช้เทคนิคของ Data Mining เดียวกันก็ได้ หรือบางปัญหาอาจจะใช้ได้เพียงเทคนิคเดียว นอกเหนือจากทางด้านเทคนิคแล้ว ยังมีการกล่าวถึงขั้นตอนในการทำ Data Mining ด้วย เพื่อเป็นหลักในการแก้ปัญหาอย่างถูกต้องต่อไป



บทที่ 3

Data Mining และทฤษฎีที่เกี่ยวข้อง

3.1 ความหมายของ Data Mining

จากคำจำกัดความ Data Mining อาจหมายถึงการที่ผู้ใช้ดึงและสังเคราะห์และตรวจสอบข้อมูลอย่างละเอียด โดยการสังเคราะห์ดังกล่าวอาจจะเป็นการเรียนรู้ข้อมูลในอดีตหรือข้อมูลในปัจจุบัน ผลลัพธ์ที่ได้ออกมาต้องมีลักษณะของข้อมูลที่มีลักษณะที่เป็นข้อมูลแบบ Unknown, ข้อมูลแบบ Valid และข้อมูลแบบ Actionable มาจากฐานข้อมูลขนาดใหญ่ซึ่งอาจมาจากรายการ Transaction, ฐานข้อมูลของฝ่ายขาย, E-Mail เพื่อนำข้อมูลดังกล่าวไปใช้เป็นพื้นฐานในการประกอบการตัดสินใจในเชิงธุรกิจ ทำให้เข้าใจแนวโน้มและรูปแบบของตลาด

ข้อมูลแบบ Unknown ข้อมูลที่ถูกใช้จะต้องเป็นข้อมูลผู้ใช้งานไม่รู้มาก่อนและไม่ชัดเจน ไม่สามารถตั้งสมมติฐานล่วงหน้าว่าควรจะเป็นแบบใด ตัวอย่างเช่น เจ้าของห้างสรรพสินค้าแห่งหนึ่งเพิ่งจะค้นพบว่าพฤติกรรมของผู้บริโภคใหม่สำหรับผู้บริโภคที่เป็นพ่อบ้านมักจะซื้อสินค้าเบียร์และผ้าอ้อมในวันศุกร์ตอนเย็น ดังนั้นเป็นสัญญาณให้เจ้าของกิจการควรจะเตรียมสินค้าไว้เพื่อจำหน่าย ซึ่งในขณะที่เขวกันห้างสรรพสินค้าคู่แข่งอาจจะไม่รู้เรื่องนี้ก็ได้ แต่ลองสังเกตคู่อีกหนึ่งตัวอย่างว่า เจ้าของร้านขายรถยนต์พบว่ารถยนต์ขนาดใหญ่ราคาแพงมักจะถูกซื้อโดยคนที่สูงอายุ ซึ่งเจ้าของไม่รู้มาก่อน แต่ข้อมูลดังกล่าวไม่เป็นลักษณะ Unknown เพราะสมมติฐานดังกล่าวมีอยู่ เพราะคนที่มีอายุมักจะมีฐานะที่ดีขึ้นเมื่อเทียบกับคนในวัยที่อายุน้อยกว่า

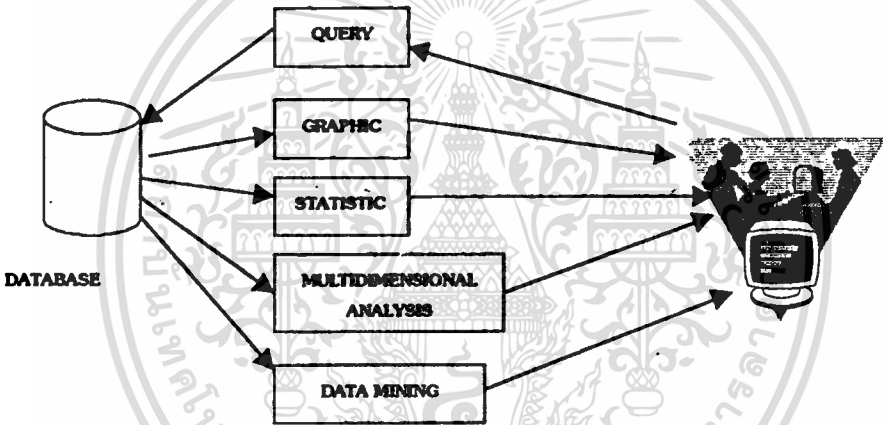
ข้อมูลแบบ Valid เมื่อผู้ใช้ได้เริ่มใช้เทคนิค Data Mining จะค้นพบสิ่งที่น่าสนใจตลอดเวลา แต่ที่ต้องพิจารณาด้วยว่าสิ่งนั้น Valid หรือไม่ เช่น ผู้ใช้มักจะพบว่ามีความสัมพันธ์ของการซื้อของ 2 สิ่งเสมอเมื่อจำนวนความหลากหลายสินค้ามากขึ้น แต่ไม่ได้หมายความว่าจะต้องให้ห้างสรรพสินค้าเก็บสินค้ามากขึ้น เพราะข้อมูลที่ได้อาจเกิดความคลาดเคลื่อน เพราะฉะนั้นจะต้องทำการ Validation และ Checking ความถูกต้องของข้อมูลและวิเคราะห์ความถูกต้องอีกครั้ง

ข้อมูลแบบ Actionable ข้อมูลจะต้องถูกแปลงออกมาและนำมาตัดสินใจให้เป็นความได้เปรียบเชิงธุรกิจ บางครั้งข้อมูลที่เราค้นพบเป็นสิ่งที่คู่แข่งได้ทำไปแล้วหรือคิดกฎหมาย ซึ่งจะต้องมีวิจรณ์ญาณในการใช้ด้วย ซึ่งข้อมูลดังกล่าวก็จะไม่มีประโยชน์อะไร

3.2 Data Mining แตกต่างจาก Tool อื่น ๆ อย่างไร

จากคำจำกัดความจะเห็นได้ว่า Data Mining ต่างจากวิธีเดิม ๆ ที่ใช้ในงานวิเคราะห์ข้อมูล เช่น พวกร Query Language, SQL Query และ On Line Analytical Processing (OLAP), Multidimensional Dimensional Analysis เพราะ Data Mining เป็นการค้นหาข้อมูลโดยปราศจากการตั้งสมมติฐานไว้ล่วงหน้า

ในงานของ Database Marketing นักวิเคราะห์ข้อมูล, และนักสถิติ มักจะมีคำถามว่าในเมื่อ Data Mining ถูกใช้ร่วมกับเทคนิคการวิเคราะห์ข้อมูลแบบเดิม ๆ มากมาย แต่มันต่างกันตรงไหน Data Mining มักถูกเอาไปปะปนกับ SQL, Multidimensional Dimensional Analysis (MDA), OLAP, แต่เทคนิคเหล่านี้ไม่ใช่ Data Mining เพราะ



รูปที่ 3.1 การสืบค้นข้อมูลและ Tool ชนิดต่าง ๆ (Cabena et al., 1997)

จากรูป 3.1 สมมติให้ผู้ใช้ฐานข้อมูลเป็นผู้จัดการภาคของร้านขายเครื่องไฟฟ้าองค์กรหนึ่ง ซึ่งมีสาขามากมายกระจายอยู่ทั่วประเทศ

- SQL Query เป็นคำถามพื้น ๆ เช่น อาจจะกำหนดจำนวนของวิดีโอที่ขายได้ในร้านทางภาคเหนือเดือนที่แล้วว่าเป็นจำนวนเท่าใด
- Visualizational Graphic อาจจะ Plot กราฟยอดขายของแต่ละร้านที่แตกต่างกันระยะเวลาหนึ่ง
- Statistics วิเคราะห์ความอ่อนไหวทางด้านราคาต่อราคาโทรทัศน์ซึ่งขายตามชานเมือง

- MDA (Multi Dimensional Analysis) เป็นเครื่องมือที่ช่วยทำให้ Query มีประสิทธิภาพมากขึ้นใช้เพื่อเปรียบเทียบยอดขายของสินค้าชนิดต่าง ๆ ในร้านต่าง ๆ ระยะเวลาต่าง กัน

จากข้างต้นจะพบว่าเทคนิคทุกอันมีสิ่งที่ยังร่วมกันอยู่ก็คือ มีสมมติฐานอยู่แล้ว ผู้จัดการรู้ว่า มีจำนวนสินค้า, จำนวนร้าน, และยอดขายอยู่ที่เท่าไร เขาสามารถเช็คดูความสัมพันธ์ได้ การวิเคราะห์แบบนี้เป็น Traditional Data Analysis เครื่องมือประเภท Query Tool นี้จะยอมให้ผู้ใช้อินเตอร์เฟซถามกับ DBMS จะให้ผู้ใช้หาความจริงใหม่ ๆ หรือความจริงที่น่าสนใจ จากข้อมูลซึ่งถูกเก็บไว้ในฐานข้อมูล คำถามที่จะเกิดขึ้นได้แก่ ยอดขายของสินค้าเป็นเท่าไรในภาคเหนือเมื่อเทียบกับภาคใต้

สำหรับ Data Mining เองนั้นแตกต่างไปโดยจะใช้คำถามเหล่านี้ซึ่งเทคนิคเดิม ๆ ตอบไม่ได้ แทนที่จะทำการเชื่อมต่อระหว่างภาคและยอดขาย Data Mining จะกำหนดปัจจัยสำคัญ ๆ ซึ่งเป็นสาเหตุที่ทำให้ยอดขายสูงหรือยอดขายต่ำ ในการศึกษากรณีแบบนี้เราไม่สามารถรู้ได้ว่าอะไรเป็นปัจจัยที่ทำให้ยอดขายสูงหรือต่ำ Data Mining จะหาปัจจัยซึ่งมีอิทธิพลที่ทำให้กระทบยอดขายโดยใช้ขั้นตอนในการแก้ปัญหา (Data Mining Algorithm) หารูปแบบและแนวโน้มของข้อมูลนั้น เครื่องมือ Data Mining ไม่ต้องการสมมติฐาน มันพยายามที่จะค้นหาความสัมพันธ์และรูปแบบที่ถูกซ่อนปิดบังไว้ทำให้ไม่เห็นภาพชัดเจน ตัวอย่างเช่น

- ทำไมคูปองส่วนลดถึงไม่ดึงดูดลูกค้า
- ทำอย่างไรจึงจะเพิ่มส่วนแบ่งตลาดต่อลูกค้า (Customer Share) ได้จากการใช้จ่ายเรื่องเครื่องไฟฟ้า
- ทำอย่างไรจึงจะให้ร้านสาขาอื่นทำยอดขายได้ดีเท่าสาขาภาคเหนือ

3.3 ชนิดของ Data Mining

มีหลายวิธีด้วยกันในการแยกแยะชนิดของ Data Mining โดยสามารถแยกแยะตามกรรมวิธีการทำ Data Mining ได้ 3 ชนิดด้วยกันคือ

3.3.1 Classification Studies (Supervised Learning)

ตัวอย่างนี้จะสร้างความเข้าใจใน Classification Study ซึ่งกรณีของตัวอย่างนี้พบได้ทั่วไปในวงการธุรกิจ นักวิเคราะห์ในองค์กรที่ทำธุรกิจเกี่ยวกับการสื่อสารแห่งหนึ่งต้องการเข้าใจว่าทำไมลูกค้าบางกลุ่มถึงยังคงซื่อสัตย์และมี Brand Loyalty สูงกับสินค้าขององค์กร แต่ในขณะที่เดียวกันลูกค้าอีกกลุ่มกลับไปหาคู่แข่งแทน ทำที่ตุนักวิเคราะห์จึงต้องการจะทำนายลักษณะและนิสัยของลูกค้าที่องค์กรจะต้องเสียไปให้คู่แข่ง

เนื่องจากขณะนี้นักวิเคราะห์มีเป้าหมายในใจเรียบร้อยแล้ว ดังนั้นนักวิเคราะห์จึงสามารถสร้าง Model ที่ข้อมูลต่าง ๆ ได้มาจากข้อมูลในอดีตของลูกค้าที่มีความซื่อสัตย์ต่อองค์กรและกลุ่มลูกค้าที่ไม่มีความซื่อสัตย์ต่อองค์กร ด้วย Model ที่สมบูรณ์ถูกต้องจะสามารถทำให้องค์กรเข้าใจและทำนายลักษณะของธุรกิจที่จะเกิดขึ้นได้

จากตัวอย่างเหตุการณ์จะสามารถอธิบายขั้นตอนของการกำหนดการศึกษาได้ การศึกษาจะกำหนดขอบเขตของกิจกรรมของ Data Mining ได้ นอกจากนี้การศึกษาจะสามารถกำหนดจุดประสงค์และข้อมูลที่ต้องการใช้ได้ทั้งหมด ด้วยการกำหนดปัญหาทางธุรกิจ นั่นก็เป็นสิ่งที่บอกให้นักวิเคราะห์ทราบได้เลยว่าขั้นตอนในการทำ Data Mining จะทำอย่างไรและจุดประสงค์ของการทำคืออะไร

ในการศึกษาต้องการหัวข้อในการศึกษา หัวข้อในการศึกษาอาจหมายถึง Data Element ของ Object ที่เราต้องการจะศึกษา เช่น เราต้องการจะศึกษาถึง Object “ลูกค้า” ซึ่งมี Data Element ที่เกี่ยวข้องคือ ชนิดของลูกค้า แนวโน้มการซื้อสินค้า ระยะเวลาที่เป็นลูกค้าขององค์กร และอื่น ๆ ซึ่ง Data Element จะเป็นตัวกำหนดลักษณะและชนิดของลูกค้าการทำ Classification Studies นั้นเราสามารถกำหนดโครงสร้างลักษณะเฉพาะหรืออุปนิสัยของลูกค้าได้โดยดูได้จากตาราง 3.1

ชื่อคอลัมน์	ชนิดของข้อมูล	ค่าที่ได้	คำอธิบาย
เบอร์ลูกค้า	ตัวเลข	ค่าเฉพาะ	ตัวกำหนดเฉพาะสำหรับลูกค้า
ระยะเวลา	ตัวเลข	จำนวนเต็ม	จำนวนที่ลูกค้าอยู่กับองค์กร
แนวโน้ม	ตัวอักษร	เพิ่มขึ้น, เหมือนเดิม, ลดลง	ตัวบ่งชี้แนวโน้มการใช้สินค้า 6 เดือนล่าสุด
สถานะ	ตัวอักษร	สูง, กลาง, ต่ำ, ไม่ทราบ	การสำรวจผลความพอใจของลูกค้า
ชนิดของลูกค้า	ตัวอักษร	ยังคงซื่อสัตย์, ไม่ซื่อสัตย์	ลูกค้ายังคงอยู่กับองค์กรหรือเสียให้คู่แข่งไปแล้ว

ตารางที่ 3.1 Data Element ของข้อมูล

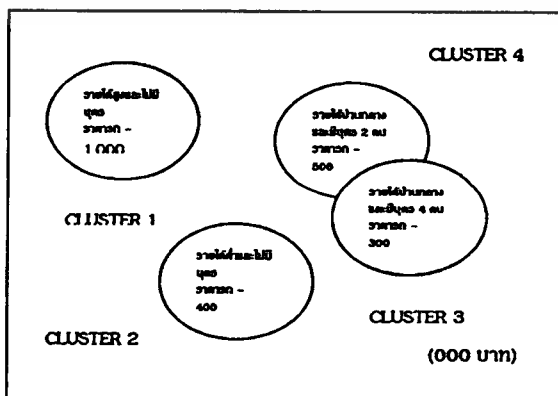
จากตัวอย่างข้างต้นเรากำหนดให้ ชนิดของลูกค้าเป็น Output หรือ Dependent Variable ซึ่งถูกใช้เป็นพื้นฐานในการศึกษาว่าจะอะไรคือสาเหตุที่ทำให้ลูกค้าซื้อผลิตภัณฑ์ขององค์กรและทำลูกค้าถึงจากองค์กรไป และเราจะใช้ Data Element ตัวอื่น ๆ มาช่วยในการอธิบายสิ่งที่เกิดขึ้น เรากำหนดให้ ชนิดของลูกค้าเป็น Training Data ถ้าเราเปลี่ยน Data Element ตัวอื่นมาเป็น Output จุดประสงค์ของการศึกษาก็จะเปลี่ยนไปด้วย

3.3.2 Clustering Studies (Unsupervised Learning)

Clustering คือวิธีการของการรวมกลุ่มของข้อมูลที่มีลักษณะเหมือนกัน รูปแบบและแนวโน้มที่เหมือนกัน การศึกษาของ Clustering ไม่มี Output หรือ Independent Variable เหมือน Classification Studies และไม่มีการจัดเป็นลักษณะโครงสร้างของ Object ใด ๆ ดังนั้นการศึกษาแบบนี้จึงถูกเรียกว่า Unsupervised Learning หรือ Segmentation การทำ Clustering เองสามารถทำบนพื้นฐานของข้อมูลในอดีตได้เหมือนกัน แต่ผลลัพธ์ที่ได้มาไม่ได้ออกจากจาก Training Data

ตัวอย่างของ Clustering เช่น องค์กรต้องการทราบความเหมือนที่มีในกลุ่มลูกค้าของตนเอง เพื่อที่ว่าองค์กรจะสามารถเข้าใจลักษณะเฉพาะของกลุ่มลูกค้านเป้าหมายขององค์กร และสร้างกลุ่มของลูกค้านที่องค์กรจะสามารถขายสินค้าได้ในอนาคต องค์กรจะทำการแยกกลุ่มของลูกค้าออกเป็นกลุ่ม ๆ

เทคนิคของ Clustering พยายามมองหาความเหมือนและความแตกต่างภายในกลุ่มของข้อมูล และแบ่งกลุ่มต่าง ๆ ออกเป็นส่วน ๆ ตัวอย่างเช่น กลุ่มของข้อมูลจะประกอบไปด้วยลูกค้าหลายรายที่มีราย ได้สูงและยัง ไม่มีบุตร และจำนวนลูกค้าอีกหลายรายที่มีราย ได้ต่ำและแต่งงานแล้ว ระหว่างการศึกษาถ้ายังมีกลุ่มของข้อมูลที่แตกต่างออกไปจาก 2 กลุ่มนี้ เราก็แบ่งกลุ่มออกให้เป็นกลุ่มย่อย ๆ ได้อีก จากตัวอย่างนี้รูปที่ 3.2 เราจะแบ่งกลุ่มของลูกค้าออกเป็น 4 กลุ่ม



รูปที่ 3.2 การแบ่งกลุ่มของลูกค้า (Groth, 1997)

3.3.3 Visualization

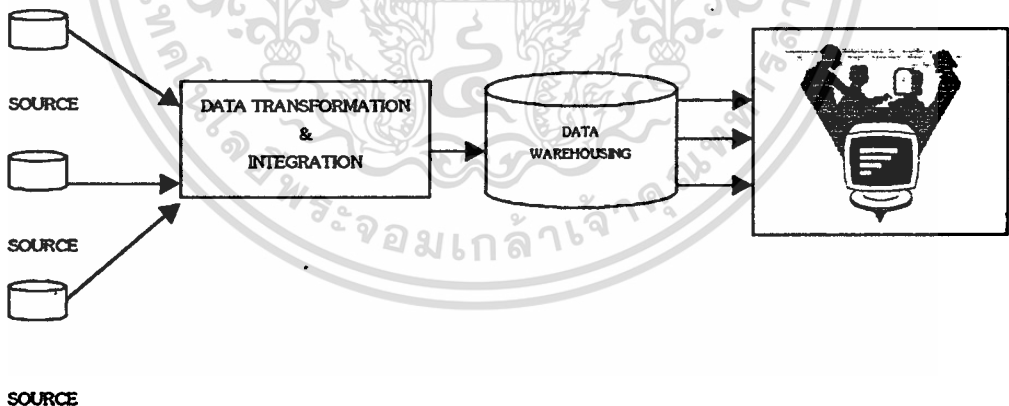
Visualization เป็นการนำเสนอข้อมูลในรูปแบบของกราฟฟิค การนำเสนอจะสามารถทำได้มากกว่า 2 มิติ ซึ่งจะสร้างความละเอียดของการนำเสนอและสร้างความเข้าใจให้มากขึ้น ตัวอย่างเช่น องค์กรต้องการที่จะหาสถานที่ในการตั้งสาขาขององค์กรในเขตพื้นที่ภาคเหนือของประเทศ ดังนั้นองค์กรจึงใช้รูปแบบที่มีการ Plot ที่ตั้งขององค์กรคู่แข่งที่มีสาขาค้างอยู่ในเขคนั้น เพื่อพิจารณาสถานที่ตั้งที่เหมาะสมที่สุด

3.4 Data Warehousing, Data Mining และ Data Mart

ก่อนที่เราจะไปไกลกว่านี้ เราควรจะมาทำความเข้าใจเกี่ยวกับคำว่า Data Warehousing เพราะว่า Data Mining มีความเกี่ยวข้องใกล้ชิดกับ Data Warehousing มาก แต่อย่างไรก็ตามในที่นี้ไม่ได้หมายความว่าถ้าไม่มี Data Warehousing แล้วจะทำให้ไม่มี Data Mining ความจริงแล้ว Data Warehousing เป็นตัวขับเคลื่อนหนึ่งซึ่งทำให้การใช้งาน Data Mining มีประสิทธิภาพสูงสุด

3.4.1 Data Warehousing

จากรูปที่ 3.3 เป็นขั้นตอนของการทำ Data Warehousing



รูปที่ 3.3 ตำแหน่งของ Data Warehousing

ขั้นตอนแรกก่อนที่จะทำ Data Mining Process คือการจัดขนาดของข้อมูลใหญ่ ๆ ให้อยู่ในรูปแบบที่ง่ายต่อการเข้าถึง, การเข้าไปใช้งานและการ Sort โดยผู้ใช้ การรวบรวมข้อมูลใน Process ของ Data Mining อาจจะยุ่งยาก เพราะว่าบ่อย ๆ พบว่าข้อมูลไม่อยู่ในรูปแบบที่เหมาะสมที่จะเข้าไปใช้งานได้

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

จุดประสงค์ของ Data Warehousing คือช่วยปรับปรุงประสิทธิภาพในการตัดสินใจที่เกี่ยวกับธุรกิจที่ข้องเกี่ยวกับตัวเลขจำนวนมาก พื้นฐานดังกล่าวตั้งอยู่บนหลักของ Informational Data (ข้อมูลที่ใช้จัดการองค์กร ซึ่งเป็นข้อมูลสรุปเพื่อการตัดสินใจ) แทนที่จะเป็น Operational Data (ข้อมูลที่ใช้ดำเนินกิจกรรมขององค์กร เช่น ข้อมูลของพวก Transaction ต่าง ๆ)

คำจำกัดความของ Data Warehousing คือการรวบรวมของข้อมูลเพื่อสนับสนุนการตัดสินใจของฝ่ายบริหาร ข้อมูลดังกล่าวถูกแบ่งเป็นระดับ ๆ หลายระดับ เพื่อให้เกิดความสามารถในการเข้าถึงข้อมูลได้อย่างรวดเร็ว

Subjected Oriented ข้อมูลใน Warehouse ถูกกำหนดในลักษณะ Business Term เช่น ถูกค้า, สินค้า, ราชอาณาจักรห้อยคชชช

Integrated Term ที่ใช้ใน Data Warehousing จะต้องถูกกำหนดให้สมบูรณ์เหมือนกันทั้งองค์กร และจะต้องถูกต้องกับแหล่งข้อมูลทั้งภายในและภายนอก

Time Variant ข้อมูลใน Data Warehousing เป็นลักษณะ Time Stamp ณ เวลาที่ข้อมูลถูกใส่เข้ามาหรือข้อมูลถูกสรุป ดังนั้นจะเป็นการบันทึกในลักษณะต่อเนื่องและมีประวัติและแนวโน้มการวิเคราะห์ที่เป็นไปได้

Non Volatile เมื่อถูกใส่เข้ามาใน Data Warehousing แล้ว ข้อมูลจะไม่ถูก Update อีก ดังนั้นจึงเป็นแหล่งที่มีรายงานถูกต้องและใช้วิเคราะห์เชิงเปรียบเทียบ

โดยมีเครื่องมือ 2 ตัวในการจัดการทำ Data Warehousing คือ (Data Transformation, Data Cleaning) และ End User Data Access เครื่องมือเหล่านี้จะทำให้มั่นใจว่า Data Warehousing จะมีความถูกต้องของข้อมูล แม่นยำ มีประสิทธิภาพและมีต้นทุนในการบริหารต่ำ

ขั้นตอนของการทำ Data Warehousing คร่าว ๆ จะเริ่มจากขั้นตอนต่อไปนี้

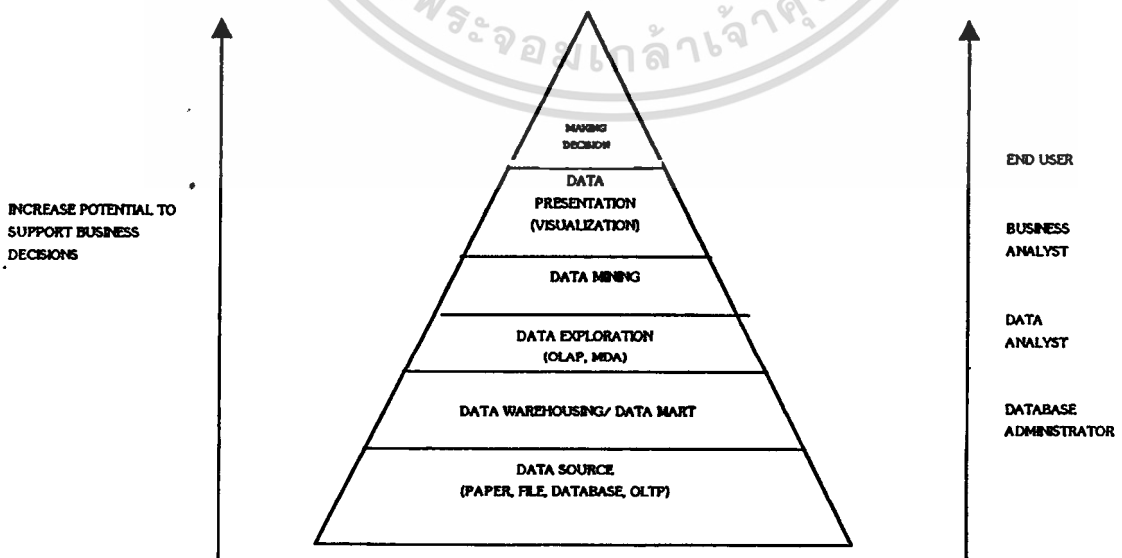
- Data Extraction จะช่วยสังเคราะห์เอาข้อมูลที่เป็นประโยชน์สำหรับ Data Mining เท่านั้น
- Sampling and Selecting จะเป็นตัวกำหนดขนาดของข้อมูล
- Aggregation จะเป็นตัวรวบรวมข้อมูลที่เกี่ยวข้องเข้าไว้ด้วยกัน
- Data Cleaning จะเป็นตัวสร้างความมั่นใจว่าข้อมูลจะสมบูรณ์ และลดความซ้ำซ้อนของข้อมูล
- Normalization จะเป็นตัวช่วยลดความซ้ำซ้อนของข้อมูล
- Overlay เช่น Demographic จะเป็นช่วยทำให้เร่ง Data Access ได้เร็วขึ้น

ปัญหาหลัก ๆ ของ Data Warehousing ก็คือคุณภาพของข้อมูล เพื่อหลีกเลี่ยงปัญหา GIGO (Garbage In Garbage Out) ข้อมูลควรมี Missing Value น้อยที่สุด เพราะอาจจะมีผลกระทบต่อ การวิเคราะห์ข้อมูลของ Data Mining ได้

3.4.2 Data Mart ปัจจุบันหลายองค์กรเริ่มหัน ไปหา Data Mart ซึ่งมีความเฉพาะเจาะจงมากกว่า และเข้าถึงได้มากกว่า แต่ขนาดเล็กลงกว่า Data Warehousing มาก Data Mart เป็นการแยกเก็บข้อมูลจาก Data Warehousing เพื่อเก็บข้อมูลให้กับแผนกเฉพาะที่มีการเรียกใช้ข้อมูลนั้น ๆ บ่อย เพื่อไม่ให้เกิดความซ้ำซ้อนและลดปริมาณข้อมูลที่ไม่เกี่ยวข้องทำให้การ Process ข้อมูลเร็วขึ้น

3.4.3 Data Mining ถือได้ว่าเป็นระดับการนำข้อมูล ไปใช้ที่สูงกว่า Data Warehousing และ Data Mart Data Mining เป็นวิธีคิดที่จะนำเอาข้อมูลมาใช้ในการวิเคราะห์ให้เกิดประโยชน์สูงสุด โดยเฉพาะอย่างยิ่งการตัดสินใจของฝ่ายบริหาร ซึ่งระบบนี้เป็นขั้นตอนต่อไปของ Data Warehousing มีระบบการทำงานอัตโนมัติ สามารถตัดสินใจแทนผู้ใช้ได้โดยอาศัยกฎเกณฑ์ต่าง ๆ ที่กำหนดขึ้นมาแล้วป้อนให้คอมพิวเตอร์คิด เครื่องมือทางธุรกิจ, เทคนิคต่างที่เราใช้เพื่อสนับสนุนการตัดสินใจทางธุรกิจนั้นมีพื้นฐานมาจาก เทคโนโลยีสารสนเทศ

จากรูปที่ 3.4 เริ่มต้นตั้งแต่ ตารางข้อมูลธรรมดา ไปจนถึงการตัดสินใจระดับสูง เราจะเห็นได้ว่า Data Mining เป็นส่วนประกอบอันใหม่ที่มีความสำคัญของเครื่องมือทางธุรกิจอย่างหนึ่ง คุณค่าของข้อมูลที่ใช้สนับสนุนการตัดสินใจจะเพิ่มขึ้นจากต่ำไปบนสุดของรูปปิรามิด จำนวนของข้อมูลและขนาด และระดับการตัดสินใจในข้อมูลมีลักษณะที่ต่าง ๆ กัน จึงมีระดับของผู้ตัดสินใจต่างกัน Database administrator จะตัดสินใจบนระดับของ Data Warehousing และแหล่งข้อมูลเท่านั้น ส่วนนักวิเคราะห์ธุรกิจและผู้บริหารจะตัดสินใจบนเหนือของปิรามิด



รูปที่ 3.4 Data Mining และ เครื่องมือทางธุรกิจต่าง ๆ (Cabena et al., 1997)

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

3.5 Data Mining Tasks and Techniques

1) Data Mining Tasks

ในงาน Data Mining นั้น สามารถแบ่งงานที่เกิดขึ้นจากการทำได้เป็น 6 ชนิดของงาน โดยงานแต่ละชนิดไม่สามารถบอกได้ว่าเทคนิคและเครื่องมือของ Data Mining ไດเหมาะสมเหมาะสมที่สุดกับงานแต่ละชนิด หลังจากที่ได้อธิบายงานทั้ง 6 อย่างแล้ว จะได้อธิบายถึงเครื่องมือที่เหมาะสมและสามารถนำไปใช้กับงานแต่ละงานได้

1.1 Classification และ Clustering อาจหมายถึงการแบ่งประเภท การแบ่งลำดับชั้น และการจัดลำดับ ของ Object การทำงานในลักษณะดังกล่าว คือการตรวจสอบลักษณะของ Object แต่ละชนิดแล้วจัดลำดับให้มันอยู่ในกลุ่มหรือชั้นที่มีลักษณะร่วมกัน ตัวอย่างของการทำ Classification และ Clustering เช่น

- การแบ่งเครดิตของผู้ขอทำบัตรเครดิต – มีความเสี่ยงต่ำ, กลาง และสูง
- สัญชาติของบุคคล
- กลุ่มของผู้มีอายุมากและร่ำรวย

1.2 Estimation ลักษณะของ Classification นั้นคำนึงถึงผลกำหนดที่ออกมาชัดเจนว่าคุณสมบัติดังกล่าวจะอยู่ในชั้นใด แต่ Estimation เป็นการประเมินที่ไม่สามารถกำหนดค่าหรือคุณสมบัติดังกล่าวให้ชัดเจน เป็นการจัดการกับค่าที่มีผลในการวัดที่ต่อเนื่อง ตัวอย่างเช่น

- การประเมินรายได้ของครอบครัว
- การประเมินความสูงของบุคคลในครอบครัว
- การประเมินจำนวนของเด็ก ๆ ในครอบครัว

1.3 Prediction เหมือนกับ Classification และ Estimation ยกเว้นว่า Record ที่ถูกแยกจัดลำดับนั้นเกิดขึ้นตามการทำนายพฤติกรรมในอนาคตหรือการทำนายค่าที่จะเกิดขึ้นในอนาคต ข้อมูลในอดีตจะถูกสร้างเป็น Model ขึ้นมาเพื่อทำนายหรืออธิบายสิ่งที่จะเกิดขึ้นในอนาคต ตัวอย่างเช่น

- การทำนายว่าลูกค้ากลุ่มใดที่องค์กรจะสูญเสียไปภายใน 6 เดือนหน้า
- การทำนายว่ายอดซื้อของลูกค้าจะเป็นเท่าใดถ้าบริษัทลดราคาสินค้า 10%

1.4 Affinity Grouping (Market Basket Analysis) Affinity Grouping คือการกำหนดว่าสิ่งของอะไรที่จะไปด้วยกันได้ โดย Concept ดังกล่าวได้ถูกนำไปใช้ในร้านซูเปอร์มาร์เก็ต ซึ่งเรียกเทคนิคดังกล่าวว่า “Market Basket Analysis” โดยจะกำหนดว่าสินค้าประเภทใดมักจะถูกซื้อควบคู่กันในการซื้อแต่ละครั้ง ทำให้ทางร้านสามารถกำหนดได้ว่าควรจะจัดเรียงสินค้าอย่างไร หรือควรจะเตรียมแค็ตตาล็อกเพื่อขายสินค้าอย่างไร นอกจากนี้ Affinity Grouping ยังสามารถ

กำหนดถึงเทคนิคการทำ Cross Selling โดยองค์กรสามารถกำหนดการขายสินค้าเป็น Package โดยนำเทคนิคดังกล่าวมาช่วยสร้างความน่าสนใจในการซื้อ Affinity Grouping จะเป็นตัวกำหนดกฎที่เกิจากข้อมูล “Associated Rules” ตัวอย่างเช่น

- คนที่ซื้ออาหารสุนัขมักจะซื้อกระดูกปลอม โดยมีโอกาสที่จะเกิดขึ้น เปอร์เซนต์
- คนที่ซื้อกระดูกปลอมมักจะซื้ออาหารสุนัข โดยมีโอกาสที่จะเกิดขึ้น เปอร์เซนต์

1.5 Description จุดประสงค์ของการทำ Data Mining การหาคำอธิบายถึงสิ่งที่จะเกิดขึ้น โดยอาศัยข้อมูลจากฐานข้อมูล ตัวอย่างเช่น

- ผู้หญิงจะสนับสนุนพรรคเดโมแครตมากกว่าผู้ชาย

2) Data Mining Tasks and Techniques

2.1 Memory Based Reasoning (MBR)

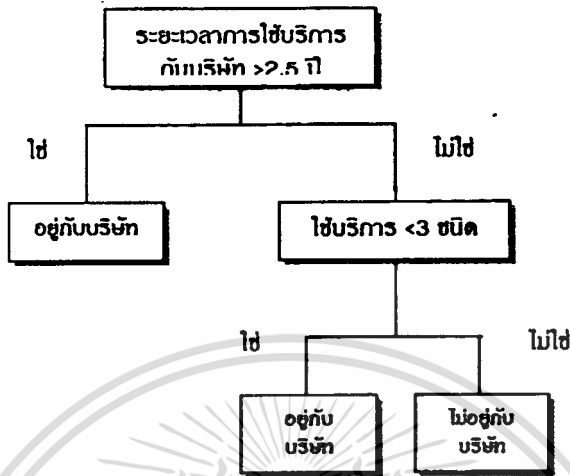
เปรียบเหมือนกับประสบการณ์การเรียนรู้ของมนุษย์ ซึ่งอาศัยการสังเกตเหตุการณ์ที่เกิดขึ้นแล้วสร้างรูปแบบของสิ่งนั้นขึ้นมา ใน Data Mining เราใช้ MBR เพื่อทำการวิเคราะห์ฐานข้อมูลที่มีอยู่ และกำหนดลักษณะพิเศษของข้อมูลที่อยู่ในนั้น แนนอนข้อมูลจะต้องมีลักษณะสมบูรณ์, การทำการสังเกตอย่างสมบูรณ์จะช่วยสร้างการทำนายอย่างละเอียดแม่นยำยิ่งขึ้น Model จะถูกบอกคำตอบที่ถูกต้องจากกรณีศึกษาที่ได้แก้ปัญหาไว้ก่อนหน้าแล้ว การทำงานแบบนี้วิธีนี้ถูกเรียกว่า “Supervised Learning”

ตัวอย่างของ นักวิเคราะห์ที่ต้องการเข้าใจทำไมลูกค้าบางกลุ่มซึ่งซื้อสัตย์แต่อีกกลุ่มบริษัทกลับเสียไป และนักวิเคราะห์จะทำนายว่าลูกค้าคนใดที่บริษัทกำลังจะเสียไปให้คู่แข่ง นักวิเคราะห์สามารถสร้าง Model จาก ข้อมูลในอดีต Model ที่ดีก็จะทำให้เรารู้ว่าลูกค้าคนใดจะอยู่กับบริษัทและลูกค้าคนใดจะเสียไป ตัวอย่างนี้เป็นขั้นตอนของการกำหนด “การศึกษา Study” การศึกษาจะเป็นตัวกำหนดขอบเขตของกิจกรรม การศึกษาจะกำหนด จุดประสงค์ให้ทั้งหมดและข้อมูลที่จะถูกใช้อาจจะไม่ต้องกำหนดจุดประสงค์ไว้ล่วงหน้า

จุดประสงค์ของการศึกษา คือ ต้องการเข้าใจว่าอะไรทำให้ลูกค้าอยู่กับบริษัทและจากบริษัทไป-จุดประสงค์นี้แตกต่างจากการถามคำถามเฉพาะเพราะเราไม่ได้กำหนดความสัมพันธ์เอาไว้เลย เทคนิคในการทำ MBR จะมีจุดประสงค์หลักคือการทำนายการคาดเดาอย่างมีหลักการเกี่ยวกับตัวแปรที่สนใจ โดยมักจะใช้เทคนิคของ Neural Network และ Decision Tree

อธิบายรูปแบบของการทำ MBR จากตัวอย่างในรูปที่ 3.5 คือบริษัทประกันภัยซึ่งมีความสนใจที่ทราบสาเหตุของการลดลงของจำนวนลูกค้าว่า โดย MBR กำหนด 2 ตัวแปรที่สนใจ คือ ระยะเวลาที่ลูกค้าอยู่กับบริษัท (ถือกรมธรรม์) และ จำนวนของบริการของบริษัทที่ลูกค้าใช้บริการ

อยู่ จะเห็นได้ชัดว่าลูกค้าที่อยู่กับบริษัทน้อยกว่า 2 ปีครึ่งและใช้บริการน้อยกว่า 3 บริการมักจะหนีไปใช้บริการของบริษัทอื่น ๆ



รูปที่ 3.5 รูปแบบการตัดสินใจของบริษัทประกันภัย (Cabena et al., 1997)

การทำงานของ MBR ตั้งอยู่บนรากฐาน 2 ประการคือ การทำ Classification และ การทำ Value Prediction ตัวอย่างของ Classification เช่น บริษัทที่ต้องการทำ Sales Promotion ซึ่งจะใช้ Mailing List จากฐานข้อมูลการซื้อของลูกค้า Mailing List ซึ่งมีการตอบรับกลับมาจากการส่ง Mail ไปครั้งก่อนหน้าจะมีการกำหนดเป็น Classification (Classification ถูกใช้กำหนดชั้นของช่องแต่ละ Record ในฐานข้อมูล จากตัวอย่าง คือ การอยู่กับบริษัท และ การไม่อยู่กับบริษัท) หรือ Profile Classification ดังกล่าวจะถูกเก็บรวบรวมและพัฒนาเพื่อที่จะบอกลักษณะของผู้ที่ตอบรับ Mail เพื่อกำหนดเป็นตัวทำนายที่จะตอบรับ และจะนำเอารายชื่อดังกล่าวไปส่ง

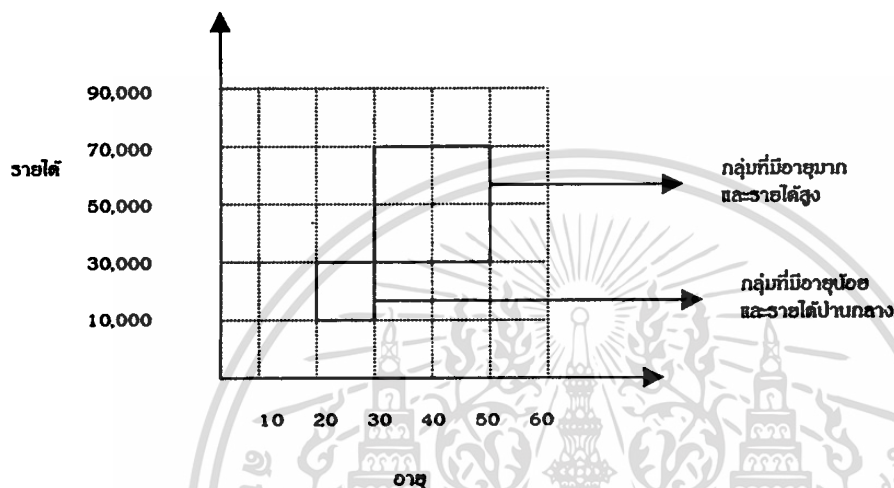
ส่วน Value Prediction จุดประสงค์เพื่อกำหนดความต่อเนื่องของมูลค่าซึ่งมีความเกี่ยวข้องกับ Record ในฐานข้อมูล ตัวอย่างเช่นการศึกษา Lifetime Customer การทำ Mining ก็จะศึกษาถึงข้อมูลที่ผ่านมาในอดีตของลูกค้าพร้อมไปถึงสถานะทางการเงินของลูกค้าในด้านนี้ยังมีตัวแปรอื่น ๆ อีก เช่น จำนวนครอบครัว รายได้ ประวัติการใช้จ่าย MBR เป็นรูปแบบที่มักจะถูกใช้อย่างกว้าง ๆ ในอุตสาหกรรมทั่ว ๆ ไป ทางธุรกิจนั้นมักจะใช้กับ Customer Retention Management, Credit Approval, Cross Selling และ Target Marketing

2.2 Cluster Detection

จุดประสงค์ของ Cluster Detection คือการแบ่งฐานข้อมูลออกเป็นส่วน ๆ หรือเราเรียกว่า Segment คือกลุ่มของ Record ที่มีความเหมือนและลักษณะที่คล้ายกัน หรือเรียกว่า "Homogeneity" ส่วน Record ที่อยู่ใน Segment อื่น ๆ ก็จะมีการแตกต่างกันหรือเรียกกลุ่มที่อยู่นอก Segment ว่า

“Heterogeneity” Cluster Detection ถูกใช้เพื่อค้นหา Sub Group ที่เหมือน ๆ กันในฐานข้อมูล เพื่อที่จะเพิ่มความถูกต้องในการวิเคราะห์ และสามารถมุ่งไปยังกลุ่มเป้าหมายได้ถูกต้อง

เรานำกราฟมาอธิบายกลุ่มของประชากรโดยเปรียบเทียบรายได้และอายุ ในรูปที่ 3.6 จะเห็นว่ากลุ่มหนึ่งเป็นกลุ่มที่มีอายุและมีรายได้สูง ส่วนอีกกลุ่มหนึ่งอายุน้อยรายได้ปานกลาง มีการศึกษา มีการแบ่งข้อมูลออกเป็น 2 Segment



รูปที่ 3.6 การแบ่งกลุ่ม Target Group ในรูปแบบของ Clustering (Cabena et al., 1997)

เทคนิค Cluster Detection เป็นวิธีการรวมกลุ่มของแถวของข้อมูลซึ่งมีส่วนร่วมที่คล้ายกันแนวโน้มและรูปแบบ Clustering Studies ไม่มี “Dependent Variable” ดังนั้นจึงไม่สามารถศึกษาได้ลงไปอย่างเจาะจง ไม่สามารถทำให้เกิดผลที่แน่นอน เราจึงเรียกการศึกษานี้ว่าเป็น “Unsupervised Learning” ตัวอย่างเช่น เราต้องการทราบว่าอะไรที่เหมือนกันในกลุ่มฐานลูกค้าของบริษัท เทคนิค Clustering ก็จะทำให้การจำแนกแยกกลุ่มให้

Cluster Detection แตกต่างจาก Data Mining เทคนิคอื่น ๆ คือ จุดประสงค์ค่อนข้างคลุมเครือเมื่อเทียบกับเทคนิคของ Data Mining ตัวอื่น ๆ

2.3 Link Analysis

Link Analysis มุ่งเน้นทำงานบน Record คือความสัมพันธ์ หรือความเกี่ยวข้องกันระหว่าง Record หรือกลุ่มของ Record ความสัมพันธ์ดังกล่าวเรียกว่า “Association” เทคนิคนี้มุ่งมองไปที่รูปแบบการซื้อหรือเหตุการณ์ที่เกิดขึ้นเป็นลำดับ โดยมีเทคนิคที่ใช้บน Link Analysis อยู่ 3 อย่าง

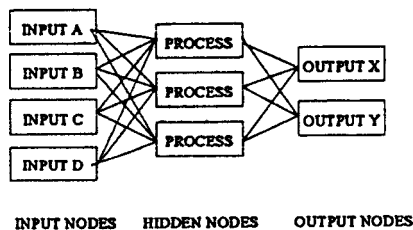
- Association Discovery ใช้วิเคราะห์การซื้อสินค้าภายในรายการเดียวกัน ศึกษาถึงความสัมพันธ์อย่างใกล้ชิดที่ถูกปิดซ่อนอยู่ของสินค้า ซึ่งสินค้าเหล่านั้นมักมีแนวโน้มที่จะถูกซื้อควบ

คู่กันไป การวิเคราะห์แบบนี้เรียกว่า “Market Basket Analysis” คือรายการทั้งหมดที่ถูกค้าซื้อ ต่อครั้งที่ซูเปอร์มาร์เก็ต สามารถใช้ Input Device โดยใช้ Bar Code Scanner มีหลายงานด้วยกัน เช่น ซูเปอร์มาร์เก็ต การเตรียม Inventory การวางแผนการเรียง Shelf การทำ Mailing List สำหรับ Direct Mail และการวางแผนเพื่อจัด Promotion สนับสนุนการขาย ตัวอย่างของ Association เช่น อาจพบว่า 75 เปอร์เซ็นต์ของผู้ซื้อน้ำอัดลมจะซื้อข้าวโพดคั่วด้วย

- Sequential Pattern Discovery ถูกใช้ระบุความเกี่ยวเนื่องกันของการซื้อสินค้าอย่างหนึ่งและจะซื้อสินค้าอีกอย่างหนึ่งในเวลาต่อมา และดูลำดับการเลือกซื้อสินค้าของถูกค้ามันมีจุดมุ่งหมายที่จะเข้าใจพฤติกรรมของการซื้อสินค้าของถูกค้าในลักษณะ Long Term เช่นผู้ขายอาจพบว่าถูกค้าที่ซื้อทีวีมีแนวโน้มที่จะซื้อวิดีโอในเวลาต่อมา
- Similar Time Sequence Discovery ใช้ค้นหาความเกี่ยวเนื่องกันระหว่างกลุ่มของข้อมูล 2 กลุ่ม ซึ่งมีการขึ้นต่อกันทางด้านเวลา โดยมีรูปแบบการเคลื่อนที่เหมือนกัน ผู้ขายสินค้ามักจะใช้เพื่อดูแนวโน้มเพื่อเตรียมสต็อก เช่นเมื่อ ไร้กี้ดามที่ชอคขายสินค้าน้ำอัดลมสูงขึ้น ชอคขายมันฝรั่งจะสูงขึ้นตาม

2.4 Artificial Neural Networks

เป็นการเลียนแบบการทำงานของระบบประสาทเทียม ซึ่งเลียนแบบการทำงานของระบบประสาทในสมองของมนุษย์ จากรูปที่ 3.7 จะอธิบายถึงการทำงานของ Neural Network แต่ละ Process จะรับ Input เข้าไปคำนวณ และสร้าง Output ออกมาในลักษณะที่ไม่ใช่เป็นการทำงานแบบเชิงเส้นตรง เพราะว่า Input แต่ละตัวจะถูกให้ลำดับความสำคัญของค่าไม่เท่ากัน ค่าของ Output ที่ได้จากการเชื่อมโยงกันนี้จะถูกนำมาเปรียบเทียบกับ Output ที่ได้ตั้งเอาไว้ ถ้าค่าที่ออกมาเกิดความคลาดเคลื่อน ก็จะนำไปสู่การปรับค่าน้ำหนักของค่าที่ใส่ไว้ให้แต่ละ Input



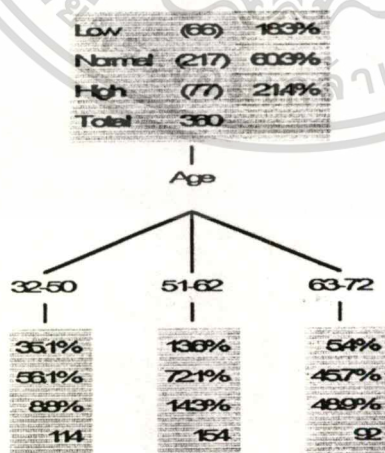
รูปที่ 3.7 รูปแบบการทำงานของ Neural Network (Groth, 1997)

2.5 Genetic Algorithm

เปรียบเสมือนเป็นการสร้างพันธุกรรมที่ดีที่สุดบนขั้นตอนของวิวัฒนาการทางชีวภาพ แนวคิดหลักก็คือเมื่อเวลาผ่านไป วิวัฒนาการของเซลล์ชีวิตจะเลือกสายพันธุ์ที่ดีที่สุด “Fittest Species” Genetic Algorithm จะมีความสามารถในการทำงานแบบรวมกลุ่มข้อมูลเข้าด้วยกัน เช่น อาจมีความต้องการที่จะแบ่งกลุ่มหรือจับรวมกลุ่มของข้อมูลเป็น 3 ชุด ขั้นตอนการทำงานของ Genetic Algorithm ก็จะเริ่มด้วยการจับกลุ่มข้อมูลเป็นกลุ่ม ๆ ด้วยการเคาะสุ่ม เปรียบเหมือนกลุ่ม 3 กลุ่มนี้เป็นเซลล์ของสิ่งมีชีวิต Genetic Algorithm จะมี “Fittest Function” ที่จะบอกว่ากลุ่มข้อมูลใดเหมาะกับกลุ่ม ๆ ไหน โดย Fittest Function จะเป็นตัวบ่งชี้ว่าข้อมูลเหมาะกับกลุ่มมากกว่าข้อมูลอื่น ๆ นอกจากนี้ในขั้นตอนต่อมา Genetic Algorithm จะมี “Operator” ซึ่งยอมให้มีการเลียนแบบและแก้ไขลักษณะของกลุ่มของข้อมูล Operator จะจำลองหน้าที่ของชีวิตที่ถูกรพบในธรรมชาติ คือชีวิตมีการแพร่พันธุ์, จับคู่ผสมพันธุ์ และเปลี่ยนรูปร่างตามต้นแบบของพันธุ์ เปรียบกับข้อมูลถ้ามีข้อมูลใดในกลุ่มของข้อมูลถูกพบว่าตรงกับคุณสมบัติของ Fittest Function แล้ว มันจะคงอยู่และถูกถ่ายเข้าไปในกลุ่มนั้น แต่ถ้าไม่ตรงกับคุณสมบัติก็ยังมีโอกาสที่จะถ่ายข้ามไปยังกลุ่มอื่น ๆ ได้

2.6 Decision Trees

เป็นรูปแบบการสร้างโครงสร้างแบบต้นไม้เพื่ออธิบายกลุ่มของข้อมูล โดยมีการกำหนด Node ที่อยู่เหนือสุดของ Decision Tree ให้เป็น Root Node และ Node ที่เรียงต่อกันลงมาเป็น Child Node แต่ละ Path ที่เรียงต่อกันมาจะมีลักษณะที่ไม่เหมือนกันเลขขึ้นอยู่กับสิ่งที่องค์กรตั้งใจจะศึกษา ตัวอย่างของ Decision Tree ดังในรูป 3.8 เป็นการศึกษาถึงผลของอายุที่มีต่อโรคความดันสูง



รูปที่ 3.8 รูปแบบการทำงานของ Decision Tree (Groth, 1997)

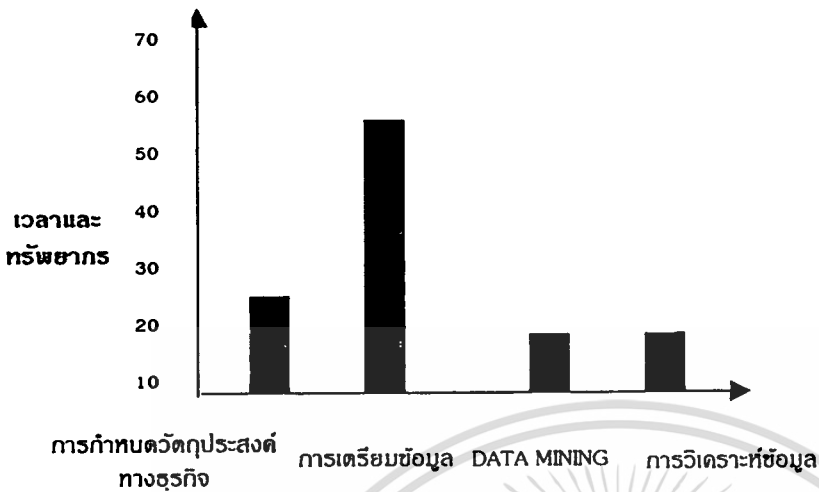
มี Algorithm หลักในการทำงานของ Decision Tree อยู่ด้วยกันหลายชนิด

- CART (Classification and Regression Trees) เป็นวิธีการสร้าง Decision Tree ที่ได้รับความนิยมมากที่สุด เป็นการทำงานในแบบ Binary Tree คือมีการแตกกิ่งออกเป็น 2 กิ่งเสมอ การทำงานจะต้องมีการกำหนด Dependent Variable หรือตัวแปรที่ต้องการจะศึกษาเสียก่อน แล้วจึงนำ Independent Variable มาแตกออกเป็นต้นไม้โดยเป็นการปลูกออกมาให้เต็มต้น เสร็จแล้วจึงทำการตัด Node โดยมีการตรวจสอบ Node ที่มี Error Rate สูงออกไป แล้วจึงเลือก Node ที่นำเสนอความเกี่ยวข้องกับเรื่องที่จะศึกษามากที่สุด
- C4.5 มีลักษณะคล้ายกับ CART คือมีการอธิบายถึงที่มาและที่ไปของ Node แต่ละ Node แต่สิ่งที่แตกต่างที่เห็นได้ชัดก็คือ CART ทำงานบน Binary Node ส่วน C4.5 ทำงานบนจำนวน Node หลาย ๆ Node ที่แตกกิ่งออกมาแต่ละครั้ง อีกความแตกต่างกันก็คือในการตัดกิ่งออกนั้น C4.5 ไม่จำเป็นต้องมีการตรวจสอบ Node ก่อน
- CHAID (Chi-Square Automatic Interaction Detection) ลักษณะการทำงานของ CHAID แตกต่างจาก 2 ชนิดแรกเนื่องจาก CHAID พยายามที่จะหยุดการเจริญเติบโตของต้นไม้ก่อนที่มันจะโตเต็มที่ แล้วค่อย ๆ สร้าง Child Node ที่เหมาะสมจนต้นไม้โตเต็มที่ สิ่งที่แตกต่างกัน 2 ชนิดแรกก็คือ CHAID ทำงานบน Data ที่เป็น Categorical เท่านั้น

3.6 ขั้นตอนการทำงานของ Data Mining (Data Mining Process)

ปกติคนมักจะเริ่มไปที่ Data Mining เลย แต่ขั้นตอนในการทำ Mining นั้น Process ทั้งหมดมีหลายหลายขั้นตอน Mining เป็นเพียงขั้นตอนหนึ่งเท่านั้น ขั้นตอนการทำ Data Mining นั้นถูกอธิบายว่าเป็นขั้นตอนของการสร้าง Model โดยสร้าง Model ของกลุ่มข้อมูล (Data Set) เพื่อสร้างความเข้าใจในแนวโน้ม, รูปแบบ และความสัมพันธ์กันของกลุ่มข้อมูล เพื่อใช้ในการทำนายบนข้อมูลนั้น ๆ

จากขั้นตอนของการทำ Data Mining จุดประสงค์ทางธุรกิจจะเป็นตัวกำหนดทิศทางของขั้นตอนการทำ Data Mining ทั้งหมด นอกจากนี้ขั้นตอนของการทำ Data Mining นั้นอาจจะดูเหมือนถูกเรียงลำดับกันแต่ว่าขั้นตอนจริง ๆ ของการทำมีลักษณะเป็นขั้นตอนที่ต้องการการย้อนกลับไปกลับมาตลอดเวลา คือมีการทำซ้ำอยู่ตลอดเวลา นอกจากนี้ในแต่ละขั้นตอนยังเป็นขั้นตอนที่ต้องใช้เวลาและความพยายามและทรัพยากรที่แตกต่างกันไปด้วย โดยสังเกตจากรูปที่ 3.9 จะเห็นว่าขั้นตอนของการเตรียมข้อมูลจะใช้ทรัพยากรสูงที่สุด



รูปที่ 3.9 เวลาในแต่ละขั้นตอนของการทำ Data Mining (Cabena et al., 1997)

1) การกำหนดจุดประสงค์ทางธุรกิจ (Business Objective Determination)

การกำหนดปัญหาให้ชัดเจนจะเป็นตัวกำหนดทิศทางการทำ Data Mining นอกจากนี้ยังต้องมีการกำหนดปัญหาทางธุรกิจ โอกาส เพื่อดูว่าจะมีงานที่ต้องใช้ Data Mining หรือไม่เพราะไม่ได้หมายความว่าทุกปัญหาสามารถแก้ได้ด้วยเทคนิค Data Mining โดยบางองค์กรมีความคาดหวังในเทคนิคของ Data Mining มากเกินไป ซึ่งอาจจะนำมาสู่ความเข้าใจผิดได้ ดังนั้นการกำหนดปัญหาไม่ถูกต้องย่อมยากที่จะนำไปสู่ความสำเร็จในการแก้ปัญหา

การกำหนดคำจำกัดความและทำความเข้าใจกับความต้องการทางธุรกิจไม่่ง่ายนัก มันต้องการความร่วมมือของหลาย ๆ ฝ่าย เช่น นักวิเคราะห์ทางธุรกิจ, ฝ่ายขายฝ่ายการตลาด, ผู้มีความรู้ความเข้าใจเกี่ยวกับธุรกิจ, นักวิเคราะห์ข้อมูล เพื่อที่จะแปลงจุดประสงค์ให้เป็งาน Data Mining ที่ประสบความสำเร็จ

ในขั้นนี้เองนักวิเคราะห์ธุรกิจและนักวิเคราะห์ข้อมูลสามารถเริ่มต้นเตรียมการกำหนด วิธีการทำ Data Mining และ ฐานข้อมูล ซึ่งเกี่ยวข้องกับจุดประสงค์ทางธุรกิจ

2) การเตรียมข้อมูล (Data Preparation)

เป็นหัวใจของขั้นตอนในการทำทั้งหมด เป็นช่วงที่ใช้เวลามากที่สุดในขั้นตอน โดยปกติแล้วต้องการเวลาประมาณ 60% ของเวลาทั้งหมดในการเตรียมข้อมูล ในขั้นตอนนี้อาจสามารถแบ่งออกได้เป็น 3 ขั้นตอนย่อยดังต่อไปนี้

2.1) การเลือกข้อมูล (Data Selection).

จุดประสงค์คือการระบุแหล่งของข้อมูลที่มี และทำการดึงเอาข้อมูลออกมาใช้สำหรับการวิเคราะห์เบื้องต้นในการเตรียมตัวสำหรับการที่จะทำการ Mining ในขั้นต่อไป การเลือกข้อมูลนั้นจะแตกต่างกันไปตามจุดประสงค์ของแต่ละธุรกิจที่ได้กำหนดไว้แต่ต้น และการเลือกข้อมูลก็ยังคงถูกกำหนดโดยลักษณะงานที่จะถูกนำมาใช้อีกด้วย

ตัวแปรที่ถูกเลือกมาแต่ละตัวนั้นจะต้องถูกทำความเข้าใจว่าตัวแปรแต่ละตัวหมายความว่าอะไร ประกอบด้วยอะไร นอกจากนี้ยังรวมไปถึงไม่เพียงแต่ค่าจำกัดความทางธุรกิจเท่านั้นแต่จะต้องมีคำอธิบายอย่างชัดเจนเกี่ยวกับชนิดของข้อมูล, ค่าที่เป็นไปได้, แหล่งกำเนิดของข้อมูล, รูปแบบของข้อมูล และลักษณะอื่น ๆ จะมี 3 ชนิดของตัวแปร

- ตัวแปรแบบ Categorical

- Nominal Variable กล่าวถึงชนิดของ Object ที่มันอ้างถึงแต่ไม่มีลำดับในค่าที่เป็นไปได้ (Possible Value) ตัวอย่างเช่น สถานะการแต่งงาน (โสด, แต่งงาน, หย่า, ไม่ทราบ), เพศ (ชาย, หญิง), ระดับการศึกษา (ปริญญาโท, ปริญญาตรี, ม. ปลาย, ปวช)
- Ordinal Variable มีลำดับสำหรับค่าที่เป็นไปได้ ตัวอย่างเช่น ลำดับของถูกค้า (ดี, ปานกลาง, ไม่ดี)

- ตัวแปรแบบ Quantitative ซึ่งมีการวัดความแตกต่างระหว่างค่าที่เป็นไปได้

- Continuous (ค่าที่ต่อเนื่อง) เช่น รายได้, เฉลี่ยจำนวนครั้งที่ซื้อ, รายได้
- Discrete (ค่าเป็นจำนวนเต็ม) เช่น จำนวนพนักงาน, เวลาปี (เดือน, ฤดู, ไตรมาส)

ตัวแปรของข้อมูลมีหลายตัวมากแต่ตัวแปรที่ถูกเลือกสำหรับทำ Data Mining นั้นถูกเรียกว่า “Active Variables” เพราะว่ามันจะถูกใช้สร้างความแตกต่างของกลุ่มย่อยต่าง ๆ และสามารถถูกนำมาทำนายผลได้ เมื่อคุณทำการเลือกข้อมูลจะต้องระวังเรื่องอายุของข้อมูลด้วย เพราะว่าสถานการณ์ภายนอกเปลี่ยนแปลงตลอดเวลาซึ่งจะทำให้ประสิทธิภาพของการทำ Mining ลดลง ตัวอย่าง รสนิยมการใช้ชีวิต, การเปลี่ยนงาน

2.2) การกลั่นกรองข้อมูล (Data Preprocessing)

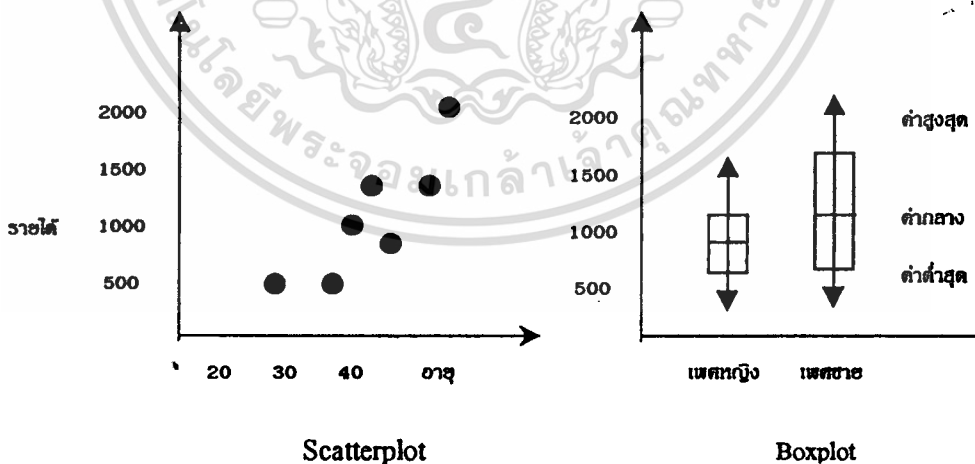
จุดประสงค์ก็เพื่อให้มั่นใจว่าคุณภาพของข้อมูลที่ถูกเลือกนั้นเหมาะสม ข้อมูลที่สมบูรณ์เป็นเครื่องประกันว่าการทำ Data Mining จะสำเร็จ ในขั้นตอนนี้เป็นขั้นตอนที่มีปัญหามากกว่าในขั้นตอนของการเตรียมข้อมูล เพราะข้อมูลส่วนใหญ่ที่มีในองค์กรไม่ได้ถูกเตรียมมาเพื่องาน Data Mining โดยเฉพาะ ข้อมูลถูกนำมาจากแหล่งต่าง ๆ ถูกจัดเก็บไม่ดี ข้อมูลที่ถูกนำมาจากภายนอกแล้วนำมาเพื่อให้เข้ากับข้อมูลภายในที่มีอยู่ ปัญหาหลักของ Data คือ คุณภาพและ Data Integrity

ในขั้นตอนนี้อ่อนอื่นจะต้องทำการทบทวนโครงสร้างของข้อมูลใหม่ และวัดคุณภาพของมัน โดยใช้วิธีทางสถิติ หรือ สุ่มตัวอย่าง

เครื่องมือที่ใช้ในการทำการกลั่นกรองข้อมูลมีดังต่อไปนี้

- ค่าตัวแปรเป็นแบบ Categorical การแบ่งความถี่ของค่าจะเป็นวิธีที่ทำให้เกิดความเข้าใจใน Data Content เครื่องมือที่ใช้เป็นเครื่องมือทางด้านกราฟฟิคจะเป็นตัวช่วยให้เห็นและกำหนดค่าที่หายไปได้
- ตัวแปรแบบ Quantitative ตัวแปรประเภทนี้มักมีการใช้การวัดตัวอย่างเช่น ค่าสูงสุด, ค่าต่ำสุด, ค่าเฉลี่ย, ค่ากลาง, ค่ามัธยฐาน และค่าอื่น ๆ ทางสถิติ เมื่อนำค่าพวกนี้มาเข้าสู่ตรรกศาสตร์ก็จะบอกถึงค่าที่ไม่สมบูรณ์ หรือค่าที่มีปัญหา

เครื่องมือทางกราฟฟิคอื่น ๆ เช่น Scatterplots คือรูป 2 มิติซึ่งแสดงความสัมพันธ์ระหว่างตัวแปร 2 ตัวแปรขึ้นไปหรือมากกว่า จากกราฟตัวอย่างจะเห็นได้ว่ามีการเปรียบเทียบรายได้กับอายุ จะเห็นได้ว่าจุดจะอยู่สูงขึ้นตามระดับของอายุ ทำให้เราพอที่จะทำนายได้ว่ารายได้ของอาชีพนี้จะสูงขึ้นเมื่ออายุสูงขึ้น ส่วน Boxplot ถูกใช้ให้เป็นประโยชน์สำหรับเปรียบเทียบศูนย์กลาง (ค่าเฉลี่ย) หรือ การกระจาย (ค่าเบี่ยงเบน) ของตัวแปรตั้งแต่ 2 ตัวแปรขึ้นไป จากกราฟตัวอย่างในรูปที่ 3.10 อธิบายถึงรายได้ของผู้ชายและผู้หญิง รูปสี่เหลี่ยมคือเรียกว่า Box และเส้นตั้ง 2 เส้นเรียกว่า Whisker จากความสูงของ Box พอจะสรุปได้ว่ารายได้ของผู้ชายสูงกว่าผู้หญิง



รูปที่ 3.10 เครื่องมือทางกราฟฟิค Scatterplot และ Boxplot (Cabena et al., 1997)

ระหว่างการทำขั้นตอนการกลั่นกรองข้อมูลจะมีปัญหาบ่อย ๆ ที่มักพบได้ ได้แก่ **Noisy Data** คือตัวแปรตัวหนึ่งหรือมากกว่ามีค่าซึ่งเกินกว่าค่าที่เราได้คาดไว้ ซึ่งอาจจะหมายถึงแง่ดีหรือแง่ร้ายก็ได้ ในแง่ดีก็คือมันจะแสดงอย่างชัดเจนถึงโอกาสซึ่งเรากำลังมองหาอยู่

ในแง่ร้ายคือมันอาจจะเป็นข้อมูลที่ไม่วสมบูรณ์ สาเหตุที่เกิดขึ้นได้ อาจจะมาจากการผิดพลาดของมนุษย์ ตัวอย่าง เช่น Operator ใส่อายุให้คน เป็น 300 ปี หรือใส่ค่าของรายได้เป็นติดลบ ค่าเหล่านี้ควรจะถูกแก้ไขหรือเอาออกจากการวิเคราะห์ ควรมีขั้นตอนการเช็คข้อมูลก่อนนำมาใช้

ค่าที่หายไป Missing Value คือค่าที่ไม่ได้แสดงในข้อมูลที่เราได้เลือกแล้ว หรือค่าที่ไม่สมบูรณ์ที่เราลบออกไประหว่างการทำ Noise Detection ค่าอาจจะหายไปเพราะเกิดจากความผิดพลาดของมนุษย์ เพราะว่าไม่มีข้อมูลนั้นระหว่างการทำ Input ข้อมูล การจัดการกับค่าที่หายไปนั้นสามารถจัดการได้ด้วยเทคนิคที่ต่าง ๆ กัน

2.3) การแปลงข้อมูล (Data Transformation)

ระหว่างขั้นตอนของการแปลงข้อมูล ข้อมูลที่ได้กลั่นกรองแล้วจะถูกแปลงให้เป็นรูปแบบของข้อมูลที่พร้อมจะถูกวิเคราะห์ รูปแบบของข้อมูลที่จะถูกวิเคราะห์คือรูปแบบของข้อมูลที่ไม่มีความขัดแย้ง ถูกจัดระเบียบมาอย่างเรียบร้อย กลั่นกรองมาจากแหล่งข้อมูลภายนอกและภายใน

ขั้นตอนนี้เป็นขั้นตอนที่สำคัญมากเนื่องจากความถูกต้องและสมบูรณ์ของผลลัพธ์สุดท้าย ซึ่งขึ้นอยู่กับว่านักวิเคราะห์ข้อมูลนั้นตัดสินใจกำหนดโครงสร้างและเสนอลักษณะของ Input อย่างไร ตัวอย่างเช่น หลักจากรูปแบบของข้อมูลถูกกำหนดแล้ว ข้อมูลที่ถูกกลั่นกรองจะเหมาะสมกับรูปแบบเฉพาะสำหรับแต่ละ กรรมวิธีของ Data Mining ที่จะถูกใช้ การแปลงข้อมูลยังรวมไปถึงการทำ Data Recoding และ Data Format Conversion เช่นการแปลงวันที่ เป็นต้น

ทางสถิติการทำกรแปลงข้อมูลยังมีเทคนิคของ Data Reduction จุดประสงค์เพื่อที่จะลดตัวแปรสำหรับการทำการ Process โดยการนำเอาตัวแปรตั้งแต่ 2 ตัวขึ้นไปมารวมกันแล้วทำการ Process ข้อดีก็คือลดจำนวนของตัวแปรลง และยังสามารถจัดการได้ง่ายขึ้น

อีกเทคนิคเรียกว่า Discretization โดยการแปลงตัวแปรแบบ Quantitative ให้เป็นแบบ Categorical โดยการแบ่งค่าของตัวแปรที่จะเป็น Input ให้เป็นช่วง ๆ เช่นการแปลงเงินเดือน อายุ

อีกเทคนิคเรียกว่า One of N โดยการแปลงตัวแปรแบบ Categorical ให้เป็น Numeric ตัวอย่างเช่น ชนิดของรถ Ford, Lincoln, Nissan ให้เป็น 100, 010, 001 ปกติแบบนี้มักจะเป็น Input ของพวก Neural Network

2.4) ขั้นตอน Data Mining

ในขั้นตอนนี้จะเริ่มทำ Data Mining กัน จุดประสงค์คือการนำเอา วิธีการทำ Data Mining ที่ได้เลือกไว้มาเริ่มทำ หรือ วิธีที่จะวิเคราะห์ข้อมูล จริง ๆ แล้วเราแยก Data Mining ออกมาเป็นอีกขั้นหนึ่ง แต่มองในแง่การทำงานแล้วค่อนข้างจะแยกออกได้ลำบากจากขั้นอื่น ๆ เนื่องจากจะต้องมีการทำงานที่เกี่ยวข้องกันและซ้อนไปซ้อนมา

โดยทางภาพภาษานั้นมีการใช้กรรมวิธี Data Mining จำนวนหนึ่งร่วมกัน ดังนั้นนักวิเคราะห์ข้อมูลจะต้องดูดูแลหลาย ๆ เรื่อง ได้แก่ โครงสร้างข้อมูล, ความหมายที่เกิดขึ้น โดยจะต้องดูแลข้อมูลที่ได้กลั่นกรองแล้ว ยังมี Meta Data รูปแบบของ Application ที่หลากหลายจะเป็นตัวกำหนดว่า จะต้องใช้กรรมวิธีทาง Data Mining กี่วิธีในการวิเคราะห์ปัญหานั้น ๆ

การพัฒนางาน Data Mining มีการใช้วิธีการทำงานหลายอย่าง (Algorithm) Data Analyst จะต้องเลือก Algorithm ที่เหมาะสมกับลักษณะของจุดประสงค์ ซึ่งแต่ละ Algorithm มีข้อดีและข้อเสียในตัวเอง ซึ่งเราจะกล่าวถึงกรรมวิธีในการทำ Data Mining อีกส่วน

2.5) Analysis of Result

การวิเคราะห์ผลในการทำ Data Mining เป็นขั้นตอนที่สำคัญมากที่สุดในทุกขั้นตอนทั้งหมด เครื่องมือทางด้าน Graphical Visualization จะช่วยทำให้นักวิเคราะห์ธุรกิจและนักวิเคราะห์ข้อมูลสามารถทำการวิเคราะห์ได้อย่างสะดวกรวดเร็วยิ่งขึ้น

แม้ขั้นตอนของการวิเคราะห์ผลเองก็ไม่สามารถแยกออกจากขั้นตอนของการทำ Data Mining ได้ เพราะมีขั้นตอนการทำงานที่เกี่ยวข้องกัน ใน Software ของ Data Mining หลาย ๆ ตัวมีเครื่องมือประเภทนี้อยู่ด้วย งานในขั้นนี้จะเฉพาะเจาะจงขนาดไหนขึ้นอยู่กับชนิดของงานที่ทำ Application ที่จะพัฒนา ในขั้นนี้ นักวิเคราะห์ข้อมูลจะถามตัวเองว่าผลที่ออกมา น่าสนใจ, สมบูรณ์ และสามารถนำไปปฏิบัติจริงได้หรือไม่ ถ้าคำตอบคือไม่ จะต้องการวนกลับไปทำงานในขั้นก่อนแล้วย้อนกลับมา ก็เป็นตัวชี้ข้อหนึ่งว่าวิธีการแบบนี้แตกต่างจากวิธีการวิเคราะห์ทางสถิติแบบดั้งเดิมเพราะทางสถิตินั้นจะพิสูจน์ว่าสมมติฐานนั้นจริงหรือไม่ แต่คำตอบที่ได้จาก Data Mining จะให้แนวทางในการปฏิบัติมากกว่า

ตัวอย่างเช่นการทำแบ่ง Customer Database Segmentation นั้นนักวิเคราะห์ทางธุรกิจและนักวิเคราะห์ข้อมูลพยายามที่จะแบ่งกลุ่มของลูกค้าที่มีลักษณะร่วมกันให้อยู่ในกลุ่มเดียวกันเพื่อการวิเคราะห์ แต่ถ้าปรากฏว่าในกลุ่มตัวอย่างกลุ่มหนึ่งมีจำนวนมากเกินไป ก็อาจจะต้องมาปรับตัวแปรบางอย่างเพื่อให้เห็นกลุ่มที่ชัดเจน

2.6) Assimilation of Knowledge

เมื่อเราตั้งจุดประสงค์ในตอนต้นของขั้นตอน ตอนนี้ก็จะเป็นขั้นตอนสุดท้ายจุดประสงค์ถูกทำให้บรรลุ ในขั้นตอนนี้มี 2 เรื่องหลักที่จะต้องพิจารณา การนำเสนอสิ่งที่ค้นพบในแนวทางที่ทำให้เชื่อว่าน่าทำและในแนวทางที่เน้นทางธุรกิจเป็นหลัก และ ค้นหาแนวทางที่จะทำการวิเคราะห์ข้อมูลที่เข้ามาใหม่ ๆ ในขั้นตอนนี้จะเป็นหน้าที่ของนักวิเคราะห์ทางธุรกิจผู้ซึ่งมีประสบการณ์ก็จะสามารถสร้างสูตรสำเร็จต่อสิ่งที่ค้นพบและผลลัพธ์ที่หามาได้ที่จะนำไปเกี่ยวข้องโดยตรงกับจุด

ประสงค์ทางธุรกิจ แต่อย่างไรก็ตามและนอกจากนี้ยังอาจจะเกิดแนวคิดทางธุรกิจใหม่ ๆ ขึ้นมาเพื่อเป็นแนวทางในการดำเนินธุรกิจต่อไปได้

3.7 การใช้งานทางด้าน Data mining (Data Mining Application)

Data Mining ถูกนำมาใช้ในงานทางธุรกิจ ได้แก่ งานทางด้านการตลาดและงานขาย, บริการทางการเงิน, การทำนายแนวโน้มทางธุรกิจ, การค้าหุ้น, การควบคุมคุณภาพ, งานธนาคาร และงานบริหารบุคคล, การดักจับการลักลอบใช้โทรศัพท์มือถือ, การโกงค่าประกันภัย และอื่น ๆ

งานที่ไม่เกี่ยวข้องกับทางด้านธุรกิจ ได้แก่ อุตสาหกรรมปิโตรเคมี, งานวิทยาศาสตร์, การป้องกันไฟฟ้า, การกำหนดโครงสร้างทางเคมี, การป้องกันอาชญากรรมและการวิเคราะห์ทางแพทย์ ซึ่งในที่นี้จะขอกกล่าวในส่วนของการขอข้อมูลที่เกี่ยวข้องกับงานทางด้านการตลาดเท่านั้น

การจัดการทางด้านการตลาดเป็นส่วนหนึ่งที่ Data Mining ถูกใช้มากที่สุด บางครั้งเราเรียกว่า “Database Marketing” จุดประสงค์ของการใช้ก็เพื่อต้องการทราบกลุ่มเป้าหมายลูกค้าที่แท้จริง และสามารถสร้างโปรแกรมส่งเสริมการขาย และ โปรแกรมทางการตลาดให้มีประสิทธิภาพมากที่สุด โดยผ่านการใช้วิเคราะห์ฐานข้อมูลขององค์กรอย่างละเอียด องค์กรสามารถรวบรวมข้อมูลได้จาก รายการยอดซื้อจากเครดิตการ์ด, ใบรับรองผลิตภัณฑ์, คู่มือส่วนลด, การจับรางวัล, Complaint จากลูกค้า, การทำสำรวจชีวิตความเป็นอยู่และรูปแบบการใช้ชีวิตของกลุ่มลูกค้า เพื่อนำข้อมูลดังกล่าวไปผสมกับข้อมูลที่ทางองค์กรศึกษาหรือมีอยู่โดยทั่วไปแล้ว

วิธีการทำงานของ Data Mining จะมองผ่านเข้าไปในข้อมูลอย่างละเอียดเพื่อมองหา Model (รูปแบบ) และพฤติกรรมของผู้บริโภคในกลุ่มที่มีลักษณะคล้ายกัน ตัวอย่างเช่น ความสนใจ, ระดับรายได้, พฤติกรรมการใช้จ่ายเงิน ทำให้องค์กรสามารถมุ่งไปสู่กลุ่มเป้าหมายได้โดยตรง โดยที่ไม่ต้องใช้ค่าใช้จ่ายมากเกิดความจำเป็นไปในค่าโฆษณาและประชาสัมพันธ์ และยังเป็นการเพิ่มอัตราการตอบรับต่อ โปรแกรมส่งเสริมการขายที่ทำได้ด้วย

รูปแบบงานของ Data Mining ที่ถูกใช้คือการกำหนดรูปแบบการซื้อสินค้าของลูกค้า ณ ระยะเวลาต่าง ๆ นักการตลาดสามารถกำหนดได้ว่าพฤติกรรมค่าใช้จ่ายเงิน และลำดับการใช้จ่ายเงินเปลี่ยนไปอย่างไรเมื่อมีสมาชิกครอบครัวเพิ่มขึ้น ในกรณีของธนาคารนั้น การที่ลูกค้าเปลี่ยนบัญชีคนเคียวเป็นบัญชีร่วมย่อมหมายถึง โอกาสในอนาคตที่จะเสนอบริการเช่าซื้อ, ประกันชีวิต, กู้ซื้อบ้าน ได้ด้วย โดยการเข้าใจพฤติกรรมดังกล่าวจะทำให้ นักการตลาดสามารถทำ “Just In Time” คือสามารถตอบสนองความต้องการของกลุ่มลูกค้าแต่ละกลุ่ม ได้ตรงเวลาและถูกต้องที่สุด ต่อไปในอนาคตการเน้นศึกษาพฤติกรรมการซื้อของลูกค้าตลอดระยะเวลา ทำให้กลยุทธ์ขององค์กรเริ่มเปลี่ยน ไปแทนที่จะเน้นไปที่ Market Share จะเปลี่ยน ไปเป็น Customer Share

This material is reserved for educational use only, not allowed for commercial use.

งานการตลาดและการขายที่เกี่ยวข้อง ตัวอย่างเช่น Direct Marketing คือการเข้าถึงกลุ่มเป้าหมายให้มากที่สุด, Point of Sales การบริการ ณ จุดขาย เพื่อระบุรูปแบบการซื้อของลูกค้า รวมไปถึงการระบุกลุ่มเป้าหมายที่จะซื้อสินค้าในอนาคต การกำหนดแนวทางในการจัดโปรแกรมโฆษณา และการแบ่งกลุ่มของลูกค้าออกเป็นกลุ่มย่อย (Segment) ทำให้องค์กรเข้าถึง ค่านิยม, ความคาดหวัง, อุปนิสัย, การใช้ชีวิต ของผู้บริโภค ได้ใกล้เคียงมากขึ้น อาจอ้างได้ถึง One-To-One Marketing

การทำ Cross Selling เป็นอีกรูปแบบงานหนึ่งซึ่ง Data Mining ถูกใช้ การทำ Cross Selling คือเมื่อผู้ให้บริการหรือผู้ขายสินค้าพยายามที่จะสร้างความประทับใจให้กับลูกค้าผู้ซึ่งซื้อสินค้าชนิดหนึ่งเพื่อที่จะซื้อสินค้าที่เกี่ยวข้อง

Market Basket Analysis ซึ่งกำหนดว่าสินค้าชนิดไหนมีแนวโน้มที่จะถูกซื้อด้วยกันในการซื้อสินค้าครั้งหนึ่ง ๆ เพื่อที่ว่านักการตลาดจะได้จัดการทุกสิ่งได้ตั้งแต่การจัดเรียงสินค้าในร้าน การออกแบบโปรแกรมส่งเสริมการขาย ตัวอย่างเช่นการ ไปซื้อสินค้าในแผนกซูเปอร์มาร์เก็ตของห้างสรรพสินค้า “กำหนดว่าลูกค้ากลุ่มใดเป็นกลุ่มที่สร้างผลกำไรให้กับองค์กรมากที่สุด”

การสร้างความสัมพันธ์กับลูกค้าโดยผ่านการจัดกลุ่มของลูกค้าที่มีความพอใจและลูกค้าที่ไม่พอใจต่อสินค้าหรือบริการขององค์กร มันอาจจะจัดกลุ่มของลูกค้าที่กำลังจะหันไปใช้สินค้าของกลุ่ม

ในบทต่อไปจะอธิบายถึงเทคนิคของ Data Mining โดยนำวิธีการทำงานของ Decision Tree มาใช้ในการดำเนินระบบขององค์กร โดยสรุปผลของขั้นตอนในการทำงานและผลลัพธ์ที่ออกมา จะพบว่าจากการวิเคราะห์การใช้เทคนิคของ Data Mining จะช่วยให้มีโอกาสสำเร็จมากขึ้นในการจัดการกับปัญหาทางธุรกิจอย่างไร

บทที่ 4

Decision Tree และขั้นตอนการดำเนินงานในระบบ

4.1 ขั้นตอนการศึกษาและตัวแปรที่เกี่ยวข้อง

ต่อไปนี้อีกร่างจะศึกษาถึงปัจจัยที่มีผลต่อพฤติกรรมการซื้อสินค้าในโปรแกรมส่งเสริมการขายที่ได้นำเสนอสู่ลูกค้า โดยมีการกำหนดเงื่อนไขของโปรแกรมน่า "ซื้อสินค้ากระดาษทรายกลมทุก ๆ 20 ม้วน จะได้รับฟรีเครื่องขัดกระดาษทรายจำนวน 1 เครื่อง" ข้อมูลของลูกค้าในฐานะข้อมูลจะถูกรวบรวมและเลือกออกมาเฉพาะ Field ที่องค์กรให้ความสนใจในการศึกษาเท่านั้น ดังตัวอย่างของตัวแปรในตารางที่ 4.1

COLUMN NAME	TYPE OF DATA	VALUE	DESCRIPTION
CUSTOMER NAME	CHARACTER	SPECIFIC NAME	CUSTOMER NAME
BUYING INDICATOR	CHARACTER	BUY, NOT BUY	PROMOTION PROGRAM BOUGHT
OWNER AGE	NUMERIC	20-30, 31-40, 41-50, 51-60	AGE OF CUSTOMER
INDUSTRY TYPE	CHARACTER	WOOD, METAL, AUTO	TYPE OF CUSTOMERS
MARKET	CHARACTER	DOMESTIC, EXPORT, DOMESTIC/EXPORT	TARGET MARKET
SALES AMOUNT	NUMERIC	<10,000 10,001-30,000 30,001-50,000 50,001-70,000 70,001-90,000 >90,000	AMOUNT OF SALES PER MONTH
NO OF EMPLOYEE	NUMERIC	1-50, 51-100, 101-150, 151 UP	NO OF EMPLOYEE
YEARS WITH COMPANY	NUMERIC	1-5, 6-10, 11-15, 16-20	NO OF YRS DEALING WITH COMPANY
LOCATION	CHARACTER	CENTRAL, NORTHERN, SOUTHERN, NORTH EASTERN, EAST	CUSTOMER LOCATION
RELATED PRODUCTS	CHARACTER	BELT, ROLL, DISC SHEET, BELT/SHEET, BELT/ROLL, BELT/DISC, BELT/ROLL/DISC, BELT/ROLL/SHEET, BELT/ROLL/SHEET/DISC, ROLL/DISC, DISC/SHEET, SHEET/ROLL, SHEET/DISC, SHEET/ROLL/DISC	CURRENT FORM OF PRODUCT BOUGHT BY CUSTOMER

ตารางที่ 4.1 ตารางการวิเคราะห์แนวโน้มการซื้อสินค้าในรายการส่งเสริมการขาย

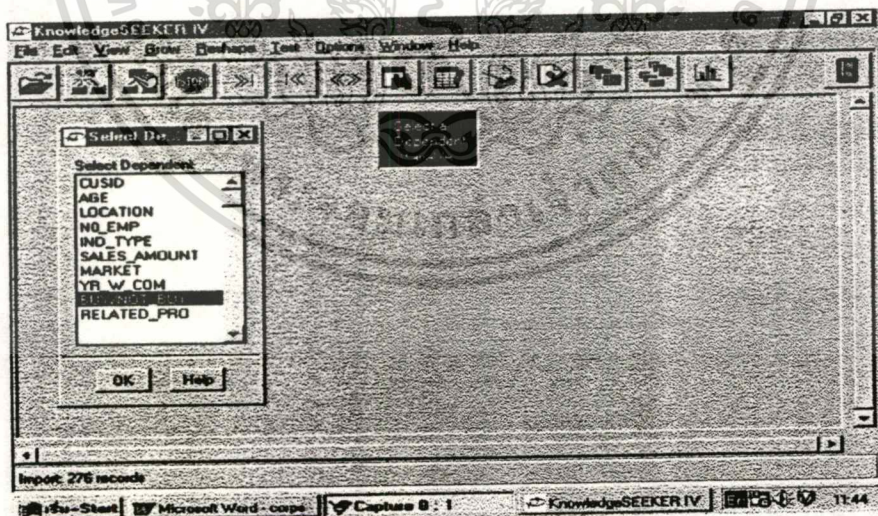
ที่มา: บริษัท 3 เอ็ม (ประเทศไทย) จำกัด

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

จากตัวแปรที่เราสนใจหรือต้องการจะศึกษาในสถานการณ์นี้คือ “การตัดสินใจซื้อหรือไม่ซื้อสินค้าในโปรแกรมส่งเสริมการขาย” ของลูกค้า จากการเก็บข้อมูลทราบว่าโปรแกรมที่องค์กรนำเสนอให้ลูกค้าจำนวนประชากรทั้งหมด 276 ลูกค้า นั้น มีการตอบการสั่งซื้อกลับมาเพียง 12.7% เท่านั้น ซึ่งลูกค้าจำนวนทั้งหมด 276 รายนั้น เป็นลูกค้าปัจจุบันขององค์กรโดยมีลูกค้าในอุตสาหกรรมไม้, อุตสาหกรรมโลหะ และอุตสาหกรรมรถยนต์ ทางองค์กรจึงต้องการทราบเป็นอย่างยิ่งว่าอุตสาหกรรมใดที่องค์กรควรมุ่งเน้น และอะไรคือปัจจัยที่ทำให้ลูกค้ากลุ่มนั้นมีความสนใจในโปรแกรม เพื่อที่จะได้ขยายผลของโปรแกรมออกไปสู่ลูกค้าที่องค์กรไม่เคยติดต่อก่อนหรือไม่ใช่ลูกค้าขององค์กร หรือออกแบบโปรแกรมให้เหมาะสมกับความต้องการของลูกค้าในกลุ่มดังกล่าวให้มากที่สุด ในที่นี้องค์กรจะนำเอาเทคนิคของการทำ Data Mining มาใช้ โดยเทคนิคดังกล่าวเป็นเทคนิคแบบ “Decision Tree”

จากการเก็บรวบรวมข้อมูลศึกษาเราจะอาศัยปัจจัยต่าง ๆ ที่มีผลกระทบต่อพฤติกรรมการซื้อสินค้าในรายการ โปรแกรมส่งเสริมการขาย มาอธิบายถึงสาเหตุว่าทำไมลูกค้ารายหนึ่งถึงซื้อสินค้าระหว่างที่ลูกค้าอีกรายหนึ่งไม่มีความสนใจในโปรแกรมขายสินค้าชนิดนี้เลย เราพอที่จะอธิบายตัวแปรต่าง ๆ ที่มีผลต่อการสั่งซื้อได้ โดยเราจะหาความสัมพันธ์ของตัวแปรอื่น ๆ ที่มีต่อตัวแปรที่เราสนใจ โดยเราจำเป็นต้องกำหนดตัวแปรที่เราสนใจขึ้นมาก่อน ดังในรูปที่ 4.1 ที่เรากำหนด Buy/Not_Buy ขึ้นมา



รูปที่ 4.1 การเลือก Dependent Variable สำหรับการศึกษ

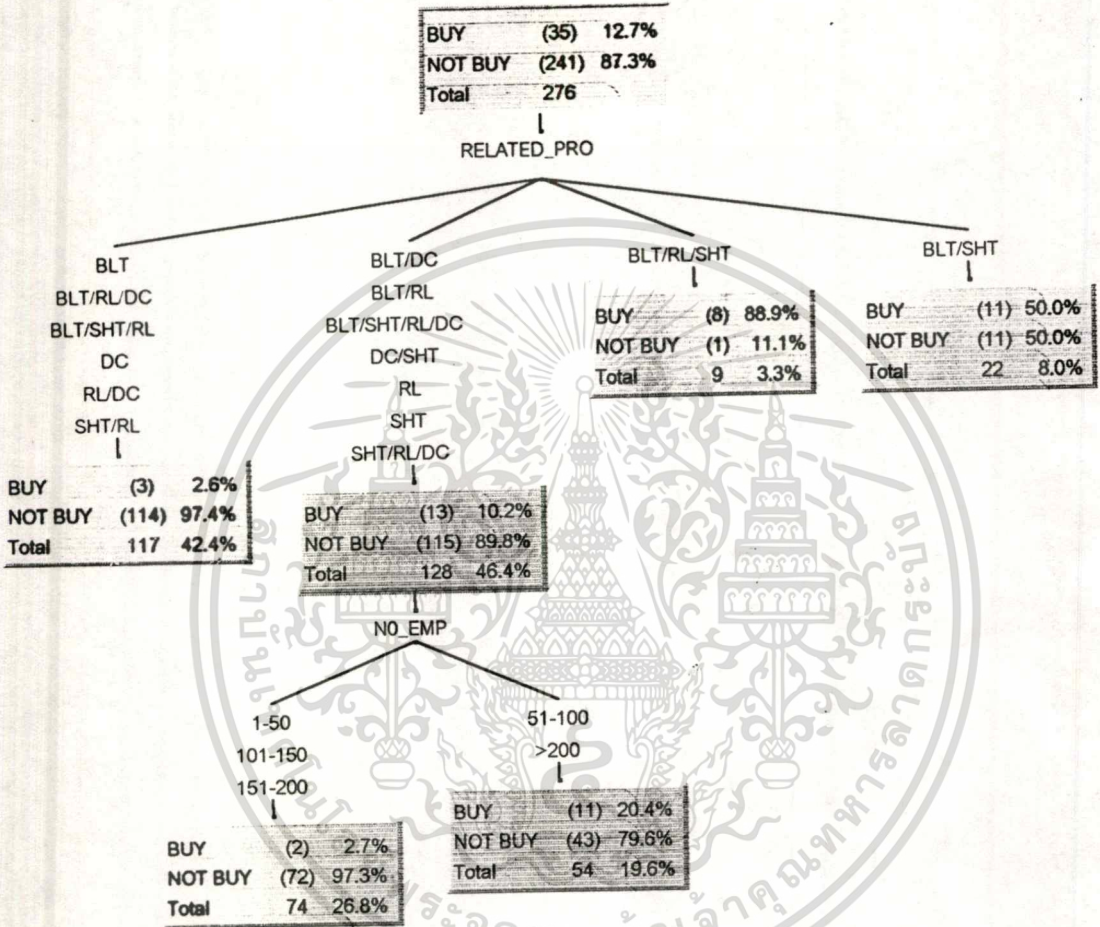
ที่มา: โปรแกรม Knowledge Seeker

จากรูปที่ 4.2 เป็นการสร้าง Decision Tree โดยอาศัย Algorithm การทำงานแบบ CART เป็นตัวเลือกตัวแปรที่ให้ผลในการทำงานมากที่สุดในการสร้างความสัมพันธ์ของข้อมูลกับตัวแปรที่

This material is reserved for educational use only, not allowed for commercial use.

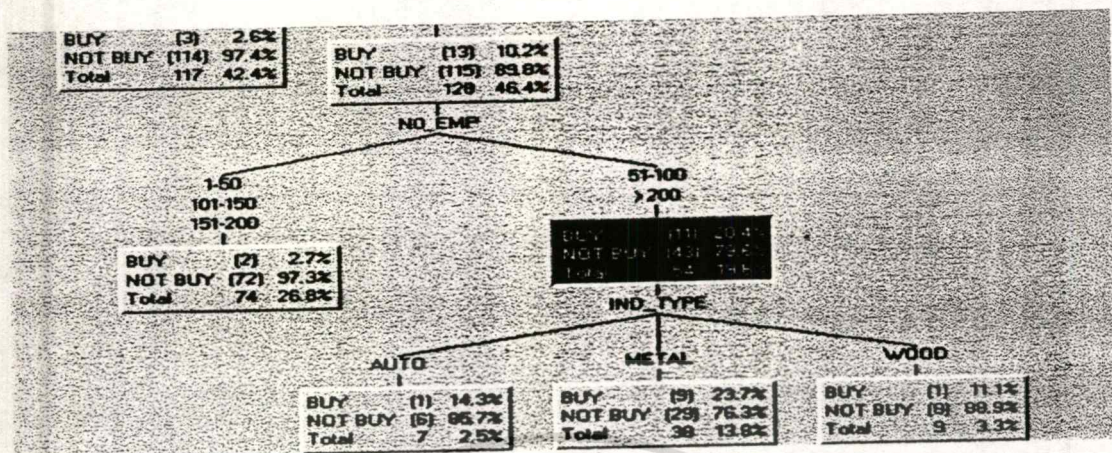
Forbidden to modify the content, and cite the document when use.

องค์กรต้องการจะศึกษา จากรูปกำหนดว่าอุตสาหกรรมที่ซื้อสินค้าในหมวด BLT/DC, BLT/RL, BLT/RL/DC/SHT, DC/SHT, RL, SHT และ SHT/RL/DC ที่มีจำนวนคนงานขนาดกลาง 51-100 คน และขนาดใหญ่ 200 คนขึ้นไป มีเปอร์เซ็นต์การซื้อสินค้าในรายการส่งเสริมการขาย 20.4%



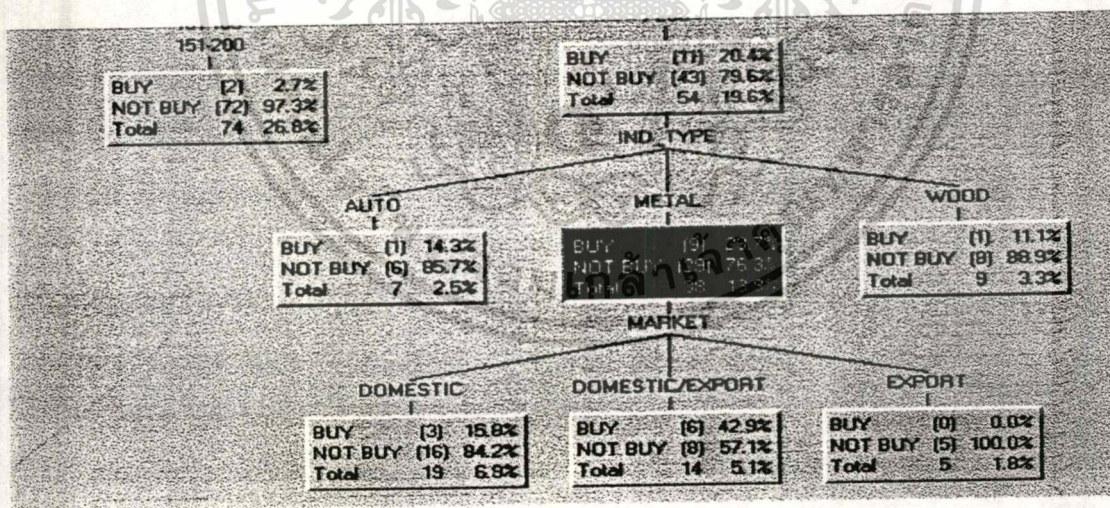
รูปที่ 4.2 รูปการทำงานของ Decision Tree โดยให้ Algorithm สร้าง Tree ขึ้นมาเอง
ที่มา: โปรแกรม Knowledge Seeker

จากรูป 4.3 เป็นการศึกษาต่อจากที่ Algorithm ได้ให้โครงร่างต้นไม้ตอนต้น เพื่อหากลุ่ม ชนิดของอุตสาหกรรมที่องค์กรควรให้ความสนใจ โดยจากรูปจะเป็นการเลือกตัวแปรเข้ามาใส่โดย องค์กรเอง มิได้คิดคำนวณโดยเครื่อง จะเห็นว่ากลุ่มตลาดอุตสาหกรรมโลหะมีเปอร์เซ็นต์การซื้อ สูงถึง 23.7% ในขณะที่ความเชื่อเดิมขององค์กรคิดว่าตลาดไม้เป็นตลาดที่สามารถสร้างรายได้จาก การขายมากกว่าตลาดอื่น ๆ เพราะโดยลักษณะของธุรกิจแล้วตลาดไม้เป็นอุตสาหกรรมที่ใช้สินค้า ประเภทนี้มากกว่าอุตสาหกรรมประเภทอื่น แต่จากการทำ Data Mining พบว่ามีได้เป็นดังนั้น เพราะเปอร์เซ็นต์การตอบรับของตลาดไม้ต่ำกว่าตลาดอุตสาหกรรมโลหะ



รูปที่ 4.3 รูปการทำงานของ Decision Tree ที่ศึกษาต่อถึงกลุ่มของชนิดของอุตสาหกรรม
ที่มา: โปรแกรม Knowledge Seeker

จากรูปที่ 4.4 จะพบว่าลูกค้าที่อยู่ในอุตสาหกรรมงานโลหะที่ซื้อสินค้าประเภท BLT/DC, BLT/RL, BLT/RL/DC/SHT, DC/SHT, RL, SHT และ SHT/RL/DC ที่มีจำนวนคนขนาดปานกลางถึงมาก และมีตลาดรองรับทั้งในและต่างประเทศจะเป็นกลุ่มที่ให้ความยอมรับต่อโปรแกรมส่งเสริมการขายสูงมาก เมื่อเทียบกับตลาดอื่น ๆ



รูปที่ 4.4 รูปการทำงานของ Decision Tree ที่ศึกษาถึงกลุ่มตลาดรองรับของอุตสาหกรรม
ที่มา: โปรแกรม Knowledge Seeker

4.2 สรุปผลการศึกษา

จากการทำงานในขั้นตอนที่ผ่านมา ทำให้องค์กรสามารถกำหนดกลุ่มเป้าหมายขององค์กรได้ว่า กลุ่มใดคือกลุ่มเป้าหมายหลักขององค์กรที่มีโอกาสจะเป็นลูกค้าในอนาคตได้ โดยสามารถกำหนดลักษณะเฉพาะของกลุ่มลูกค้าที่มีโอกาสซื้อในอนาคตว่ามีลักษณะที่ร่วมกันอยู่อย่างไร และแบ่งลักษณะเฉพาะออกเป็นอย่างไร เพื่อใช้ในการวิเคราะห์อย่างถูกต้องเมื่อได้ข้อมูลดังกล่าวแล้ว องค์กรก็จะสามารถขยายผลของกิจกรรมทางการตลาด โดยนำเสนอโปรแกรมส่งเสริมการขายไปสู่กลุ่มลูกค้าที่ไม่เคยได้รับโปรแกรมมาก่อน ซึ่งอาจจะเป็นทั้งลูกค้าขององค์กรหรือลูกค้าที่ไม่ใช่ขององค์กรก็ได้

จากการศึกษาทางองค์กรเลือกเอาเทคนิค Data Mining เข้ามาใช้เพื่อกำหนดลักษณะของกลุ่มโดยใช้ซอฟต์แวร์ Knowledge Seeker ของบริษัท Angoss ซึ่งซอฟต์แวร์ดังกล่าวมีการทำงาน CART สาเหตุที่เลือกใช้เทคนิค Data Mining เพราะมีความสามารถในการอธิบายถึงการจัดกลุ่ม (Classification) ได้เป็นอย่างดี ง่ายต่อการทำความเข้าใจ การจัดกลุ่มเป็นขั้นตอนของการใช้ข้อมูลในอดีตมาสร้างรูปแบบเพื่อการทำนายผล หลังจากที่ได้เลือก Model แล้ว ก็ได้มีการเลือกตัวแปรที่สำคัญ ๆ ที่จะมีผลต่อการทำนายออกมาจากฐานข้อมูลขององค์กร นอกจากนี้ยังมีการรวบรวมข้อมูลเพิ่มเติมบางประเภทที่ไม่สามารถหามาได้ ข้อมูลของตัวแปรที่ได้มานั้นจะต้องถูกจัดให้อยู่ในรูปแบบที่เครื่องมือ Data Mining สามารถวิเคราะห์ได้ หลังจากนั้นก็ได้เลือกตัวแปรที่ต้องการจะศึกษาคือ “การซื้อและไม่ซื้อสินค้าในรายการส่งเสริมการขาย” เพื่อมาทำการวิเคราะห์เพิ่มเติมออกไปโดยนำตัวแปรตัวอื่น ๆ เข้ามาเกี่ยวข้องสร้างความสัมพันธ์ และดูผลลัพธ์ที่ซ่อนอยู่ในข้อมูลนั้น ๆ จากนั้นก็กำหนดเป็นข้อสรุปเพื่อนำเสนอต่อผู้บริหารต่อไป

ลักษณะของกลุ่มเป้าหมายที่กำหนดออกมาคือ ลูกค้าที่มีโอกาสใช้สินค้าได้แก่ลูกค้าในอุตสาหกรรมที่มีคนงานจำนวน 51-100 และ 200 คนขึ้นไป โดยซื้อสินค้าประเภทต่อไปนี้ BLT/DC, BLT/RL, BLT/RL/DC/SHT, DC/SHT, RL, SHT และ SHT/RL/DC นอกจากนี้ยังสามารถกำหนด Profile ของลูกค้าได้ว่า

ประเภท: อุตสาหกรรมโรงงานโลหะ

จำนวนคนงาน: 51-100 คน และ 200 คนขึ้นไป

ตลาดรองรับสินค้า: ในประเทศและต่างประเทศ

บทที่ 5

บทสรุป

จากการศึกษามาผ่านมาพบว่านักการตลาดหรือองค์กรใดที่ต้องการใช้เทคนิคของ Data Mining เข้ามาช่วยในการทำงานนั้นจำเป็นอย่างยิ่งที่จะต้องเข้าใจลักษณะที่แท้จริงของเนื้องานเสียก่อนเพื่อที่จะได้มีความสามารถในการกำหนดปัญหาและขอบเขตปัญหาได้ ถ้าการกำหนดปัญหาถูกต้องก็จะนำไปสู่ขั้นตอนการทำ Data Mining ที่ถูกต้องและนำไปสู่ผลลัพธ์ที่ต้องการได้จากปัญหาขององค์กรก็คือ มีความต้องการหากกลุ่มลูกค้าที่แท้จริงขององค์กรที่มีความสนใจต่อการซื้อสินค้าในรายการส่งเสริมการขาย หลังจากที่ได้กำหนดปัญหาและขอบเขตของปัญหาเรียบร้อยแล้วก็จะมาสู่ขั้นตอนการรวบรวมข้อมูล ปัญหาขององค์กรจะเป็นตัวกำหนดว่าต้องการข้อมูลประเภทใด ซึ่งจากปัญหาดังกล่าวพบว่าไม่สามารถที่จะใช้ข้อมูลจากภายในองค์กรเพียงอย่างเดียวได้ จะต้องอาศัยข้อมูลจากการรวบรวมมาจากภายนอกด้วย ข้อมูลดังกล่าวต้องถูกแปลงให้อยู่ในรูปแบบที่เทคนิค Data Mining สามารถรับเข้าไป Process ได้ ตัวปัญหาเองนั้นจะกำหนดว่าควรจะใช้เทคนิคการทำ Data Mining แบบใด ผลลัพธ์ที่ได้จากการทำ Mining จะเป็นผลลัพธ์ที่ออกมาจากการสร้างของ Algorithm เอง ซึ่งผลที่ออกมาอาจจะแตกต่างจากแนวคิดที่องค์กรได้คาดไว้ตั้งแต่ต้น

ในขั้นตอนของการปฏิบัติงานจริงองค์กรใช้ Algorithm ของ Decision Trees แบบ Modified CART (Modified Classification and Regression Tree) เป็นเครื่องมือในการแยกแยะกลุ่มและหาความสัมพันธ์ของข้อมูลและหาลักษณะเฉพาะของข้อมูลลูกค้าที่ซื้อสินค้าและไม่ซื้อสินค้าในรายการส่งเสริมการขาย จากกรรมวิธีการทำงานของ Decision Trees นั้น โปรแกรมได้ดึงเอาตัวแปรการซื้อสินค้าที่เกี่ยวข้องและตัวแปรจำนวนคนงาน เพื่อมาหาผลสรุปและนำผลดังกล่าวไปใช้ในประโยชน์ทางธุรกิจ จากการทำงานนั้นจะเห็นได้ว่ามีตัวแปรเพียง 2 ตัวเท่านั้นที่ Algorithm เลือกมาวิเคราะห์ ซึ่งผู้ใช้งานจะพบว่าตัวแปรที่ถูกดึงมานั้นอาจจะไม่ใช่ตัวแปรที่ผู้ใช้งานคาดไว้ในการศึกษางานในลักษณะอื่น ๆ ซึ่งมีข้อมูลจำนวน Record และตัวแปรมาก ๆ นั้น จะพบว่า Algorithm ของ Decision Tree อาจจะดึงเอาตัวแปรจำนวนมากออกมา ทำให้ผู้ใช้งานจะต้องใช้วิธีการอนุมานและประสิทธิภาพในงานธุรกิจนั้น ๆ ในการวิเคราะห์ว่าตัวแปรที่ Algorithm เลือกมานั้นมีความเหมาะสมและความเกี่ยวข้องกับตัวแปรต่าง ๆ มากเพียงใด มีความเกี่ยวข้องกับตัวแปรที่ต้องการจะศึกษามากเพียงใด ถ้าหากพบว่าตัวแปรที่ Algorithm นั้นเลือกมาไม่เหมาะสมแล้ว ก็จะต้องใช้ประสิทธิภาพของผู้ใช้งานในการเลือกตัดตัวแปรบางตัวออกจากกลุ่มต้นไม้ หรือเรียกว่า การ Pruning เพื่อหาผลสรุปจากกรรมวิธีการทำงานของ Decision Tree และเพื่อเพิ่มประสิทธิภาพ

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

ของผลในการทำนายให้มีความถูกต้องมากขึ้น แต่ในการศึกษานี้ไม่ได้มีการทำ Pruning เพราะตัวแปรที่ Algorithm เลือกออกมานั้นมีความสมเหตุสมผลและเกี่ยวข้องกับตัวแปรที่องค์กรต้องการจะศึกษาอยู่แล้ว จากนั้นทางองค์กรก็ได้เลือกเอาตัวแปรอื่น ๆ ที่คาดว่าจะเกี่ยวข้องมาสร้างความสัมพันธ์เอง จากการกระทำดังกล่าวเป็นการกระทำที่นอกเหนือจากการทำงานที่แท้จริงของ Data Mining เช่น การนำเอาประเภทของอุตสาหกรรมมาศึกษาต่อเพิ่มเติมจากตัวแปรจำนวนของพนักงานเพื่อกำหนดรายละเอียดของกลุ่มลูกค้าที่ต้องการให้ชัดเจนยิ่งขึ้น

ผลที่ได้จากการศึกษาพบว่าแนวคิดเดิมที่องค์กรคิดว่าโรงงานอุตสาหกรรมผลิตไม้และเฟอร์นิเจอร์ไม้้นั้นที่โดยปกติใช้สินค้ากระดาษทรายเป็นวัตถุดิบหลักในการผลิตนั้น จะตอบสนองต่อโปรแกรมส่งเสริมการขายมากที่สุด ซึ่งขัดแย้งกับผลลัพธ์ที่ได้จากกรรมวิธีการทำงานของ Decision Tree ซึ่งพบว่าโรงงานอุตสาหกรรมที่จะตอบสนองต่อโปรแกรมดีที่สุดคือโรงงานอุตสาหกรรมโลหะ จึงนำมาสู่ข้อสรุปว่าองค์กรควรเปลี่ยนความสนใจจากตลาดอุตสาหกรรมไม้มาสู่อุตสาหกรรมโลหะ

แต่ผลลัพธ์ที่ได้ออกมานั้นคงยังจะไม่อาจนำมายืนยันถึงความสำเร็จได้ 100 เปอร์เซ็นต์ในการทำโปรแกรมส่งเสริมการขาย แต่ก็ยังเป็นเครื่องยืนยันอย่างหนึ่งถึงโอกาสของความสำเร็จที่จะเกิดขึ้น แต่ขั้นตอนในการทำ Data Mining นั้นเป็นขั้นตอนของการทำงานย้อนไปและย้อนมาเพื่อตรวจสอบความถูกต้อง ดังนั้นผลลัพธ์ที่ได้จากการทำ Data Mining ครั้งนี้อาจจะถูกนำไปใช้และนำผลลัพธ์ของมันมาเป็น Input ของการทำ Data Mining ครั้งต่อไปเพื่อหาผลสรุป

ผลสำเร็จอีกอย่างหนึ่งของโครงการนี้อาจจะขึ้นอยู่กับจำนวนของข้อมูลที่ใช้ในการศึกษาค้น เนื่องจากข้อจำกัดของจำนวนข้อมูลที่มีในโครงการศึกษานี้คือมี Record เพียง 276 Record และมีตัวแปรที่เกี่ยวข้องกับการศึกษาเพียง 10 ตัวแปรเท่านั้น เนื่องจากจำนวนข้อมูลที่น้อยทำให้อาจเกิดผลลัพธ์ที่ออกมาคลาดเคลื่อนได้

บรรณานุกรม

Berry, M. J.A. and Linoff, G., Data Mining Techniques: for marketing, sales, and customer support, Wiley Computer Publishing, 1997.

Cabena, P., et.al., Discovering Data Mining: from concept to implementation, Prentice Hall Publishing, 1998.

Groth, R., Data Mining: a hands on approach for business professionals, Prentice Hall Publishing, 1997.

“Knowledge Seeker.” (ออนไลน์). เข้าถึงได้จาก: <http://www.angoss.com/>, 1998.



ภาคผนวก

ฐานข้อมูลที่ใช้ในการวิเคราะห์ทำ Data Mining เพื่อหาปัจจัยที่มีผลต่อการทำโปรแกรมส่งเสริมการขายของสินค้ากระดาษทรายและเครื่องขัดกระดาษทรายนั้น ได้รับข้อมูลมาจากฐานข้อมูลของบริษัท 3 เอ็ม ประเทศไทย จำกัด ส่วนหนึ่ง และเกิดจากการรวบรวมข้อมูลเพิ่มเติมจากพนักงานขายของบริษัทฯ โดยเกิดจากการสัมภาษณ์และรวบรวมข้อมูลโดยอาศัยข้อมูลเดิมมาประกอบกันดังตารางต่อไปนี้



CUSID	CUS_NAME	AGE	LOCATION	NO_EMP	IND_TYPE	SALES_AMOUNT	MARKET	YR_W_COM	BUY/NOT_BUY	RELATED_PRO
1001	ALLOY INDUSTRIAL	41-50	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
1002	ANAN CROMRUM	31-40	CENTRAL	51-100	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	RL
1003	BANCHA FACTORY	31-40	CENTRAL	51-100	METAL	<10,000	DOMESTIC	6-10	NOT BUY	BLT/SHT
1004	BANGKOK STEEL IND	41-50	CENTRAL	151-200	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	SHT
1005	BANJONG INDUSTRIAL	31-40	EASTERN	51-100	METAL	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	SHT/RL
1006	BAOME CHEMICAL	31-40	CENTRAL	101-150	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/SHT/RL/DC
1007	BKK ART & CRAFT	31-40	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	BUY	BLT/SHT
1008	BKK COIL CENTER	20-30	EASTERN	51-100	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT/RL
1009	BKK FURNITECH	31-40	CENTRAL	>200	METAL	10,001-30,000	DOMESTIC/EXPORT	6-10	BUY	BLT/DC
1010	BKK NARUPHAN	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	DC/SHT
1011	BKK PRESS PART	20-30	CENTRAL	>200	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	RL/DC
1012	BKK SPRING	31-40	CENTRAL	51-100	METAL	<10,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/SHT
1013	CHAI CHAROEN GRINDING WHEEL	31-40	EASTERN	1-50	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT/RL/DC
1014	CHAI HIRANCHOT	20-30	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/RL/DC
1015	CHAI PATANASIN	31-40	NORTH EASTERN	1-50	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/RL
1016	CHALEE LOHAKARN	31-40	CENTRAL	1-50	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	BLT/RL/DC
1017	CHAROENMIT WATER TANK	41-50	EASTERN	1-50	METAL	30,001-50,000	DOMESTIC	11-15	NOT BUY	DC
1018	CHEVAL ELECTRONIC	31-40	EASTERN	51-100	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	RL
1019	CHROMALLOY	31-40	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	SHT
1020	CHU CHIN HUA	41-50	NORTH EASTERN	101-150	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
1021	CK METAL & TRADING	31-40	CENTRAL	1-50	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	BLT/SHT/RL/DC
1022	COLUMBIA CHROME	31-40	CENTRAL	51-100	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/SHT/RL/DC
1023	COMPRESOR THAI IND.	31-40	CENTRAL	51-100	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	BLT
1024	COSMO INDUSTRY	31-40	SOUTH	51-100	METAL	10,001-30,000	EXPORT	11-15	NOT BUY	BLT/DC
1025	DATA LOOP	31-40	NORTH EASTERN	1-50	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	DC
1026	DRESSEN AIRCRAFT	20-30	NORTH	1-50	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	RL
1027	FAH SAEANG	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	DC/SHT
1028	FUJITSU	31-40	CENTRAL	51-100	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	SHT/RL/DC
1029	FUTURE TECH	41-50	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT
1030	GRAND PEWTER	31-40	CENTRAL	151-200	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	RL
1031	GREEN SPOT	31-40	CENTRAL	101-150	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	DC/SHT
1032	HENGER MFG	31-40	CENTRAL	101-150	METAL	30,001-50,000	EXPORT	6-10	NOT BUY	BLT
1033	HONGKONG FURNISHING	41-50	CENTRAL	101-150	METAL	>90,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
1034	INTER-CUT	41-50	CENTRAL	101-150	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	SHT
1035	INTERTOOL TECHNOLOGY	31-40	EASTERN	1-50	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	DC/SHT
1036	KACHIT PEWTER	51-60	CENTRAL	1-50	METAL	10,001-30,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
1037	KHT CENTRAL SUPPLY	31-40	CENTRAL	1-50	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/SHT/RL/DC
1038	KINAREE	41-50	CENTRAL	151-200	METAL	30,001-50,000	EXPORT	1-5	NOT BUY	DC
1039	KLUAY NAM THAI ENGINEERING	31-40	CENTRAL	51-100	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	BLT/DC
1040	KLUAY NAM THAI KARNCHAENG	31-40	CENTRAL	1-50	METAL	30,001-50,000	DOMESTIC	6-10	NOT BUY	BLT/DC
1041	KYO DO DIE WORK	31-40	NORTH	1-50	METAL	30,001-50,000	EXPORT	6-10	NOT BUY	DC/SHT
1042	KYOKAM	31-40	CENTRAL	51-100	METAL	<10,000	DOMESTIC	6-10	NOT BUY	BLT
1043	LAO LIANG HUA	41-50	SOUTH	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1044	LEE RUNG RUEANG STEEL	31-40	CENTRAL	>200	METAL	10,001-30,000	DOMESTIC/EXPORT	6-10	BUY	BLT/DC
1045	LOHAKIT STEEL SERVICE CENTER	31-40	EASTERN	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	SHT/RL/DC
1046	LOHAPATEEP IND	20-30	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	11-15	NOT BUY	RL
1047	LOHASLP ENGINEERING	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT/DC
1048	LOLAND INDUSTRY	41-50	CENTRAL	151-200	METAL	10,001-30,000	EXPORT	6-10	NOT BUY	BLT
1049	LT WORK	31-40	EASTERN	1-50	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	SHT
1050	MAELO	41-50	CENTRAL	51-100	METAL	>90,000	DOMESTIC	6-10	NOT BUY	DC/SHT
1051	MAHAJAK IND	31-40	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	DC
1052	MARY JEWELRY	31-40	CENTRAL	151-200	METAL	30,001-50,000	EXPORT	1-5	NOT BUY	BLT/SHT/RL/DC
1053	MDA	31-40	EASTERN	1-50	METAL	10,001-30,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT/DC
1054	METAL	31-40	SOUTH	151-200	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT/SHT/RL/DC
1055	MIGHTY IND	31-40	CENTRAL	>200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	RL
1056	MINAMI & VERBINA	31-40	EASTERN	101-150	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	DC/SHT
1057	MODERNFORM GROUP	41-50	CENTRAL	>200	METAL	30,001-50,000	DOMESTIC/EXPORT	6-10	BUY	SHT/RL/DC
1058	MPN STAINLESS STEEL	51-60	EASTERN	101-150	METAL	<10,000	DOMESTIC	1-5	NOT BUY	BLT/SHT/RL/DC
1059	NASTOA	20-30	CENTRAL	101-150	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	SHT/RL/DC
1060	NATIONAL THAI	51-60	CENTRAL	>200	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
1061	NEWLIFE	41-50	SOUTH	1-50	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/DC
1062	NIKKEI THAI ALUMINIUM	31-40	CENTRAL	>200	METAL	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	SHT/RL/DC
1063	NOVOLUX WATCH	31-40	EASTERN	>200	METAL	30,001-50,000	EXPORT	1-5	NOT BUY	RL
1064	NSK IND	31-40	CENTRAL	101-150	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	SHT
1065	NSK STEEL	41-50	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	DC
1066	NSL INTERGROUP	31-40	CENTRAL	>200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/SHT/RL/DC

CUSID	CUS_NAME	AGE	LOCATION	NO_EMP	IND_TYPE	SALES_AMOUNT	MARKET	YR_W.COM	BUY/NOT_BUY	RELATED_PRO
1067	NUMSILP ELECTRIC	31-40	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/DC
1068	NVC TECHNOLOGY NYC IND	31-40	NORTH EASTERN	101-150	METAL	<10,000	DOMESTIC/EXPORT	6-10	NOT BUY	SHT/RL/DC
1069	ORIENTAL STAINLESS	31-40	CENTRAL	101-150	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	BLT
1070	PATKOL	31-40	CENTRAL	>200	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	DC/SHT
1071	PCS FOUNDRY	31-40	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1072	PERMA STEEL	31-40	EASTERN	151-200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	RL
1073	PETCH IND	20-30	EASTERN	1-50	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT
1074	PORNTHPEKTPAISAN	20-30	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1075	PR PROGRESS	31-40	CENTRAL	151-200	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	SHT/RL/DC
1076	PRASEKT ANODZE	41-50	NORTH EASTERN	1-50	METAL	30,001-50,000	EXPORT	1-5	NOT BUY	BLT
1077	PRATULEAK THAI	41-50	CENTRAL	1-50	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	SHT/RL/DC
1078	PRECISCO PROPACING	31-40	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT/DC
1079	PRECISION MFG	31-40	CENTRAL	1-50	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	DC
1080	PROGRESS WATCH	31-40	CENTRAL	51-100	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	SHT
1081	PTT ENGINEERING	31-40	EASTERN	51-100	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	RL
1082	PTT POMPATAN KARFAFAH	20-30	CENTRAL	51-100	METAL	>90,000	DOMESTIC/EXPORT	1-5	NOT BUY	RL/DC
1083	Q INDUSTRY	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1084	RIVER TECH	31-40	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/SHT
1085	ROYAL FLAME	41-50	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT
1086	SAITONG KARNCHANG	31-40	NORTH	151-200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/SHT
1087	SAMART MACHINERY	41-50	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	SHT/RL/DC
1088	SAMSON EYEGASS	31-40	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT/RL
1089	SERVICE ENGINEERING	31-40	EASTERN	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	RL
1090	SHIKEN CORPORATION	31-40	NORTH	101-150	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/DC
1091	SIAM CARGO CONTAINER	31-40	CENTRAL	101-150	METAL	>90,000	DOMESTIC	1-5	NOT BUY	DC
1092	SIAM CEMENT	31-40	CENTRAL	101-150	METAL	<10,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
1093	SIAM CITY BRASS WARE	31-40	CENTRAL	51-100	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
1094	SIAM COOLER IND	31-40	NORTH EASTERN	101-150	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	SHT/RL/DC
1095	SIAM GROUP DIE & PART	31-40	SOUTH	51-100	METAL	10,001-30,000	DOMESTIC	11-15	NOT BUY	SHT
1096	SIAM OKAMURA STEEL	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	BUY	SHT/RL/DC
1097	SIAM SANITARY FITTING	31-40	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/SHT/RL/DC
1098	SIAM STAINLESS PIPE FITTING	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	RL
1099	SIAM STAINLESS STEEL	31-40	EASTERN	>200	METAL	30,001-50,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT
1100	SIAM STEEL SERVICE	41-50	EASTERN	>200	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	DC
1101	SIAM WILAT IRON SHEET	31-40	CENTRAL	101-150	METAL	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	DC
1102	SINCO TECHNOLOGY	31-40	CENTRAL	51-100	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	BLT
1103	SINDUMRONG	41-50	NORTH	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	DC
1104	SODICK	31-40	EASTERN	101-150	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/DC
1105	SOECON ENGINEERING	31-40	CENTRAL	101-150	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	BLT
1106	SOMBOON LOHAKIT	51-60	EASTERN	101-150	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	BLT
1107	SRICHAROEN INDUSTRY	31-40	CENTRAL	>200	METAL	30,001-50,000	DOMESTIC/EXPORT	6-10	BUY	RL
1108	SRIBONCHOT	41-50	CENTRAL	101-150	METAL	<10,000	DOMESTIC	6-10	NOT BUY	BLT
1109	SRSURASAK STAINLESS	31-40	CENTRAL	51-100	METAL	30,001-50,000	DOMESTIC	11-15	BUY	SHT
1110	SRIJITAI KITCHENWARE	41-50	NORTH EASTERN	>200	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	SHT
1111	STANLEY WORK	41-50	NORTH EASTERN	101-150	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
1112	SUMMIT ADVANCE MATERIAL	31-40	CENTRAL	51-100	METAL	30,001-50,000	DOMESTIC	6-10	BUY	BLT/SHT/RL/DC
1113	SUPASIT	41-50	CENTRAL	51-100	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	BLT/DC
1114	SURACHAI LOHAKARN	41-50	NORTH	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	SHT
1115	TAIYO INDUSTRIAL	31-40	EASTERN	>200	METAL	>90,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
1116	TAT THAILAND	31-40	CENTRAL	51-100	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	DC
1117	THAI AIRWAY	41-50	CENTRAL	>200	METAL	30,001-50,000	DOMESTIC	11-15	BUY	RL
1118	THAI BURANA MFG	31-40	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/RL
1119	THAI FRAME & ACCESSORIES	31-40	NORTH	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/RL
1120	THAI INDUSTRIAL GAS TRADING	31-40	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC	1-5	BUY	BLT/RL
1121	THAI MAJOR	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT/RL
1122	THAI MAX WELL ELECTRIC	31-40	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/RL
1123	THAI OJMA PRECISION	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	DC
1124	THAI PART FUTOR	31-40	NORTH	151-200	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	SHT
1125	THAI PIYA KARLEK	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT/DC
1126	THAI PRESENT	41-50	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT/RL
1127	THAI SENG	41-50	NORTH EASTERN	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/SHT/RL/DC

This material is reserved for educational use only. Not allowed for commercial use.

CUSID	CUS_NAME	AGE	LOCATION	NO_EMP	IND_TYPE	SALES_AMOUNT	MARKET	YR_W_COM	BUY/NOT_BUY	RELATED_PRO
1128	THAI STAINLESS ENGINEERING	31-40	CENTRAL	101-150	METAL	<10,000	DOMESTIC	11-15	NOT BUY	BLT/RL
1129	THAI STEEL DOOR	31-40	CENTRAL	51-100	METAL	<10,000	DOMESTIC	1-5	NOT BUY	BLT/RL
1130	THAI STEEL SERVICE CENTER	31-40	EASTERN	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1131	THAI SUNAE	41-50	CENTRAL	101-150	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	DC
1132	THAI YONG ENGINEERING	31-40	CENTRAL	1-50	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	RL
1133	THAINOX STEEL	31-40	EASTERN	101-150	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
1134	THAPANIN METAL	31-40	EASTERN	51-100	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT
1135	THONGCHAI MACHINE	31-40	EASTERN	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	RL
1136	TITPAKORN	41-50	NORTH EASTERN	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1137	TOR KOLLOHA	31-40	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT/SHT
1138	TOSTEM THAI	31-40	CENTRAL	>200	METAL	10,001-30,000	EXPORT	1-5	NOT BUY	BLT/SHT
1139	TOWA SESAKUCHO	31-40	NORTH	>200	METAL	30,001-50,000	DOMESTIC/EXPORT	6-10	BUY	SHT
1140	TPW DIECASTING INDUSTRY	31-40	CENTRAL	1-50	METAL	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
1141	TSP	41-50	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	DC
1142	TWC PERMIT	41-50	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1143	UDOMRIT LOHAPHANT	31-40	CENTRAL	101-150	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/SHT/RL/DC
1144	UDOMSAK HARDWARE	31-40	CENTRAL	101-150	METAL	30,001-50,000	DOMESTIC	1-5	BUY	BLT
1145	UNION THAILAND	31-40	CENTRAL	>200	METAL	30,001-50,000	DOMESTIC/EXPORT	1-5	BUY	BLT/SHT
1146	UNION TOMITA	31-40	CENTRAL	151-200	METAL	>90,000	EXPORT	1-5	NOT BUY	BLT/RL
1147	UNION ZOJIRUSHI	31-40	CENTRAL	151-200	METAL	>90,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT/SHT/RL/DC
1148	VAROPRAKORN	41-50	CENTRAL	1-50	METAL	30,001-50,000	DOMESTIC	6-10	NOT BUY	BLT/SHT/RL
1149	VICHITPHOLCHOCK	31-40	NORTH EASTERN	51-100	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1150	VICHITRAPORNBRANGSEE	20-30	CENTRAL	151-200	METAL	10,001-30,000	DOMESTIC	11-15	BUY	BLT/RL
1151	VICTORY	41-50	CENTRAL	1-50	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
1152	VOR PORNSIN IND	31-40	CENTRAL	1-50	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT
1153	VR SAHAKIT	31-40	EASTERN	151-200	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/DC
1154	WATPAISARN ENG	31-40	CENTRAL	151-200	METAL	30,001-50,000	DOMESTIC	1-5	NOT BUY	RL
1155	WENMANN & SCHEIDER	31-40	CENTRAL	1-50	METAL	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT/DC
1156	WILASTHAI IRON	41-50	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	11-15	NOT BUY	SHT
1157	ZIP LOHAKIT	41-50	CENTRAL	51-100	METAL	10,001-30,000	DOMESTIC	6-10	NOT BUY	RL
2001	3IN IND	41-50	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC	11-15	NOT BUY	BLT
2002	A&P WOOD DESIGN	41-50	CENTRAL	101-150	WOOD	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	SHT
2003	ACME IND	31-40	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC/EXPORT	16-20	BUY	BLT/RL/SHT
2004	ACMEN INTERNATIONAL	31-40	CENTRAL	>200	WOOD	50,001-70,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT/RL
2005	ANCHALEE PRODUCTS	41-50	NORTH	51-100	WOOD	70,001-90,000	DOMESTIC	6-10	NOT BUY	BLT/RL
2006	AP WOOD	51-60	CENTRAL	151-200	WOOD	>90,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT/RL
2007	APINA IND	31-40	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2008	APS WOOD PRODUCT	31-40	CENTRAL	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2009	AUGUSTI FURNITURE & DEC	41-50	EASTERN	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2010	AUTO TRIM INTER	51-60	CENTRAL	101-150	WOOD	10,001-30,000	DOMESTIC	6-10	NOT BUY	BLT
2011	BANGKOK CABINET	41-50	NORTH	51-100	WOOD	50,001-70,000	EXPORT	6-10	BUY	BLT/SHT
2012	BANGKOK SIAM PRODUCT	20-30	CENTRAL	101-150	WOOD	10,001-30,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2013	BB HOME DESIGN	20-30	CENTRAL	101-150	WOOD	10,001-30,000	DOMESTIC/EXPORT	1-5	NOT BUY	RL
2014	BEST THAI HI-TECH	31-40	CENTRAL	101-150	WOOD	10,001-30,000	EXPORT	6-10	NOT BUY	BLT
2015	BESTA	41-50	CENTRAL	51-100	WOOD	10,001-30,000	DOMESTIC	11-15	BUY	BLT/SHT
2016	BS PARAWOOD PRO	31-40	SOUTH	51-100	WOOD	50,001-70,000	DOMESTIC	11-15	NOT BUY	BLT
2017	BUMRUNG THAI	41-50	EASTERN	51-100	WOOD	70,001-90,000	DOMESTIC	6-10	NOT BUY	BLT
2018	CITY WOOD	31-40	EASTERN	101-150	WOOD	50,001-70,000	EXPORT	6-10	NOT BUY	BLT
2019	DAIMOND FURNICON	31-40	NORTH EASTERN	151-200	WOOD	70,001-90,000	EXPORT	6-10	NOT BUY	BLT
2020	DD ALL WOOD	31-40	CENTRAL	101-150	WOOD	50,001-70,000	DOMESTIC	11-15	BUY	BLT
2021	DENCHAI PARQUET	41-50	NORTH	101-150	WOOD	50,001-70,000	EXPORT	11-15	NOT BUY	BLT
2022	EXTRA FORM IND	20-30	NORTH	151-200	WOOD	50,001-70,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT
2023	FANCY VENEER CORP	31-40	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2024	FUTURE WOOD	41-50	NORTH	101-150	WOOD	30,001-50,000	EXPORT	1-5	NOT BUY	BLT
2025	GOLDEN THAI TEAK	31-40	CENTRAL	51-100	WOOD	<10,000	DOMESTIC/EXPORT	6-10	BUY	BLT/RL/SHT
2026	HAWAI THAI EXPORT	41-50	CENTRAL	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT/SHT
2027	HOME CRAFTS	41-50	EASTERN	101-150	WOOD	30,001-50,000	DOMESTIC	6-10	NOT BUY	RL
2028	KAVEEPOJ	41-50	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2029	KK WOOD	31-40	CENTRAL	51-100	WOOD	50,001-70,000	DOMESTIC/EXPORT	11-15	BUY	BLT
2030	KONGKA MFG	20-30	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	RL
2031	LEE BROTHERS PICTURE FRAME	31-40	CENTRAL	101-150	WOOD	10,001-30,000	EXPORT	1-5	NOT BUY	SHT/RL
2032	LUANG THAVORN FURNITURE	41-50	NORTH	101-150	WOOD	10,001-30,000	EXPORT	1-5	NOT BUY	BLT
2033	MADERA PRODUCT	41-50	NORTH	151-200	WOOD	30,001-50,000	EXPORT	11-15	NOT BUY	SHT
2034	MANEEYA SHOP	41-50	CENTRAL	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	SHT

CUSID	CUS_NAME	AGE	LOCATION	NO_EMP	IND_TYPE	SALES_AMOUNT	MARKET	YR_W_COM	BUY/NOT_BUY	RELATED_PRO
2035	MASTER FLOORING	41-50	CENTRAL	51-100	WOOD	50,001-70,000	DOMESTIC	11-15	NOT BUY	BLT
2036	MASTER WOOD	31-40	CENTRAL	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	11-15	NOT BUY	SHT/RL
2037	MEE DEE	41-50	CENTRAL	151-200	WOOD	50,001-70,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT
2038	MEKN	31-40	EASTERN	151-200	WOOD	50,001-70,000	EXPORT	6-10	NOT BUY	BLT
2039	METROPOLITAN PLYWOOD	31-40	EASTERN	>200	WOOD	30,001-50,000	DOMESTIC	6-10	NOT BUY	RL
2040	MODERN PRODUCT	41-50	CENTRAL	151-200	WOOD	70,001-90,000	DOMESTIC	6-10	NOT BUY	BLT/SHT
2041	MODERNFORM GROUP	41-50	CENTRAL	>200	WOOD	10,001-30,000	DOMESTIC/EXPORT	1-5	BUY	BLT/SHT
2042	MP WOOD	41-50	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC	1-5	NOT BUY	BLT
2043	NEW MUANG-TONG FURNITURE	31-40	NORTH	51-100	WOOD	10,001-30,000	DOMESTIC	1-5	NOT BUY	BLT
2044	NEW SIAM FRAME	31-40	CENTRAL	151-200	WOOD	10,001-30,000	EXPORT	1-5	NOT BUY	BLT
2045	NUNG SIAM FURNITURE	20-30	CENTRAL	1-50	WOOD	10,001-30,000	DOMESTIC	1-5	NOT BUY	RL
2046	PACIFIC WOOD	31-40	CENTRAL	51-100	WOOD	50,001-70,000	EXPORT	1-5	NOT BUY	BLT/SHT
2047	PAN INDUSTRY PRODUCT	41-50	CENTRAL	101-150	WOOD	30,001-50,000	EXPORT	6-10	NOT BUY	BLT
2048	PANAWOOD	31-40	CENTRAL	51-100	WOOD	<10,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2049	PARTICLE PLANNER	41-50	SOUTH	51-100	WOOD	10,001-30,000	DOMESTIC/EXPORT	11-15	NOT BUY	SHT
2050	PICOTEE INTERIOR CENTER	41-50	NORTH EASTERN	151-200	WOOD	50,001-70,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT/RL
2051	PODIUM HOLDING GROUP	20-30	CENTRAL	101-150	WOOD	50,001-70,000	DOMESTIC/EXPORT	6-10	BUY	BLT/RL/SHT
2052	POLLARD IND	41-50	CENTRAL	51-100	WOOD	50,001-70,000	EXPORT	6-10	NOT BUY	BLT
2053	PRIME BOX MFG	41-50	EASTERN	151-200	WOOD	<10,000	DOMESTIC/EXPORT	1-5	BUY	BLT/SHT
2054	PROSPRO INTERTRADE	31-40	CENTRAL	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2055	QC PARAWOOD	51-60	CENTRAL	101-150	WOOD	10,001-30,000	DOMESTIC/EXPORT	16-20	BUY	BLT/RL/SHT
2056	ROYAL FRAME	31-40	CENTRAL	101-150	WOOD	10,001-30,000	EXPORT	6-10	NOT BUY	BLT/SHT
2057	RUANGUTAI WOOD	31-40	CENTRAL	51-100	WOOD	70,001-90,000	EXPORT	1-5	BUY	SHT
2058	RUNGRUANG PARQUET	31-40	NORTH EASTERN	101-150	WOOD	70,001-90,000	DOMESTIC/EXPORT	1-5	BUY	BLT/RL/SHT
2059	RUNGRUENG THAI WOOD	31-40	CENTRAL	101-150	WOOD	10,001-30,000	EXPORT	1-5	NOT BUY	BLT
2060	RYOKO & ASSOCIATED	20-30	CENTRAL	151-200	WOOD	30,001-50,000	EXPORT	11-15	NOT BUY	RL
2061	SANG THONG FRAME IND.	31-40	SOUTH	51-100	WOOD	30,001-50,000	EXPORT	6-10	BUY	BLT/RL/SHT
2062	SIAM FRAME IND	41-50	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC/EXPORT	11-15	BUY	BLT/SHT
2063	SIAM RICH WOOD	41-50	CENTRAL	51-100	WOOD	70,001-90,000	EXPORT	11-15	BUY	BLT/RL/SHT
2064	SIAM UNITED FRAME	31-40	CENTRAL	151-200	WOOD	<10,000	EXPORT	6-10	NOT BUY	SHT
2065	SIAM WOODEN PRODUCT	31-40	SOUTH	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT
2066	SINPRAPA	20-30	CENTRAL	1-50	WOOD	30,001-50,000	DOMESTIC	6-10	NOT BUY	BLT/RL
2067	SP WOODEN WORK	31-40	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT/RL
2068	SRI DHANACHAI	31-40	NORTH	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT/SHT
2069	SRIWATTANA WOODING IND	20-30	CENTRAL	51-100	WOOD	50,001-70,000	EXPORT	1-5	BUY	BLT/SHT
2070	STA GROUP	31-40	SOUTH	>200	WOOD	70,001-90,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT/RL/SHT
2071	STAR MARK MFG	41-50	CENTRAL	>200	WOOD	30,001-50,000	EXPORT	11-15	NOT BUY	RL
2072	SUN CABINET	20-30	CENTRAL	101-150	WOOD	>90,000	DOMESTIC/EXPORT	1-5	NOT BUY	BLT/RL
2073	SUN PARATECH	31-40	SOUTH	101-150	WOOD	>90,000	DOMESTIC/EXPORT	11-15	BUY	BLT/SHT
2074	SUNWOOD IND	51-60	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2075	SURINA IND	20-30	CENTRAL	151-200	WOOD	<10,000	EXPORT	6-10	NOT BUY	RL
2076	TAMAC	31-40	SOUTH	151-200	WOOD	70,001-90,000	EXPORT	1-5	NOT BUY	BLT
2077	TASK	41-50	CENTRAL	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2078	TCK FURNITURE	31-40	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	SHT/RL
2079	TEAK FURNITURE	41-50	CENTRAL	151-200	WOOD	10,001-30,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2080	TEE KARINCHANG	41-50	CENTRAL	1-50	WOOD	30,001-50,000	DOMESTIC	6-10	NOT BUY	SHT
2081	THAI BAMBOO & WOOD PRODUCT	31-40	CENTRAL	51-100	WOOD	50,001-70,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2082	THAI FANCY WOOD	41-50	NORTH EASTERN	151-200	WOOD	<10,000	DOMESTIC/EXPORT	11-15	NOT BUY	SHT/RL
2083	THAI INTORN ENTERPRISE	31-40	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC	1-5	BUY	BLT/RL/SHT
2084	THAI ITOKI	31-40	CENTRAL	>200	WOOD	30,001-50,000	EXPORT	6-10	NOT BUY	BLT
2085	THAI LOCAL PRODUCT	31-40	CENTRAL	1-50	WOOD	30,001-50,000	DOMESTIC	6-10	NOT BUY	BLT
2086	THAI PLYWOOD FACTORY	41-50	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC	6-10	NOT BUY	BLT/RL
2087	THAI SOFT WOODS	31-40	CENTRAL	>200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2088	THAIPONG SAWMILL	31-40	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC	11-15	NOT BUY	BLT
2089	TONGCHERN FURNITURE	41-50	CENTRAL	101-150	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2090	TPA IMPORT EXPORT	41-50	SOUTH	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2091	TRP FURNITURE	20-30	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2092	TVT WOOD IND	41-50	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC	11-15	NOT BUY	BLT
2093	TWC FURNIT	31-40	CENTRAL	>200	WOOD	30,001-50,000	EXPORT	6-10	NOT BUY	BLT
2094	UNIFURN	41-50	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	RL

CUSID	CUS_NAME	AGE	LOCATION	NO_EMP	IND_TYPE	SALES_AMOUNT	MARKET	YR_W_COM	BUY/NOT_BUY	RELATED_PRO
2095	UNIVERSAL PARAWOOD	20-30	CENTRAL	151-200	WOOD	10,001-30,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT
2096	VANACHAI GROUP	31-40	CENTRAL	>200	WOOD	10,001-30,000	DOMESTIC/EXPORT	6-10	NOT BUY	BLT
2097	VANACHAROEN	41-50	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	SHT
2098	VP WOOD PRODUCT	20-30	EASTERN	51-100	WOOD	30,001-50,000	DOMESTIC	11-15	BUY	BLT/SHT
2099	WONDER WORLD PRODUCT	31-40	CENTRAL	51-100	WOOD	30,001-50,000	DOMESTIC/EXPORT	11-15	BUY	BLT/SHT
2100	WOODY TOYS	51-60	CENTRAL	101-150	WOOD	<10,000	EXPORT	11-15	NOT BUY	BLT
2101	WORLD CRAFT	41-50	CENTRAL	151-200	WOOD	30,001-50,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT
2102	YANGMA SPORT TECH	41-50	CENTRAL	151-200	WOOD	<10,000	DOMESTIC/EXPORT	11-15	NOT BUY	BLT
3001	AUTO ALLIANCE THAILAND	31-40	EASTERN	>200	AUTO	30,001-50,000	DOMESTIC/EXPORT	1-5	NOT BUY	DC/SHT
3002	BANGCHAN GENERAL ASSEMBLY	41-50	CENTRAL	>200	AUTO	50,001-70,000	DOMESTIC	6-10	NOT BUY	DC
3003	EUROPEAN AUTOMOBILE	31-40	CENTRAL	151-200	AUTO	10,001-30,000	DOMESTIC	1-5	NOT BUY	DC/SHT
3004	HONDA CAR MANUFACTURING	31-40	CENTRAL	>200	AUTO	10,001-30,000	DOMESTIC/EXPORT	6-10	NOT BUY	DC/SHT
3005	ISUZU MOTOR THAILAND	41-50	CENTRAL	>200	AUTO	10,001-30,000	DOMESTIC	1-5	NOT BUY	DC/SHT
3006	MMC SITTIPOI (LAEM CHABANG)	31-40	EASTERN	>200	AUTO	30,001-50,000	DOMESTIC/EXPORT	6-10	NOT BUY	DC/SHT
3007	MMC SITTIPOI (LATKRABANG)	41-50	CENTRAL	>200	AUTO	50,001-70,000	DOMESTIC	6-10	NOT BUY	SHT
3008	SIAM KOLAKARN AND NISSAN	31-40	CENTRAL	>200	AUTO	30,001-50,000	DOMESTIC	6-10	NOT BUY	DC
3009	SIAM NISSAN AUTOMOBILE	31-40	CENTRAL	>200	AUTO	30,001-50,000	DOMESTIC	6-10	NOT BUY	DC/SHT
3010	SUKOSOL MAZDA MOTOR	31-40	CENTRAL	151-200	AUTO	50,001-70,000	DOMESTIC	6-10	NOT BUY	DC/SHT
3011	THAI AUTO WORK	41-50	CENTRAL	151-200	AUTO	10,001-30,000	DOMESTIC	6-10	NOT BUY	DC
3012	THAI HINO INDUSTRY	31-40	CENTRAL	151-200	AUTO	30,001-50,000	DOMESTIC	6-10	NOT BUY	RL
3013	THAI RUNG UNION CAR	31-40	CENTRAL	151-200	AUTO	50,001-70,000	DOMESTIC	6-10	NOT BUY	DC/SHT
3014	THAI SWEDISH ASSEMBLY	31-40	CENTRAL	151-200	AUTO	30,001-50,000	DOMESTIC	6-10	NOT BUY	DC/SHT
3015	THAI VARNYONT	41-50	CENTRAL	>200	AUTO	10,001-30,000	DOMESTIC	1-5	NOT BUY	DC
3016	THONBURI PRAKORBYONT	31-40	CENTRAL	151-200	AUTO	50,001-70,000	DOMESTIC	6-10	NOT BUY	DC
3017	TOYOTA MOTOR THAILAND	20-30	CENTRAL	>200	AUTO	50,001-70,000	DOMESTIC/EXPORT	1-5	BUY	DC/SHT

ประวัติผู้เขียน

ชื่อผู้เขียน	นายเอนก กองอมรภิญโญ
วัน เดือน ปีเกิด	วันที่ 6 พฤศจิกายน พ.ศ. 2510
สถานที่เกิด	กรุงเทพฯ
วุฒิการศึกษา	ปริญญาตรี บริหารธุรกิจ
สถานที่สำเร็จการศึกษา	มหาวิทยาลัยอัสสัมชัญ
ปีที่สำเร็จการศึกษา	ปีการศึกษา 2531
ประสบการณ์การทำงาน พ.ศ. 2531-2533	บริษัท มิคซุซ จำกัด
	ตำแหน่ง พนักงานขาย
พ.ศ. 2533-2537	บริษัท ทีไอเอ จำกัด
	ตำแหน่ง พนักงานขาย
พ.ศ. 2537-ปัจจุบัน	บริษัท 3 เอ็ม ประเทศไทย จำกัด
	ตำแหน่ง ผู้ประสานงานทางการตลาด