ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.

ระบบสารสนเทศเพื่อการควบคุมและเฝ้าระวังการติดเชื้อในโรงพยาบาล

Information Systems for Nosocomial Infection Control

โดย

นาง ดวงรัตน์ รักศักดิ์ศรี

รหัส 41067263

H002732

อาจารย์ที่ปรึกษา

รศ.ดร. วิเชียร เปรมชัยสวัสดิ์

วัน เดือน ปี 02 ธ.ค. 2550
เลขทะเบียน 02732
เลขเรียกหนังสือ วิท ด 2555 ร 542
"ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล."

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการศึกษากรณีพิเศษ
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
ภาคเรียนที่ 2 ปีการศึกษา 2542
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อหัวข้อ	ระบบสารสนเทศเพื่อการควบคุมและเฝ้าระวังการติดเชื้อในโรงพยาบาล
นักศึกษา	นาง ดวงรัตน์ รักศักดิ์ศรี
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร.วิเชียร เปรมชัยสวัสดิ์
ระดับการศึกษา	วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
แขนงวิชา	การจัดการเทคโนโลยีสารสนเทศ
ปีการศึกษา	2542

บทคัดย่อ

ในปัจจุบันได้มีการนำเทคโนโลยีสารสนเทศเข้ามาใช้ในธุรกิจต่างๆอย่างกว้างขวางโดยเฉพาะในโรงพยาบาล ระบบสารสนเทศได้เข้ามามีบทบาท ทั้งในด้านการรักษาพยาบาล การรวบรวมข้อมูลผู้รับบริการและการประเมินสุขภาพผู้รับบริการซึ่งมีความสำคัญมากต่อการให้การบริการ นอกจากนี้แล้วในทุกโรงพยาบาลที่มีการควบคุมโรคติดเชื้อในโรงพยาบาลที่ดีจะช่วยลดภาระค่าใช้จ่ายของผู้รับบริการและสถานพยาบาลลงตลอดจนระยะเวลาในการพักรักษาตัวด้วย

จากการนำระบบสารสนเทศเข้ามาใช้ในระบบการควบคุมและป้องกันการติดเชื้อในโรงพยาบาล ปัจจัยเสี่ยงที่มีผลต่อการติดเชื้อในโรงพยาบาล โดยพบว่าในกลุ่มผู้ป่วยที่ไม่ได้รับยาปฏิชีวนะกับกลุ่มผู้ป่วยที่ได้รับยาปฏิชีวนะมีความสัมพันธ์กับการที่ผู้ป่วยไม่พบการติดเชื้อในโรงพยาบาลและมีความสัมพันธ์ต่อการมีชีวิตหลังได้รับการรักษาโดยมีค่าความเชื่อมั่น 0.984 นอกจากนี้ยังพบว่าปัจจัยเสี่ยงในผู้ป่วยเพศหญิงมีความสัมพันธ์กับการไม่ติดเชื้อในโรงพยาบาลโดยมีค่าความเชื่อมั่น 0.951 และที่น่าสนใจคือศึกษาพบว่าอายุรแพทย์และศัลย์แพทย์ที่ทำงานเต็มเวลาจำนวน4.9เปอร์เซนต์สัมพันธ์กับผู้ป่วยที่พบการติดเชื้อในโรงพยาบาลโดยมีค่าความเชื่อมั่น 0.951

ผลจากการศึกษาแสดงให้เห็นถึงปัจจัยเสี่ยงที่มีผลต่อการติดเชื้อในโรงพยาบาลและสามารถแสดงลักษณะนิสัยของบุคลากรทางด้านการแพทย์ที่ให้การรักษาผู้ป่วย นอกจากนี้ยังเป็นการแสดงข้อมูลปัจจัยเสี่ยง ความสัมพันธ์ของปัจจัยต่างๆ และเป็นการกระตุ้นเตือนบุคลากรให้ตระหนักถึงความสำคัญ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการควบคุมปัจจัยเสี่ยงต่างๆดังนั้นจึงควรมีการสนับสนุนให้มีการใช้ระบบสารสนเทศเพื่อการควบคุมและป้องกันการติดเชื้อในโรงพยาบาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Information Systems for Nosocomial Infection Control
Student	Mrs. Duangrat Raksaksri
Advisor	Assoc.Prof.Dr. Wichian Premchaiswadi
Level of Study	Master of Science in Information Technology
Major	Information Technology Management
Academic Year	1999

ABSTRACT

Nowsdays, information technology has been widely used in various branches of business. Especially in hospitals, it plays a very important roles in medical treatment, collecting all patients' information and assessing theirs health condition which is a very important service. Besides this, a good control of nosocomial infection of every hospital will help to reduce both patients'cost and hospitals'expenses, including the length of stay in hospitals.

Using information systems to control and prevent nosocomial infection has revealed the risk factors affected to the infection. This study found that in groups of patients who did not take antibiotic medicine and the ones who did are related to the non-infectious patients and are also related to their staying alive after being cured at 0.984 confidence level. In addition, the risk factors of female patient are related to the non-infection at 0.951 confidence level. It is also interesting to find that 4.9 percent of full time medical doctors and surgeons is related to nosocomial infectious patients at 0.951 confidence level.

The result from the study shows the risk factors affected directly to nosocomial infection and the character of medical personnels who treat patients. It also indicate information about risk factors, the relationship of various factors,and encourage personnels to recognize the importance of

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

the control of many risk factors. Therefore, there should be a support to use information technology to and prevent nosocomial infection.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

โครงการศึกษากรณีพิเศษนี้สำเร็จได้ด้วยความกรุณาจากรองศาสตราจารย์ ดร. วิเชียร เปรมชัยสวัสดิ์ ที่ได้กรุณาให้คำปรึกษาคำแนะนำที่เป็นประโยชน์อย่างยิ่ง ตลอดจนตรวจสอบแก้ไขข้อบกพร่องต่าง ๆ ผู้เขียนขอกราบขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้

ขอกราบขอบพระคุณคณาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทความรู้และประสบการณ์ต่างๆ ขอกราบขอบพระคุณผู้ทรงคุณวุฒิทุกท่านที่กรุณาให้ข้อเสนอแนะที่เป็นประโยชน์ต่อการศึกษาโปรแกรมให้สมบูรณ์ขึ้น ขอขอบพระคุณโรงพยาบาลเกษมราษฎร์ ประชาชื่น ที่กรุณาอนุญาตให้ข้อมูลในการศึกษาครั้งนี้

ขอกราบขอบพระคุณผู้อำนวยการโรงพยาบาลเกษมราษฎร์ ประชาชื่น ที่ได้กรุณาอนุญาตให้เก็บรวบรวมข้อมูล ขอกราบขอบพระคุณผู้อำนวยการฝ่ายการพยาบาล หัวหน้าและบุคลากรพยาบาลหอผู้ป่วย หัวหน้างานและบุคลากรงานจุลชีวะวิทยา หัวหน้าหน่วยจ่ายกลาง โรงพยาบาลเกษมราษฎร์ประชาชื่น ตลอดจนผู้ป่วยในหอผู้ป่วยทุกท่านที่กรุณาให้ความกรุณาในการเก็บรวบรวมข้อมูล

สุดท้ายขอกราบขอบพระคุณ คุณพ่อ คุณแม่ และน้องๆ ที่ให้กำลังใจสนับสนุนและช่วยเหลืออย่างดียิ่ง ขอขอบคุณพี่ๆ เพื่อนๆ และน้องๆ ที่ร่วมงานที่ให้กำลังใจและการช่วยเหลือเสมอมา

ดวงรัตน์ รักศักดิ์ศรี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ III

กิตติกรรมประกาศ V

สารบัญ VI

สารบัญภาพประกอบ VII

บทที่

1. บทนำ 1

1.1 ความสำคัญและเหตุผลในการศึกษา 1

1.2 วัตถุประสงค์ 2

1.3 ขอบเขตของการศึกษา 2

1.4 ขั้นตอนการศึกษา 3

1.5 ประโยชน์ที่คาดว่าจะได้รับการศึกษา 3

1.6 คำจำกัดความที่ใช้ในการศึกษา 3

2. การติดเชื้อในโรงพยาบาลและระบบสารสนเทศ 6

2.1 ความหมายการติดเชื้อในโรงพยาบาล 6

2.2 หลักเกณฑ์ในการวินิจฉัยโรคติดเชื้อในโรงพยาบาล 7

2.3 วิทยาการระบาดของการติดเชื้อในโรงพยาบาล 8

2.4 ปัจจัยส่งเสริมการติดเชื้อในโรงพยาบาล 9

2.5 ทำไมต้องใช้ Data Mining 9

2.6 กรอบแนวความคิดที่ใช้ในการศึกษา 11

3. DATA MINING และทฤษฎี 12

3.1 ความหมายของ Data Mining 12

3.2 ชนิดของ Data Mining 13

3.3 Data Warehouse และ Data Mining 14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

หน้า

3.4 Star Schemas 16

3.5 เทคนิคของ Data Mining 23

3.6 กรณีศึกษาขบวนการทำงานของ Data Mining 26

4. Decision Tree และขบวนการดำเนินในระบบ 33

4.1 ลักษณะของประชากรที่ทำการศึกษา 33

4.2 เครื่องมือที่ใช้ในการศึกษา 33

4.3 การรวบรวมข้อมูล 33

4.4 การวิเคราะห์ข้อมูลโดยใช้ Data Mining 34

4.5 ขอบเขตของระบบที่นำมาศึกษา 38

4.6 การอภิปรายผลจากขบวนการ Data Mining 38

5. บทสรุป 73

5.1 สรุปผลการศึกษา 73

5.2 ประโยชน์ที่ได้รับจากการศึกษา 76

5.3 ข้อเสนอแนะในการนำผลการศึกษาไปใช้ 76

5.4 ข้อเสนอแนะในการทำการศึกษาครั้งต่อไป 76

บรรณานุกรม 78

ภาคผนวก 79

ภาคผนวก ก ODBC 80

ภาคผนวก ข สถิติที่ใช้ในการวิเคราะห์ 83

ภาคผนวก ค แบบฟอร์มรายงานการเฝ้าระวังโรคแบบต่อเนื่อง 85

ประวัติผู้เขียน 86

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

หน้า

ภาพที่

1 แสดงความสัมพันธ์ปัจจัยทางระบาดวิทยาของการติดเชื้อในโรงพยาบาล 11

2 แสดง Star Schemas ของปัจจัยที่ต้องการศึกษา 17

3 แสดงความสัมพันธ์ 3 มิติ ของปัจจัยที่ต้องการศึกษา 18

4 แสดงการ Slice และDice เพื่อหาความสัมพันธ์ของปัจจัยที่ต้องการศึกษา 19

5 แสดงลำดับขั้นของฟิลด์ที่ทำการศึกษาปัจจัยที่เกี่ยวข้องกับการติดเชื้อในโรงพยาบาล 20

6 แสดงการนอร์มัลไลซ์ตารางมิติความสัมพันธ์ 21

7 แสดงตาราง Multiple Fact Tables 22

8 รูปแบบการวิเคราะห์ข้อมูลโดยใช้ขบวนการเครือข่ายระบบประสาท 24

9 แสดงรูปแบบเครือข่ายระบบประสาทวิเคราะห์โมเลกุลยา 24

10 แสดงความสัมพันธ์ของการศึกษาการดูดซึมของยาโดยการทำนายเป็นกราฟ 25

11 แสดงรูปแบบของการจัดกลุ่มยีนส์ของเชื้อราแบบ Cluster 25

12 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DX และ TY_RM 39

13 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง PT_NI_DX และ PT_DEP 40

14 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_DRUG_ANT 41

15 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ DR_ID 42

16 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_STATUS_DIS 43

17 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_AGE 44

18 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_SEX 45

19 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_TY_NI 45

20 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ AD_DAY 46

21 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ AD_MONTH 47

22 แสดงรายงานความถี่ของแพทย์ที่รับรักษาผู้ป่วย (DR_ID) 48

23 แสดงรายงานความถี่ของกลุ่มผู้ป่วยที่ทำการศึกษาการเฝ้าระวังโรค 48

24 แสดงรายงานความถี่ของการวินิจฉัยการติดเชื้อในโรงพยาบาล 49

25 แสดงรายงานความถี่ของของผลการรักษากลุ่มผู้ป่วยที่ทำการศึกษา 49

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

หน้า

ภาพที่

1 แสดงความสัมพันธ์ปัจจัยทางระบาดวิทยาของการติดเชื้อในโรงพยาบาล 11

2 แสดง Star Schemas ของปัจจัยที่ต้องการศึกษา 17

3 แสดงความสัมพันธ์ 3 มิติ ของปัจจัยที่ต้องการศึกษา 18

4 แสดงการ Slice และDice เพื่อหาความสัมพันธ์ของปัจจัยที่ต้องการศึกษา 19

5 แสดงลำดับขั้นของฟิลด์ที่ทำการศึกษาปัจจัยที่เกี่ยวข้องกับการติดเชื้อในโรงพยาบาล 20

6 แสดงการนอร์มัลไลซ์ตารางมิติความสัมพันธ์ 21

7 แสดงตาราง Multiple Fact Tables 22

8 รูปแบบการวิเคราะห์ข้อมูลโดยใช้ขบวนการเครือข่ายระบบประสาท 24

9 แสดงรูปแบบเครือข่ายระบบประสาทวิเคราะห์โมเลกุลยา 24

10 แสดงความสัมพันธ์ของการศึกษาการดูดซึมของยาโดยการทำนายเป็นกราฟ 25

11 แสดงรูปแบบของการจัดกลุ่มยีนส์ของเชื้อราแบบ Cluster 25

12 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DX และ TY_RM 39

13 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง PT_NI_DX และ PT_DEP 40

14 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_DRUG_ANT 41

15 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ DR_ID 42

16 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_STATUS_DIS 43

17 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_AGE 44

18 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_SEX 45

19 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ PT_TY_NI 45

20 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ AD_DAY 46

21 แสดงความสัมพันธ์ในรูปแบบ 3 มิติระหว่าง TY_NI_DXและ AD_MONTH 47

22 แสดงรายงานความถี่ของแพทย์ที่รับรักษาผู้ป่วย (DR_ID) 48

23 แสดงรายงานความถี่ของกลุ่มผู้ป่วยที่ทำการศึกษาการเฝ้าระวังโรค 48

24 แสดงรายงานความถี่ของการวินิจฉัยการติดเชื้อในโรงพยาบาล 49

25 แสดงรายงานความถี่ของของผลการรักษากลุ่มผู้ป่วยที่ทำการศึกษา 49

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ(ต่อ)

ภาพที่ หน้า

26 แสดงความถี่ของจำนวนผู้ป่วยเปรียบเทียบจำนวนวันที่นอนในโรงพยาบาล 50

27 แสดงรายงานการจุดความถี่ระหว่างผู้ป่วยกับจำนวนวันที่นอนในโรงพยาบาล 50

28 แสดงรูปแบบของระบบในการออกรายงานการวิเคราะห์ข้อมูล 51

29แสดงรูปแบบของระบบในการสร้างความสัมพันธ์เชิงกรณีที่ใช้ในการวิเคราะห์ 51

30 แสดงรายงานการจุดความถี่ระหว่างจำนวนผู้ป่วยต่อแพทย์ที่ทำการรักษาผู้ป่วย 52

31 ภาพแสดงรายงานตัวอย่างตารางข้อมูลที่ทำการวิเคราะห์ข้อมูล 52

32 ภาพแสดงการเลือกขอบเขตค่าในการวิเคราะห์ข้อมูลแบบ ไม่จัดกลุ่ม 53

33 ภาพแสดงการกำหนดค่าของการเลือกกฎในการสร้างความสัมพันธ์ 53

34 ภาพแสดงการเชื่อมโยงข้อมูลโดยใช้ Microsoft Access 54

35 ภาพแสดงผลการวิเคราะห์โดยใช้ความสัมพันธ์เชิงกรณี 54

36 ภาพแสดง Table View และกราฟของ TY_NI_DX ที่ทำการวิเคราะห์ 55

37 ภาพแสดงขั้นตอนในการเข้าระบบเพื่อเชื่อมต่อกับฐานข้อมูลผู้ป่วยขั้นตอนที่ 1 55

38ภาพแสดงขั้นตอนในการเข้าระบบเพื่อเชื่อมต่อกับฐานข้อมูลผู้ป่วยขั้นตอนที่ 2 56

39ภาพแสดงขั้นตอนในการเข้าระบบเพื่อวิเคราะห์แบบโครงสร้างต้นไม้โดยใช้ Tree Miner 56

40 ภาพแสดงการคัดเลือกกิ่งที่ทำการศึกษาโดยการแสดงหัวข้อให้เลือก 57

41 ภาพแสดง Node แรกที่ทำการคัดเลือกกิ่งของต้นไม้ 57

42 ภาพแสดงการแตกกิ่งโครงสร้างต้นไม้คัดเลือกระหว่าง PT_DRUG_ANT และPT_CLASS 58

43 ภาพแสดงการแตกกิ่งโครงสร้างต้นไม้ระหว่างPT_DRUG_ANT , PT_CLASS,TY_NI_DX 58

44 ภาพแสดงการแตกกิ่งโครงสร้างต้นไม้ระหว่าง PT_SEX, PT_DRUG_ANT, PT_CLASS, TY_NI_DX 59

45 แสดงค่าความสัมพันธ์เชิงเงื่อนไขที่วิเคราะห์ได้ 60

46 ภาพแสดงการแตกกิ่งโครงสร้างต้นไม้ระหว่าง TY_NI_DX , PT_DEP, DR_ID 60

47 แสดงการเข้าระบบเพื่อสร้างโครงสร้างต้นไม้โดยใช้ Tree Miner 61

48 แสดงกำหนดขอบเขตค่าในการวิเคราะห์ความสัมพันธ์แบบจัดกลุ่ม 61

49 ภาพแสดงความสัมพันธ์ระหว่างเงื่อนไขระหว่าง PT_CLASS, PT_DRUG_ANT, TY_NI_DX 62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ(ต่อ)

ภาพที่	หน้า
50 ภาพแสดงความสัมพันธ์ระหว่างเงื่อนไขที่ทำการศึกษาแบบที่ 1 PT_DEP,TY_NI_DX, PT_STATUS_DIS	63
51 ภาพแสดงโครงสร้างต้นไม้โดยพิจารณาเงื่อนไขแบบที่ 2PT_DEP,TY_NI_DX, PT_STATUS_DIS	63
52 ภาพแสดงโครงสร้างต้นไม้โดยพิจารณาเงื่อนไข PT_AGE, TY_NI_DX, PT_STATUS_DIS	65
53ภาพแสดงโครงสร้างต้นไม้โดยพิจารณาเงื่อนไข PT_DRUG_ANT, PT_TY_NI, TY_NI_DX แบบที่ 1	66
54 ภาพแสดงโครงสร้างต้นไม้โดยพิจารณาเงื่อนไข PT_DRUG_ANT, PT_TY_NI, TY_NI_DX แบบที่ 2	67
55 แสดงการจัดแยกกลุ่มผู้ป่วยที่ทำการศึกษาที่มีค่าแบบ Discrete	68
56 แสดงการจัดแยกกลุ่มผู้ป่วยที่ทำการศึกษาเป็น 2 กิ่งสาขา ค่าที่แยกคือDR_ID	68
57 ภาพแสดงโครงสร้างต้นไม้โดยพิจารณาเงื่อนไข AD_DAY, PT_DEP, TY_NI_DX	69
58 ภาพแสดงโครงสร้างต้นไม้โดยพิจารณาเงื่อนไข PT_SEX, DR_ID, PT_DEP, TY_NI_DX	70
59 ภาพแสดงโครงสร้างต้นไม้โดยพิจารณากลุ่มผู้ป่วยที่ทำการศึกษาแยกตาม TY_NI_DX TY_RM	71
60 ภาพแสดงโครงสร้างต้นไม้โดยพิจารณาเชิงกรณีระหว่าง TY_RM, PT_CLASS TY_NI_DX	72

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความสำคัญและเหตุผลในการศึกษา

ในการดำเนินธุรกิจโรงพยาบาลในสถานการณ์ปัจจุบันมีการการตื่นตัวจากกระแสการปฏิรูประบบบริการสุขภาพ มีการแข่งขันกันทางด้านการดำเนินธุรกิจบริการด้าสุขภาพ การแข่งขันกันอย่างรุนแรงนี้เพื่อให้องค์กรอยู่รอดไม่ว่าจะเป็นการแข่งขันในด้านราคาคุณภาพการให้บริการตลอดจนประสิทธิภาพในการรักษาพยาบาล โรงพยาบาลต่างๆต้องมีการปรับเปลี่ยนกลยุทธ์ทางธุรกิจใหม่เพื่อให้ผู้รับบริการเกิดความพึงพอใจจึงได้มีการศึกษารูปแบบการจัดบริการมาใช้เป็นแนวทางในการกำหนดค่าใช้จ่ายการใช้ทรัพยากรต่างๆประกอบกับแนวคิดในการพัฒนาคุณภาพโรงพยาบาลมีการรับรองโรงพยาบาลคุณภาพ (Hospital Accreditation) ที่มุ่งเน้นการพัมนาคุณภาพการดูแลอย่างต่อเนื่องโดยพิจารณาเน้นในด้านการประเมินตนเอง (Self Assessment) กระบวนการรักษาพยาบาลและผู้ป่วยพิจารณาผลกระทบที่มีต่อองค์กรการกำหนดเครื่องชี้วัดคุณภาพ (Quality Indicator) ที่มีความถูกต้องน่าเชื่อถือและใช้ประโยชน์ได้ดีในทางปฏิบัติเป็นสิ่งที่มีความสำคัญเป็นอย่างสำหรับธุรกิจโรงพยาบาลกระบวนการคุณภาพการดูแลนั้นสามารถประเมินได้จากผลลัพธ์ (Outcome) เช่นอัตราการตาย การเกิดภาวะแทรกซ้อนในโรงพยาบาล จำนวนวันที่นอนในโรงพยาบาล อัตราการครองเตียง อัตราการติดเชื้อในโรงพยาบาล อัตราการเข้ารับการรักษาซ้ำในโรงพยาบาล สิ่งที่สำคัญประการหนึ่งคือการพิจารณาศึกษาปัจจัยที่มีผลต่อการเกิดผลลัพธ์เหล่านี้ สิ่งที่มีความจำเป็นคือการศึกษาปัจจัยที่มีผลต่อการติดเชื้อในโรงพยาบาลซึ่งจะมีกระทบต่อต้นทุนและคุณภาพของการให้บริการ ปัจจัยต่างๆที่อาจส่งผลต่อการติดเชื้อในผู้ป่วย

ดังนั้นโรงพยาบาลหลายแห่งจึงได้นำเทคโนโลยีสารสนเทศเข้ามามีบทบาทในองค์กรเพิ่มขึ้น เช่น ในการจัดการคลังข้อมูล (Data Warehouse) ทั้งฐานข้อมูลผู้ป่วย ยา เวชภัณฑ์ตลอดจนข้อมูลทางด้านการแพทย์ การรักษาต่าง ๆ เป็นการรวบรวมปัจจัยที่สนใจจากหลากหลายระบบสิ่งที่สำคัญสำหรับการดำเนินธุรกิจโรงพยาบาลคือ การทำให้ผู้รับบริการปลอดภัยและหายจากการเจ็บป่วยได้อย่างรวดเร็วแต่เนื่องจาก โรงพยาบาลเป็นแหล่งที่รวมของผู้รับบริการที่เจ็บป่วยทั้งที่เกิดจากการติดเชื้อ โรคไร้เชื้อตลอดจนมีบุคลากรทางด้านการแพทย์หลายระดับทีให้บริการผู้ป่วย ประกอบกับสภาพแวดล้อมของโรงพยาบาลที่ส่งเสริมให้เกิดจากการติดเชื้อในโรงพยาบาลซึ่งมีผลกระทบต่อผู้รับบริการ,ญาติ, บุคลากรต่าง ๆ การศึกษาปัจจัยที่มีผลต่อการควบคุมและการป้องกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การติดเชื้อในโรงพยาบาลและค้นหาความสัมพันธ์ของปัจัยต่างๆจะช่วยให้สถานพยาบาลสามารถเฝ้าระวังปัจจัยเหล่านั้นไว้หรือหาวิธีในการป้องกันให้กับผู้ป่วยเพื่อให้เกิดความปลอดภัยจากการติดเชื้อนั้นๆ

การนำระบบสารสนเทศเข้ามาช่วยในการวิเคราะห์ปัจจัยเหล่านี้นั้นต้องคำนึงถึงว่าปัจจัยต่างๆมีมากมายและบางครั้งเป็นปัจจัยที่ไม่สามารถควบคุมได้ การพิจารณาโดยนำเอาเทคนิค Data Mining เข้ามาใช้ในการออกแบบและพัฒนาระบบการควบคุมและการป้องกันการติดเชื้อในโรงพยาบาลจึงมีความจำเป็นอย่างยิ่ง

1.2วัตถุประสงค์

การนำเอาเทคนิค Data Mining มาใช้ในการออกแบบการระบบการควบคุมและการป้องกันการติดเชื้อในโรงพยาบาลเพื่อใช้ในการวิเคราะห์ปัจจัยที่มีผลต่อการติดเชื้อในโรงพยาบาลให้เกิดประสิทธิภาพและสามารถนำผลที่ได้ไปวางแผนเพื่อป้องกันและควบคุมการติดเชื้อในโรงพยาบาลให้มีคุณภาพการรักษาพยาบาลที่ได้มาตราฐานโดยมีการกำหนดเป้าหมายเพื่อศึกษาค้นหาความสัมพันธ์ของปัจจัยที่มีผลต่อการการติดเชื้อในโรงพยาบาลโดยนำผลที่ได้จากการศึกษามาวางแผนในการป้องกันการติดเชื้อในโรงพยาบาลตลอดจนเป็นการลดค่าใช้จ่ายในการจัดทำระบบสารสนเทศเพื่อการควบคุมและป้องกันการติดเชื้อในโรงพยาบาล

1.3 ขอบเขตของการศึกษา

การติดเชื้อในโรงพยาบาลอาจเกิดขึ้นได้กับผู้ป่วยทุกรายแต่ประชากรที่ทำการศึกษาคือกลุ่มที่ทำการเฝ้าระวังการติดเชื้อโรคในโรงพยาบาลเกษมราษฎร์โดยแบ่งกลุ่มประชากรที่ทำการศึกษาออกเป็น4กลุ่มคือผู้ป่วยหนักที่ต้องได้รับการดูแลอย่างใกล้ชิด ผู้ป่วยที่นอนในโรงพยาบาลนานเกิน 3วันผู้ป่วยที่ได้รับการสอดใส่อุปกรณ์เข้าไปในร่างกายและผู้ป่วยหลังได้รับการผ่าตัด/คลอดโดยรวบรวมข้อมูลผู้ป่วยในระหว่างวันที่ 1 มกราคม 2542 - 31ธันวาคม 2542

ขอบเขตของการศึกษานี้จะพิจารณาการติดเชื้อในโรงพยาบาลนี้ได้ทำการศึกษาโดยการนำเอาซอฟท์แวร์สำเร็จรูป (Data Mining Technique) มาปรับปรุงเพื่อใช้ในการออกแบบระบบสารสนเทศเพื่อควบคุมและป้องกันการติดเชื้อในโรงพยาบาลโดยการแปลงฐานข้อมูลเดิมของโรงพยาบาลเป็นฐานข้อมูลใหม่ที่ใช้ในการศึกษาค้นหาความสัมพันธ์ของปัจจัยที่มีผลต่อการติดเชื้อในโรงพยาบาลและพิจารณาร่วมกับการรายงานข้อมูลการเฝ้าระวังโรคติดเชื้อในโรงพยาบาลประจำเดือน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4 ขั้นตอนการศึกษา

เพื่อให้การศึกษาบรรลุตามวัตถุประสงค์ที่กำหนดไว้ภายใต้ขอบเขตของการศึกษาจึงได้กำหนดขั้นตอนในการศึกษาไว้ดังนี้

ขั้นที่ 1 เก็บรวบรวมข้อมูล

ขั้นที่ 2 ศึกษาทฤษฎีที่เกี่ยวข้อง

ขั้นที่ 3 ศึกษาวิธีการทำงานและแนวทางในการนำเอาเทคนิคใหม่มาใช้

ขั้นที่ 4 สรุปผลการศึกษา

1.5 ประโยชน์ที่คาดว่าจะได้รับการศึกษา

จากการศึกษาระบบสารสนเทศโดยนำซอฟท์แวร์สำเร็จรูป (Data Mining Technique) มาใช้ในระบบสารสนเทศเพื่อการควบคุมและการป้องกันการติดเชื้อในโรงพยาบาล โดยคาดหวังว่าจะได้รับประโยชน์ดังนี้

1.เพื่อพัฒนาระบบสารสนเทศการควบคุมและป้องกันการติดเชื้อในโรงพยาบาลให้มีประสิทธิภาพ

2.เพื่อลดค่าใช้จ่ายในการจัดทำระบบสารสนเทศเพื่อการควบคุมและป้องกันการติดเชื้อในโรงพยาบาล

3. ค้นหาปัจจัยความเสี่ยงต่อการติดเชื้อในโรงพยาบาลที่มีความสัมพันธ์กัน

4. เพื่อนำผลที่ได้นำไปการวางแผนกำหนดมาตราการในการควบคุมและป้องกันการติดเชื้อในโรง พยาบาล

5.ช่วยให้ผู้รับบริหารได้รับข้อมูลที่มีคุณภาพช่วยในการตัดสินใจ

6.ช่วยให้ผู้รับบริการได้รับการดูแลที่มีคุณภาพและมีประสิทธิภาพ (Reinforcing Practices) จากการพัฒนาระบบการเฝ้าระวังโรคติดเชื้อในโรงพยาบาล

1.6 คำจำกัดความที่ใช้ในการศึกษา

TY_RM หมายถึงประเภทหอผู้ป่วยที่รับผู้ป่วยไว้พักรักษาแบ่งเป็น 5 หอผู้ป่วยคือ

ICU หมายถึงหน่วยงานของโรงพยาบาลที่ให้บริการผู้ป่วยในซึ่งอยู่ในภาวะวิกฤตหรือต้องการการดูแลอย่างใกล้ชิดโดยมีจำนวนเตียง 7เตียงและรับผู้ป่วยจากทุกแผนกในโรงพยาบาล

W5A หมายถึงหน่วยงานของโรงพยาบาลที่ให้บริการผู้ป่วยในโดยมีจำนวนเตียงที่ให้บริการจำนวน26 เตียงมีลักษณะห้องเป็นห้องเดี่ยว

W5C หมายถึง หน่วยงานของโรงพยาบาลที่ให้บริการผู้ป่วยในทั่วไปมีห้องพักที่พร้อมให้บริการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวน 8 ห้องโดยมีลักษณะห้องเป็นห้องคู่จำนวน 3ห้อง ห้องรวม 4เตียงจำนวน 5 ห้อง

W4A หมายถึง หน่วยงานในโรงพยาบาลที่ให้บริการผู้ป่วยในทั่วไป มีห้องพักที่พร้อมให้บริการ จำนวน 12 ห้องมีลักษณะเป็นห้องเดี่ยว

NR หมายถึงหน่วยงานของโรงพยาบาลที่ให้บริการทารกแรกเกิดและทารกที่มีอาการผิดปกติ หรือต้องได้รับการดูแลอย่างใกล้ชิดมีเตียงรองรับผู้ป่วยได้จำนวน 10 คลิบ

TY_NI_DX หมายถึงการวินิจฉัยว่าผู้ป่วยได้รับการติดเชื้อในโรงพยาบาลหรือไม่โดยกำหนดดังนี้

INF หมายถึงการติดเชื้อโรคในโรงพยาบาลเนื่องมาจากผู้ป่วยได้รับเชื้อจุลชีพขณะอยู่ในโรงพยาบาลโดยผู้ป่วยไม่ปรากฏอาการและอาการแสดงของการติดเชื้อและไม่ได้อยู่ในระยะฟักตัวของเชื้อขณะเข้ารับการรักษาการวินิจฉัยการติดเชื้อในโรงพยาบาลใช้เกณฑ์การวินิจฉัยของศูนย์ควบคุมโรคสหรัฐอเมริกา ปีคศ.1988

None หมายถึงไม่พบว่าผู้ป่วยได้รับการติดเชื้อโรคในโรงพยาบาล

PT_DEP หมายถึงแผนกผู้ป่วยที่ทำการรักษาแบ่งได้ดังนี้

Med	แผนกอายุรกรรม	Surg	แผนกศัลย์กรรม
Ped	แผนกกุมารเวช	Orthro	แผนกศัลย์กรรมกระดูก
Obgyn	แผนกสูติ-นรีเวช	Eye	แผนกตา
ENT	แผนกหู คอ จมูก	NS	แผนกศัลย์กรรมระบบประสาท

PT_STATUS_DIS หมายถึงสถานะผู้ป่วยภายหลังได้รับการรักษาในโรงพยาบาล

ALIVE ผู้ป่วยที่ได้รับการรักษาในโรงพยาบาลแล้วมีชีวิต

DEAD ผู้ป่วยที่ได้รับการรักษาในโรงพยาบาลแล้วเสียชีวิต

PT_HN หมายถึงรหัสประจำตัวผู้ป่วย

PT_AN หมายถึงรหัสประจำตัวผู้ป่วยใน

PT_SEX หมายถึงเพศของผู้ป่วยที่ทำการศึกษา (F=เพศหญิง) (M=เพศชาย)

PT_AGE หมายถึงอายุของผู้ป่วยที่เข้ารับการรักษาโดยแบ่งเป็นช่วงอายุ (0-1) (2-11) (12-20) (21-30) (31-50) (51-60) (>60) หน่วยเป็นปี

DR_ID หมายถึง รหัสของแพทย์แต่ละท่านที่ทำการรักษาผู้ป่วย

AD_DAY หมายถึง จำนวนวันที่ผู้ป่วยนอนในโรงพยาบาลหน่วยเป็นวัน

PT_TY_NI หมายถึงประเภทของผู้ป่วยที่ทำการศึกษาการเฝ้าระวังโรคติดเชื้อในโรงพยาบาลโดยกำหนดเป็นหมายเลขดังนี้

1= ผู้ป่วยที่นอนรักษาในห้องผู้ป่วยหนัก

2= ผู้ป่วยที่นอนในโรงพยาบาล>3วัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3= ผู้ป่วยที่ได้รับการสอดใส่อุปกรณ์ทางการแพทย์

4= ผู้ป่วยที่ได้รับการผ่าตัด/คลอด

PT_CLASS หมายถึงการแบ่งประเภทผู้ป่วยโดยแบ่งเป็น 2 ประเภทคือ

1 = ผู้ป่วยทั่วไป 2=ผู้ป่วยประกันสังคม

PT_DRUG_ANT หมายถึงการได้รับยาปฏิชีวนะในผู้ป่วยที่รับการรักษาในโรงพยาบาล

1= ผู้ป่วยได้รับยาปฏิชีวนะ 0=ผู้ป่วยไม่ได้รับยาปฏิชีวนะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
การติดเชื้อในโรงพยาบาลและระบบสารสนเทศ

2.1 ความหมายการติดเชื้อในโรงพยาบาล

ในประเทศไทย กองระบาดวิทยา กระทรวงสาธารณสุขได้จัดการประชุมเชิงปฏิบัติการเพื่อกำหนดนิยามสำหรับการวินิจฉัยโรคติดเชื้อในโรงพยาบาลขึ้นเมื่อ พ.ศ. 2534 โดยเรียบเรียงจากการวินิจฉัยโรคติดเชื้อในโรงพยาบาล ศูนย์ควบคุมโรค (CDC) ของสหรัฐอเมริกา ปี ค.ศ. 1988 โดยให้คำนิยามว่า โรคติดเชื้อในโรงพยาบาลหมายถึง การติดเชื้ออันเป็นผลจากการที่ผู้ป่วยได้รับเชื้อโรคขณะที่อยู่ในโรงพยาบาล ซึ่งเป็นเชื้อที่อยู่ในร่างกายผู้ป่วยเอง (Endogenous Organism) หรือเป็นเชื้อที่จากภายนอกร่างกายผู้ป่วย (Exogenous Organism) โดยขณะที่ผู้ป่วยเข้ารับการรักษาตัวในโรงพยาบาล ผู้ป่วยไม่มีอาการและอาการแสดงของการติดเชื้ออยู่และไม่ได้อยู่ในระยะฟักตัวของเชื้อ หากไม่ทราบว่าระยะฟักตัวของเชื้อให้ถือว่าเป็นโรคติดเชื้อในโรงพยาบาลหากพบว่าการติดเชื้อของโรคนั้นปรากฏอาการหลังจากผู้ป่วยเข้ารับการรักษาตัวในโรงพยาบาล 48 ชั่วโมงการติดเชื้อในโรงพยาบาลอาจเกิดขึ้นได้แทบทุกระบบทางร่างกาย การติดเชื้อที่เกิดขึ้นจะทำให้ผู้ป่วยมีอาการที่เฉพาะเจาะจง ตลอดจนอาการและอาการแสดงที่แตกต่างกันไปโดยแบ่งการติดเชื้อในโรงพยาบาลได้ดังนี้

1. การติดเชื้อระบบทางเดินปัสสาวะ
2. การติดเชื้อระบบทางเดินหายใจ
3. การติดเชื้อระบบหัวใจและหลอดเลือด
4. การติดเชื้อที่ตำแหน่งผ่าตัด
5. การติดเชื้อในกระแสโลหิตแบบปฐมภูมิ
6. การติดเชื้อระบบประสาทส่วนกลาง
7. การติดเชื้อที่ผิวหนัง
8. การติดเชื้อที่ปอด
9. การติดเชื้อที่กระดูกและข้อ
10. การติดเชื้อระบบทางเดินอาหาร
11. การติดเชื้อที่ตา หู จมูก คอและปาก
12. การติดเชื้อที่ระบบสืบพันธุ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 หลักเกณฑ์ในการวินิจฉัยโรคติดเชื้อในโรงพยาบาล

ข้อมูลที่ใช้ในการพิจารณาการติดเชื้อในผู้ป่วยจะได้มาจากข้อมูลต่างๆทั้งจากการติดตามเวชระเบียน จากการติดตามอาการผู้ป่วย การบันทึกของแพทย์และเจ้าหน้าที่พยาบาล รวมทั้งจากการซักถามหรือคำบอกเล่าของผู้ป่วยเองจะทำให้สามารถวินิจฉัยได้ว่าผู้ป่วยติดเชื้อในโรงพยาบาลโดยพิจารณาข้อมูลดังต่อไปนี้

1. อาการและอาการแสดงของผู้ป่วย การติดเชื้อทำให้เกิดอาการผิดปรกติในร่างกายอาการและอาการแสดงขึ้นอยู่กับเชื้อที่เป็นสาเหตุและตำแหน่งที่เกิดจากติดเชื้อ ได้แก่
   - 1.1 เจ็บคอ ไอ
   - 1.2 เสมหะมาก เสมหะเปลี่ยนสี
   - 1.3 ไข้ขึ้นอุณหภูมิสูงกว่า 37.5 องศาเซนติเกรด
   - 1.4 หายใจเร็ว
   - 1.5 ผื่นขึ้น ผิวหนังอักเสบ
   - 1.6 อุจจาระร่วง
   - 1.7 คลื่นไส้ อาเจียน
   - 1.8 น้ำหนักลด
   - 1.9 มีหนองที่แผลผ่าตัด
   - 1.10 ปัสสาวะแสบ ปัสสาวะขัด
2. ผลการตรวจทางห้องปฏิบัติการ ได้แก่ ผลการตรวจเพาะเชื้อ การตรวจหาตัวเชื้อ (Antigen) การตรวจหาระดับภูมิคุ้มกัน (Antibody) การตรวจ Complete Blood Count เป็นต้น
3. การตรวจวินิจฉัยอื่นๆที่มีส่วนช่วยบ่งชี้ว่าผู้ป่วยมีการติดเชื้อ ได้แก่การถ่ายภาพรังสี(X- ray) การตรวจด้วยการส่องกล้อง (Endoscopic procedures) การตรวจชิ้นเนื้อ (Biopsy) การทำ Needle aspiration การทำ Skin tests การทำ Ultrasound การทำ Computed tomography (CT Scan) การทำ Magnetic Resonance Imaging (MRI) เป็นต้น
4. ระยะเวลาที่ผู้ป่วยนอนในโรงพยาบาล การติดเชื้อในโรงพยาบาลต้องไม่ใช่การติดเชื้อที่ตรวจพบตั้งแต่แรกรับผู้ป่วยไว้ในโรงพยาบาลหรือผู้ป่วยกำลังอยู่ในระยะฟักตัวของโรค ผู้ป่วยอาจปรากฏอาการของโรคติดเชื้อในโรงพยาบาลหลังจากกลับออกจากโรง พยาบาล คือ ไปปรากฏอาการที่บ้านก็ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 วิทยาการระบาดของการติดเชื้อในโรงพยาบาล

องค์ประกอบของการติดเชื้อในโรงพยาบาลในผู้ป่วยสามารถอธิบายตามหลักวิทยาการระบาดซึ่งประกอบด้วยผู้ป่วย,เชื้อโรคและสิ่งแวดล้อมโดยมีหนทางการแพร่กระจายเชื้อเป็นตัวเชื่อมโยงการติดเชื้อในโรงพยาบาลเป็นปัญหาที่สำคัญที่พบว่าทำให้เพิ่มต้นทุนในด้านการให้บริการและมีผลกระทบต่อกระบวณการให้บริการที่มุ่งเน้นในด้านความพึงพอใจของผู้รับบริการตลอดจนมีผลกระทบต่อขบวณการคุณภาพของโรงพยาบาลในด้านการควบคุมการติดเชื้อในโรงพยาบาลที่กำหนดเป็นตัวชี้วัดไว้ในการควบคุมการติดเชื้อทั้งในประเทศและต่างประเทศมีการกำหนดปัจจัยต่างๆ ที่สำคัญที่มีผลต่อการติดเชื้อไว้โดยปัจจัยที่มีผลต่อการติดเชื้อที่ทำการศึกษาแบ่งได้ 3 ประเภท คือ

1. ผู้ป่วย จะพบว่าผู้ป่วยที่รับการรักษาในโรงพยาบาลพบว่ามีทั้งผู้ป่วยศัลยกรรม อายุรกรรม ผู้ป่วยเด็กซึ่งส่วนมากจะมีโรคเดิมอยู่หรือมีภาวะที่ทำให้ผู้ป่วยมีภูมิต้านทานต่อเชื้อโรคลดลงโดยจะพบความแตกต่างตามลักษณะของผู้ป่วยดังต่อไปนี้

   1.1 ผู้ป่วย ระยะวิกฤตผู้ป่วยเหล่านี้มักจะได้รับการสอดใส่อุปกรณ์หรือเครื่องมือต่างๆ เข้าสู่ร่างกาย รวมทั้งมักมีการคาสายสวนต่างๆเช่นสายสวนปัสสาวะ สายสวนหัวใจ สายให้อาหารเป็นต้นซึ่งจะทำให้เกิดความเสี่ยงต่อการติดเชื้อในโรงพยาบาลเช่น คราเวนและคณะ (Craven et al.,1988) ทำการศึกษาตัวแปรที่มีผลต่อการติดเชื้อในโรงพยาบาลของหอผู้ป่วยหนัก (ICU) จะพบว่าการสวนคาอุปกรณ์ต่างๆในผู้ป่วยจะเพิ่มความเสี่ยงต่อการติดเชื้อในโรงพยาบาล

   1.2 อายุ การติดเชื้อในโรงพยาบาลพบว่าในผู้สูงอายุในสหรัฐอเมริกามีการติดเชื้อในโรงพยาบาลสูงกว่าผู้ป่วยในช่วงอายุอื่นๆเนื่องจากภูมิต้านทานของร่างกายต่ำ

   1.3 เพศ การติดเชื้อในโรงพยาบาลจากการศึกษาได้พบว่าระบาดวิทยาการติดเชื้อในระบบทางเดินปัสสาวะจะพบในเพศหญิงสูงกว่าเพศชาย แต่สำหรับการวิจัยการติดเชื้อในผู้ป่วยของ (Richter, Hottiroff&Nissenkorn, 1991) พบว่าผู้ชายที่ต่อมลูกหมากโตจะเกิดภาวะการติดเชื้อระบบทางเดินปัสสาวะสูงกว่าเพศหญิง

   1.4 โรคประจำตัว ในผู้ป่วยที่มีโรคประจำตัวเช่นโรคเบาหวาน จะพบว่ามีการติดเชื้อได้ง่ายกว่าผู้ป่วยที่ไม่ได้เป็นโรคเบาหวาน โดยมีอัตราการติดเชื้อสูงกว่าถึง 2.5เท่า

2. สิ่งแวดล้อม ที่มีผลต่อการติดเชื้อในโรงพยาบาลประกอบไปด้วยสิ่งแวดล้อมที่ไม่มีชีวิตเช่น อุปกรณ์เครื่องมือเครื่องใช้พบว่าผู้ป่วยที่ได้รับการรักษาโดยการสอดใส่อุปกรณ์พิเศษมีโอกาสเกิดการติดเชื้อได้มากกว่า เจ้าหน้าที่บุคลากรประจำหน่วยงานนั้นๆ สถานที่ที่ผู้ป่วยพักรักษาตัวถ้ามี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การออกแบบไม่ดีจะทำให้สกปรกและมีเชื้อโรคปะปนได้รวมทั้งสถานที่ที่มีการรับผู้ป่วยแออัดเกินไป

3. เชื้อโรค เชื้อโรคที่เป็นสาเหตุของการติดเชื้อในโรงพยาบาลพบว่ามีทั้งแกรมลบและแกรมบวก ซึ่งเป็นเชื้อแบคทีเรีย นอกจากนี้ยังมีการติดเชื้อรา เชื้อไวรัสและอื่นๆผลกระทบจากการติดเชื้อในโรงพยาบาลแต่ละระบบทำให้ผู้รับบริการต้องอยู่ในโรงพยาบาลนานขึ้น เสียค่าใช้จ่ายเพิ่มขึ้น หรือเพิ่มความรุนแรงของการเจ็บป่วยจนกระทั่งทำให้เกิดอัตราตายได้จากการติดเชื้อในระบบนั้นๆหรือระบบอื่นๆตามมา

2.4 ปัจจัยส่งเสริมการติดเชื้อในโรงพยาบาล

1. ระยะเวลาในการนอนพักรักษาตัวในโรงพยาบาล ผู้ป่วยเรื้อรังที่อยู่โรงพยาบาลนานจะทำให้ภูมิต้านทานในโรงพยาบาลต่อการติดเชื้อของร่างกายลดลงทำให้เชื้อแบคทีเรียโดยเฉพาะเชื้อแบคทีเรียแกรมลบ
2. การได้รับยาต้านจุลชีพที่ออกฤทธิ์กว้างเป็นระยะเวลานาน ซึ่งจะมีผลต่อการทำให้มีการเปลี่ยนแปลงของเชื้อโรคทำให้เชื้อก่อโรคมีความสามารถในการก่อโรคสูงขึ้น
3. การตรวจรักษาโดยการใส่อุปกรณ์เครื่องมือทางการแพทย์เนื่องจากอุปกรณ์เครื่องมือเหล่านี้มักก่อให้เกิดการบาดเจ็บของเนื้อเยื่อไม่มากก็น้อยซึ่งจะทำให้เชื้อโรคสามารถเกาะติดกับเนื้อเยื่อที่บาดเจ็บเหล่านั้นได้ง่าย รวมทั้งกรณีที่อุปกรณ์เหล่านั้นมีการปนเปื้อนก็จะนำเชื้อโรคเข้าสู่ร่างกาย

2.5 ทำไมต้องใช้ Data Mining

ในสหรัฐอเมริกาได้มีการพัฒนาโปรแกรมการควบคุมการติดเชื้อในโรงพยาบาลต่างๆตั้งแต่ปีคศ1960 ซึ่งเนื่องมาจากการได้รับการกระตุ้นจาก The Joint Commission on Accreditation of Hospital:TCAH, The American Hospital Association:AHA และศูนย์ควบคุมโรคสหรัฐอเริกาดำเนินการเฝ้าระวัง พัฒนามาตรการการควบคุมการติดเชื้อโดยได้ทำการศึกษาโรคติดเชื้อในโรงพยาบาลขึ้นมามีการศึกษาวิจัยปัจจัยต่างๆที่เข้ามามีผลต่อการติดเชื้อของผู้ป่วย ซึ่งมีสาเหตุจากปัจจัยหลายๆอย่าง ศูนย์ควบคุมโรคสหรัฐอเมริกาพบว่าโรงพยาบาลจะลดอัตราการติดเชื้อในโรงพยาบาลได้ประมาณร้อยละ 32 ถ้าโปรแกรมการเฝ้าระวังและควบคุมการติดเชื้อมีองค์ประกอบครบถ้วน องค์ประกอบที่สำคัญมี 5 อย่างด้วยกันคือ (Soule,1983; Haley et al.,1985)

1. มีการเฝ้าระวังการติดเชื้อในโรงพยาบาล (Survillance)
2. มีกิจกรรมป้องกันและควบคุมการติดเชื้อในโรงพยาบาล (Control Measure)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. มีแพทย์ควบคุมโรคติดเชื้อ (Infection Control Physician)

4. มีพยาบาลควบคุมโรคติดเชื้อ (Infection Control Nurse)

5. มีระบบการรายงานอัตราการติดเชื้อแก่บุคลากรที่เกี่ยวข้อง

จากหลักการดังกล่าวข้างต้นนี้จึงมีแนวคิดในการนำระบบสารสนเทศที่ใช้ในการวิเคราะห์ข้อมูล เข้ามาวิเคราะห์ปัจจัยต่างๆเพื่อกำหนดแนวทางในการทำกิจกรรมป้องกันและควบคุมการติดเชื้อในโรงพยาบาลและค้นหาลักษณะของปัจจัยที่ไม่เคยรู้มาก่อนหน้านี้ รวมถึงความสัมพันธ์ การพึ่งพา หรือแนวโน้มอื่น ๆ เข้ามาช่วยในการวิเคราะห์เพื่อสร้างระบบรายงานการติดเชื้อแก่บุคลากรที่เกี่ยวข้อง จากนั้น การค้นหาความสัมพันธ์เหล่านี้จะกลายเป็นส่วนหนึ่งของระบบซึ่งจะนำไปสู่การตัดสินใจ

ในปัจจุบันในหน่วยงานของโรงพยาบาลมีการจัดการการออกแบบระบบฐานข้อมูล (Database) เพื่อนำข้อมูลเหล่านั้นไปจัดทำรายงานเพื่อนำเสนอต่อผู้บริหาร ในการนำเสนอรายงานทางด้านการติดเชื้อในโรงพยาบาลมีความจำเป็นที่จะต้องศึกษาถึงปัจจัยที่มีผลต่อการติดเชื้อในผู้ป่วย มีการนำเครื่องมือต่างๆเข้ามาช่วยในการวิเคราะห์เช่น ระบบช่วยในการตัดสินใจ (Decision Support Systems)โดยมีการเชื่อมโยงเข้ากับระบบคลังข้อมูล (Data Warehouse) มีการระบบเหมืองข้อมูลเข้าใช้ในระบบทางด้านการแพทย์

เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลที่เป็นต้นแบบนี้ขึ้นอยู่กับผู้ใช้ (End users) ที่จะกำหนดขอบเขตของปัญหารวมถึงการเลือกปัจจัย และการเริ่มต้นวิเคราะห์ปัจจัยที่เหมาะสมเพื่อที่จะก่อให้เกิดขบวณการความรู้ที่ช่วยในการสร้างแบบ (Model) และแก้ปัญหาที่ระบบเปิดเผยให้ทราบถ้าผู้ใช้ประสบความล้มเหลวในการรู้ปัญหาที่แท้จริงก็จะไม่มีปฏิกิริยาใดๆเกิดขึ้น ปัจจุบันนี้มีระบบสารสนเทศที่ช่วยในการตัดสินใจ (Decision Support Systems) บางตัวที่สนับสนุนรูปแบบที่เน้นในการช่วยตัดสินใจอย่างหลากหลาย ในระบบเหล่านี้จะมีการสร้างระบบสัญญาณเตือนภัยแบบอัตโนมัติ สัญญาณเตือนภัยเหล่านี้คือตัวแทนของ Software ที่ควบคุม Parameter ตัวนั้น ๆ อย่างสม่ำเสมอ เช่นการพิจารณาเรื่องความดันโลหิตสูงโดยมี Parameter ต่าง ๆเช่น เพศ อายุ การสูบบุหรี่ การรับประทานอาหาร การออกกำลังกาย การพิจารณาการติดเชื้อโดยค้นหาความสัมพันธ์ต่างๆเช่น แผนกผู้ป่วย อายุของผู้ป่วย เพศ การได้รับยาปฏิชีวนะ

สำหรับระบบสารสนเทศที่พิจารณานำมาใช้ในการวิเคราะห์ปัจจัยการควบคุมและป้องกันการติดเชื้อในโรงพยาบาลนี้ ได้พิจารณาเครื่องมือที่มีประสิทธิภาพสูงที่ สามารถทำการพยากรณ์หรือทำนายแนวโน้มที่อาจเกิดขึ้นได้จากรูปแบบ (Pattern) ที่มีอยู่ซึ่งจะพิจารณาแตกต่างจากการใช้สถิติโดยทั่วไป เราเรียกระบบนี้ว่า Data Mining

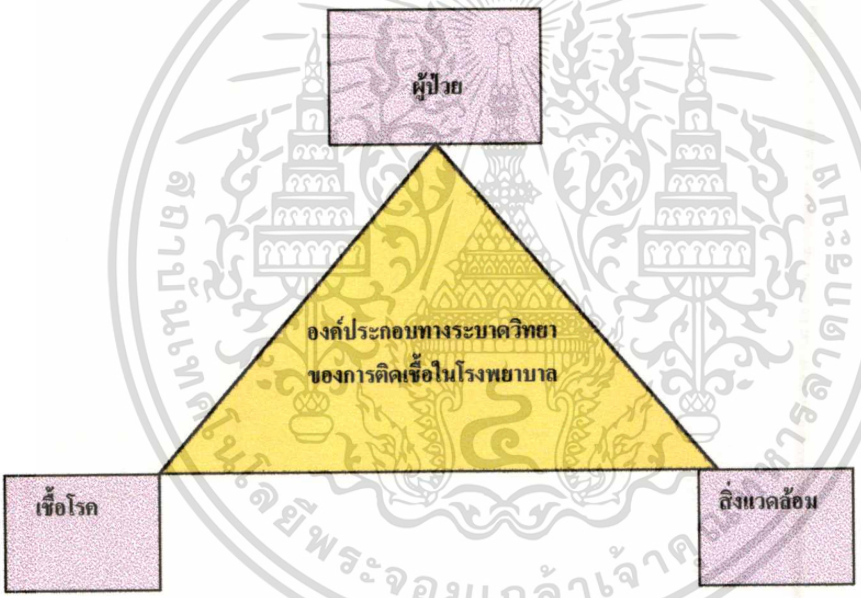
จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องพบว่า การเกิดการติดเชื้อในโรงพยาบาลต้องเกิดจากองค์ประกอบหลายอย่างด้วยได้แก่ ผู้ป่วยที่มีภูมิไวในการรับเชื้อ เชื้อโรค สิ่งแวดล้อม โดยมี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หนทางการแพร่กระจายเชื้อเป็นตัวเชื่อมโยงให้เกิดการติดเชื้อและวิธีหนึ่งที่สามารถทำได้โดยการลดปัจจัยเสี่ยงทางด้านสิ่งแวดล้อมคือให้บุคลากรตระหนักถึงความสำคัญและให้ข้อมูลที่จะเป็นประโยชน์แก่บุคลากร ให้เกิดความเข้าใจในปัจจัยเสี่ยงของสถานพยาบาลตนเอง และนำผลจากปัจจัยเสี่ยงของการติดเชื้อในโรงพยาบาลไปวางแผนเพื่อลดการติดเชื้อในอนาคต

2.6 กรอบแนวคิดที่ใช้ในการศึกษา

จาการศึกษาเอกสารที่เกี่ยวข้องกับการวิเคราะห์ระบบสารสนเทศเพื่อการควบคุมและป้องกันการติดเชื้อในโรงพยาบาลผู้ศึกษาได้นำมาสร้างกรอบแนวคิดในการทำการศึกษาครั้งนี้ตามภาพที่ 1



ภาพที่ 1 ภาพแสดงความสัมพันธ์ระหว่างปัจจัยทางระบาดวิทยาของการติดเชื้อในโรงพยาบาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

Data Mining และทฤษฎี

3.1 ความหมายของ Data Mining

Data Mining หมายถึง ขบวนการค้นหาความหมายความสัมพันธ์ใหม่ ๆ โดยการเลือกผ่านข้อมูลที่รวบรวมเก็บไว้โดยใช้รูปแบบของเทคโนโลยีเข้ามาช่วยในการคำนวน ทางสถิติและทางคณิตศาสตร์ หลักการในการทำ Data Mining คือการวิเคราะห์และตรวจสอบข้อมูลเปิดเผยให้เห็นถึงปัญหา หรือ โอกาสที่ซ่อนอยู่ในข้อมูล (Data) สร้างรูปแบบความสัมพันธ์ของข้อมูลและจากนั้นก็จะใช้ Model ตัวนั้นในการพยากรณ์แนวโน้ม โดยที่ End users จะเข้ามามีส่วนแทรกแซงน้อยที่สุด ดังนั้น End users สามารถที่จะใช้หาข้อมูลที่มีประโยชน์ในการวิเคราะห์เพื่อให้เกิดประสิทธิภาพในการควบคุมโรคติดเชื้อในโรงพยาบาลได้ Data Mining จะอธิบายถึงเครื่องมือชนิดใหม่ที่สนับสนุนการตัดสินใจเฉพาะด้าน ซึ่งควบคุมการวิเคราะห์ข้อมูล พูดง่าย ๆ ก็คือ Data Mining Tool นอกจากนี้เป็นวิธีที่ถูกออกแบบเพื่อก่อให้เกิดความรวดเร็วในการค้นพบความรู้ (Knowledge-Discovery) ที่เหนือกว่าข้อมูลของ Database และ มีการแทรกแซงของ End users น้อยที่สุดระหว่างช่วงของการค้นพบความรู้จริงมี 3 ลักษณะคือ

1.ข้อมูล (Data) ข้อมูลจะก่อรูปร่างขึ้นมาก่อนแสดงให้เห็นว่าอะไรคือองค์ประกอบที่มีมากที่สุดในการเก็บรวบรวมข้อมูลไว้ในฐานะข้อมูลที่เกี่ยวกับการปฏิบัติ

2.ข้อมูลที่ผ่านกระบวนการ (Information) ประกอบด้วยข้อมูลที่แสดงให้เห็นถึงข้อมูลที่ผ่านกระบวนการพิเศษ ข่าวสารนี้จะสร้างรากฐานสำหรับการตัดสินใจและความเข้าใจในทางธุรกิจ

3.ความรู้ (Knowledge) คือความรู้ที่ค้นพบได้และแสดงให้เห็นถึงข่าวสารหรือข้อมูลที่เฉพาะเจาะจงอย่างยิ่ง คุณลักษณะของ Data Mining Tool ในปัจจุบันประกอบด้วยการออกแบบมากมายและหลากหลายของการประยุกต์ใช้ให้เหมาะกับความต้องการของ Data Mining ก็มีมากมายเช่นกัน แต่ความผันผวนมากมายยังคงมีอยู่ เนื่องจากไม่มีการตั้งมาตราฐานขึ้นมาควบคุมดูแล การสร้าง Data Mining Tool จะถูกควบคุมดูแลโดยศูนย์รวมและการเข้าถึงที่แตกต่างกันออกไป ดังนั้น การสร้างประเภทของ Data Mining Tool จึงมุ่งประเด็นไปที่สภาวะตลาด เช่น การตลาด การขายปลีก การเงิน การดูแลสุขภาพ การลงทุน การประกันภัย การธนาคารและอื่นๆ Data Mining Toolสามารถใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการคำนวนได้อย่างแน่นนอน และการคำนวนเหล่านั้นสามารถถูกนำมาดำเนินให้เป็นไปตามแผนได้ในวิธีที่แตกต่างกันและหรือประยุกต์ใช้กับข้อมูลที่แตกต่างกันได้ทั่วไป

3.2 ชนิดของ Data Mining

การทำ Data Mining สามารถแบ่งตามวิธีการทำ Data Mining ได้ 3 ชนิดด้วยกันคือ

1. การศึกษาโดยการแบ่งประเภท Classification

กรณีศึกษานี้จะสร้างความเข้ามาใน Classification Study นักวิเคราะห์จะต้องการเข้าใจว่าข้อมูลบางตัวมีผลต่อการวิเคราะห์และนักวิเคราะห์จะต้องทำนายลักษณะของข้อมูลที่มีผลต่อการศึกษานั้นๆเนื่องจากวิธีนี้นักวิเคราะห์มีเป้าหมายในใจแล้วดังนั้นนักวิเคราะห์จึงสามารถสร้างโมเดล (Model) ที่ข้อมูลต่างๆได้มาจากข้อมูลในอดีตที่มีผลต่อการศึกษานั้นๆ โมเดลที่สมบูรณ์ถูกต้องจะสามารถทำให้เข้าใจและทำนายลักษณะผลที่จะเกิดขึ้นจากการศึกษาได้ จากตัวอย่างเหตุการณ์ที่ศึกษาจะสามารถอธิบายขั้นตอนของการกำหนดได้ การศึกษาจะกำหนดของเขตของกิจกรรมของ Data Mining ได้ นอกจากนี้การศึกษาจะสามารถกำหนดจุดประสงค์และข้อมูลที่ต้องการใช้ได้ทั้งหมด ด้วยการกำหนดปัญหานั่นก็เป็นสิ่งที่บอกให้นักวิเคราะห์ทราบได้เลยว่าขั้นตอนในการทำ Data Mining จะทำอย่างไรและจุดประสงค์ของการทำคืออะไรจากตัวอย่างข้างต้นเรากำหนดให้จากขบวนการศึกษาวิธีนี้ตัวแปรที่ได้จะเป็นตัวแปรไม่อิสระชนิด Dependent Variable ใช้เป็นพื้นฐานในการศึกษาว่าอะไรคือสาเหตุที่ทำให้เกิดผลต่อการศึกษาจะใช้ Data Element ตัวอื่นๆ มาช่วยในการอธิบายสิ่งที่เกิดขึ้น ถ้าเราเปลี่ยน Data Element ตัวอื่น จุดประสงค์ของการศึกษาก็จะเปลี่ยนไปด้วย

2. การศึกษาโดยการรวมกลุ่ม แบบ Clustering

การศึกษาโดยการรวมกลุ่มแบบ Clustering มีรูปแบบลักษณะที่เหมือนกันการศึกษาของ Clustering ไม่มีผลลัพธ์หรือตัวแปรไม่อิสระแบบ Classification การทำ Clustering เองสามารถทำบนพื้นฐานของข้อมูลในอดีตได้เหมือนกันแต่ผลลัพธ์ที่ได้มาไม่ได้ออกจาก Training Data เทคนิคของ Clustering พยายามมองหาความเหมือนและความแตกต่างภายใน กลุ่มของข้อมูลและแบ่งกลุ่มต่างๆ ออกเป็นส่วน ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. การศึกษาโดยการนำเสนอรูปแบบกราฟฟิค Visualization

Visualization เป็นการนำเสนอข้อมูลในรูปแบบของกราฟฟิค การนำเสนอจะสามารถทำได้มากกว่า 2 มิติซึ่งจะสร้างความละเอียดของการนำเสนอและสร้างความเข้าใจให้มากขึ้น

3.3 Data Warehouse และData Mining

Data Mining จะแตกต่างจากวิธีเดิม ๆ ที่ใช้ในการวิเคราะห์ข้อมูล เช่น Query language,SQL และ Online Analytical processing (OLAP),Multi-dimentional Dimension Analysis เพราะเป็นการค้นหาข้อมูลโดยปราศจากการตั้งสมมุติฐานไว้ล่วงหน้า ในงานของ Database นักวิเคราะห์ข้อมูลและนักสถิติ มักจะมีคำถามว่าในเมื่อ Data Mining ถูกใช้รวมกับเทคนิควิเคราะห์ข้อมูลแบบเดิมๆ มากมาย แต่จะมีความหมายแตกต่างกันเนื่องจาก SQL Query เป็นคำถามพื้น ๆ เช่น อาจจะกำหนดจำนวนของที่ขายในร้านใน Location หนึ่งในเดือนที่ผ่านมาว่าเป็นเท่าใด
การจุดกราฟแสดงยอดขายของแต่ละสถานที่ในระยะเวลาหนึ่งเรียกว่า Visualizational graphic
การวิเคราะห์ความอ่อนไหวของสินค้าหนึ่งๆต่อหน่วยในสถานที่หนึ่ง ๆเรียกว่า Statistic
เครื่องมือที่ช่วยทำให้การค้นหามีประสิทธิภาพมากขึ้นใช้เปรียบเทียบยอดขายของสินค้าชนิดหนึ่งหรือสินค้าชนิดต่างๆในแต่ละLocationในช่วงเวลาที่แตกต่างกันจากข้างต้นเรียกว่าMulti-dimentional analysis เป็นจะพบว่าเทคนิคทุกอันมีสิ่งที่เหมือนกันคือมีการกำหนดสมมุติฐานไว้แล้ว ผู้บริหารรู้ว่ามีจำนวนสินค้า จำนวนร้าน ยอดขายอยู่เท่าไร สามารถเช็คดูความสัมพันธ์ได้ การวิเคราะห์นี้เป็นการวิเคราะห์ โดยใช้เครื่องมือประเภทค้นหา (Query Tool) ซึ่งจะยอมใช้คำถามกับ DBMS โดยให้ผู้ใช้หาความจริงใหม่ๆหรือหาความจริงที่น่าสนใจจากข้อมูลซึ่งเก็บไว้ในฐานข้อมูล สำหรับ Data Mining นั้นแตกต่างไปโดยจะใช้เทคนิคเดิมๆตอบคำถามซึ่งไม่ได้แทนที่จะทำการเชื่อมต่อระหว่างสถานที่และยอดขาย
Data Mining จะกำหนดปัจจัยสำคัญ ๆซึ่งเป็นสาเหตุที่ทำให้ยอดขายสูงหรือยอดขายต่ำในการศึกษาแบบนี้เราไม่สามารถรู้ได้ว่าอะไรเป็นปัจจัยที่ทำให้ยอดขายสูงหรือต่ำ Data Mining จะหาปัจจัยซึ่งมีอิทธิพลที่ทำให้มีผลต่อยอดขาย โดยใช้ขั้นตอนในการแก้ปัญหาหารูปแบบและแนวโน้มของข้อมูลนั้น โดยจะค้นพบความสัมพันธ์ และรูปแบบที่ถูกส่วนปิดบังไว้ทำให้ไม่เห็นภาพชัดเจน Data Warehouse คือการรวบรวมของข้อมูลเพื่อสนับสนุนการตัดสินใจของฝ่ายบริหารข้อมูลจะถูกแบ่งเป็นระดับๆหลายระดับเพื่อให้เกิดความสามารถในการเข้าถึงข้อมูลได้อย่างรวดเร็วจุดประสงค์ของ Data Warehouse คือช่วยปรับปรุงประสิทธิภาพในการตัดสินใจที่เกี่ยวกับธุรกิจที่ต้องเกี่ยวข้องกับตัวเลขจำนวนมาก โดยมีพื้นฐานอยู่บนหลักของข้อมูลที่ใช้จัดการองค์กร ซึ่งเป็นข้อมูลสรุปเพื่อการตัดสินใจ แทนที่จะเป็นข้อมูลที่ใช้ดำเนินกิจกรรมประจำองค์กรData Warehouse เป็นขั้นตอนก่อนที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะทำ Data Mining โดยการจัดขนาดของข้อมูลใหญ่ ๆ ให้อยู่ในรูปแบบที่ง่ายต่อการเข้าถึง หรือการเข้าไปถึงข้อมูล รูปแบบของ Data Warehouse มี 4 ลักษณะคือ

1. การรวมเข้าด้วยกัน (Integrate) มีลักษณะของ Data Warehouse ที่เป็นศูนย์กลาง (Centralized) โดยรวมข้อมูลที่ได้มาจากองค์กร ทั้งภายในและภายนอกซึ่งมาจากแหล่งข้อมูลหลายแหล่ง ที่มีความแตกต่างกัน การรวมกันสามารถใช้เวลา และทำให้ผู้บริหารเกิดความเข้าใจในการดำเนินการของธุรกิจได้ดีและคาดการโอกาศทางธุรกิจ
2. การจัดรูปแบบเชิงวัตถุ (Subject-Oriented) ในคลังความรู้จะถูกจัดการและเปรียบเทียบวิธีการหาคำตอบจากแหล่งข้อมูลที่แตกต่างกันข้อมูลจะถูกกำหนดในลักษณะเชิงวัตถุและหลากหลาย เช่น ข้อมูลการขาย ชนิดของสินค้า ประเภทของลูกค้า เป็นต้นสำหรับแต่ละข้อใน Data Warehouse บรรจุข้อมูลเฉพาะที่น่าสนใจ เช่น ผลิตภัณฑ์ ลูกค้า แผนก สถานที่ การโฆษณาอื่น ๆแบบฟอร์มของข้อมูลจะแตกต่างจากหน้าที่หรือขบวนการของระบบ Transaction เฉพาะ
3. การเปลี่ยนแปลงตามเวลา (Time-Variant) ข้อมูลใน Data Warehouse มีลักษณะระบุเวลาที่ข้อมูลถูกใส่เข้ามาหรือเป็นข้อมูลสรุปดังนั้นจะเป็นการบันทึกข้อมูลในลักษณะต่อเนื่องจะมีประวัติสามารถวิเคราะห์แนวโน้มที่เป็นไปได้ ข้อมูลนี้จะชี้ Transaction Data Warehouse แสดงการไหลของข้อมูลและบรรจุข้อมูลซึ่งแสดงผลเกิดในอาทิตย์ที่ผ่านมา เดือนที่แล้ว หรือ 5 ปีที่แล้ว เป็นต้นนอกจากนี้สามารถบรรจุข้อมูลรูปแบบทางสถิติเช่นเมื่อข้อมูลการขายในอาทิตย์ที่ผ่านมาจะเข้าไปอยู่ใน Data Warehouse และในช่วงเวลานั้นจะรวมข้อมูล ผลิตภัณฑ์ ลูกค้า คลังพัสดุ และอื่นๆ เข้าด้วยกัน
4. ไม่ลืมเลือน (Non-Volatile) ข้อมูลที่ใส่ไว้ใน Data Warehouse จะไม่ถูกลบหรือเปลี่ยนแปลง เพราะข้อมูลในคลังความรู้แสดงประวัติของบริษัท ข้อมูลในการปฏิบัติการแสดงประวัติในอดีต

ข้อมูลจะไม่ถูกลบและจะมีข้อมูลใหม่ที่ถูกใส่เข้าไปใน Data Warehouse ทำให้ Data Warehouse มีขนาดใหญ่ขึ้นดังนั้นจึงเป็นแหล่งข้อมูลที่มีรายงานความถูกต้องและสามารถนำไปใช้วิเคราะห์เชิงเปรียบเทียบมีเครื่องมือสองแบบที่ใช้ในการจัดการทำ Data Warehouse คือการใช้การแปลงข้อมูล (Data transformation) การตรวจสอบความถูกต้องของข้อมูล (cleaning) ซึ่งทำให้มั่นใจว่า Data Warehouse มีความถูกต้องและแม่นยำ มีประสิทธิภาพและมีต้นทุนในการจัดทำต่ำขั้นตอนของการทำ Data Warehouse คร่าว ๆ จะเริ่มจากขั้นตอนต่อไปนี้

การสังเคราะห์เอาข้อมูลที่เป็นประโยชน์สำหรับการทำ Data Mining เรียกว่า Data Extraction

การกำหนดขนาดของข้อมูลโดยการคัดเลือกตัวอย่างเรียกว่า Sampling and Selecting

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การรวบรวมข้อมูลที่เกี่ยวข้องไว้ด้วยกันเรียกว่า Aggregation
การตรวจสอบข้อมูลเป็นการสร้างความมั่นใจว่าข้อมูลจะสมบูรณ์และลดความซ้ำซ้อนของข้อมูล เรียกว่า Data Cleaning
การนอร์มัลไลซ์ (Normalization) เป็นกระบวนลดความซ้ำซ้อนของข้อมูล
กระบวนการช่วยทำให้สามารถเข้าหาข้อมูล (Data Access) ได้เร็วขึ้นเรียกว่า Overlay

3.4 Star Schemas

Star Schemas เป็นเทคนิคในการออกแบบที่ใช้ใน Multidimensional Decision Support Data ในฐานข้อมูลแบบความสัมพันธ์ (Relational) เหตุผลในการพัฒนา Star Schemas เนื่องจากเทคนิคในการออกแบบ E-Rdiagram และการนอร์มัลไลซ์ไม่ช่วยในการวิเคราะห์ข้อมูลได้ดีเท่าที่ควร Star Schema จะเป็นรูปแบบง่ายในการติดตั้งสำหรับการวิเคราะห์ข้อมูลหลายมิติ Multidimensional และรักษาโครงสร้างความสัมพันธ์ซึ่งสร้างระบบ Operational database พื้นฐาน ของ Star Schemas มีส่วนประกอบ 4 ส่วนคือ

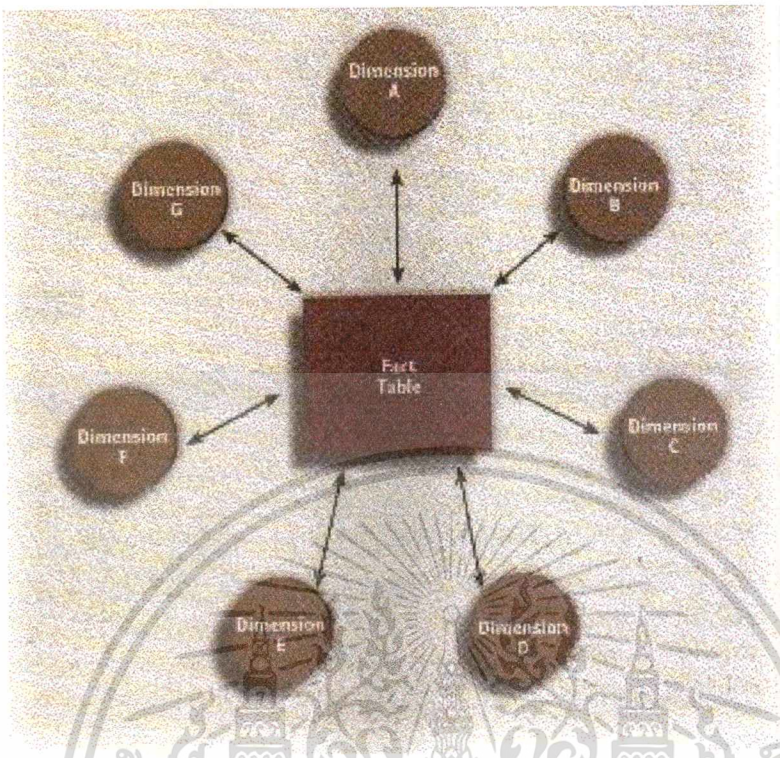
1. ค่าความจริง (Fact)
2. มิติแห่งความสัมพันธ์ (Dimension)
3. ข้อมูลที่แสดงลักษณะและคุณสมบัติ (Attributes)
4. การแตกสาขาของลักษณะและคุณสมบัติข้อมูล (Attribute hierarchies)

1. ค่าความจริง (Fact) เป็นตัววัดค่าในลักษณะจำนวนซึ่งจะแสดงค่าของตัวเลขธุรกิจเฉพาะกิจกรรมนั้นๆตัวอย่างเช่น ตัวเลขการขายเป็นตัววัดจำนวนการขายผลิตภัณฑ์ ค่าความจริงที่ใช้ในการวิเคราะห์ข้อมูลทางธุรกิจเป็นหน่วย (Unit) เช่นราคา รายได้ ค่าความจริงจะบรรจุตารางที่เป็นศูนย์กลางของ Star Schemas ตารางความจริงบรรจุความจริงและถูกเชื่อมในมิติต่างๆค่าความจริงสามารถคำนวนหรือค้นหาได้ทันเวลาการคำนวนหรือการสืบค้นความจริงเรียกว่าเครื่องวัด (Metric) เพื่อแยกความแตกต่างข้อมูลที่จัดเก็บ ตารางความจริงจะทำให้ทันสมัย ทุกวัน ทุกอาทิตย์ ทุกเดือน ด้วยข้อมูลจากฐานข้อมูลของระบบปฏิบัติการ

2. มิติความสัมพันธ์ (Dimension) เป็นคุณลักษณะที่มีคุณภาพจัดใส่ตารางโดยเริ่มต้นจากมิติที่สนใจ เพราะข้อมูล DSS จะมีความสัมพันธ์กับข้อมูลอื่นๆใน View เดียวกัน เช่น การขายอาจจะเปรียบเทียบผลิตภัณฑ์ในสถานที่หนึ่งกับอีกสถานที่หนึ่งและจากเหตุการณ์หนึ่งไปยังเหตุการณ์ข้างหน้า รูปแบบของปัญหาจะถูกจัดทำโดยระบบการตัดสินใจ (DSS) เช่นเปรียบเทียบจำนวนการขายสินค้าในสถานหนึ่งสำหรับไตรมาสแรกในปี 1990 ถึง 1997 ตัวอย่าง การขายมีผลิตภัณฑ์สถานที่ และเวลาในมิติ แต่ละมิติจะถูกเก็บไว้ในตารางมิติความสัมพันธ์ ตามภาพที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.



ภาพที่ 2 แสดง STAR SCHEMAS ของการขายสินค้า (Peter Rob Carlos Coronel, 1997)

3. ข้อมูลที่แสดงลักษณะและคุณสมบัติ (Attribute) แต่ละตารางมิติบรรจุ Attribute ใช้ในการค้นหา การกรองข้อมูล (Filter) หรือแบ่งประเภทของความจริง มิติจะจัดหาความหมายของคุณลักษณะเดียวกับความจริง ผ่านมายัง Attribute ดังนั้นการออกแบบ Data warehouse ต้องแสดงAttribute ให้ชัดเจน ซึ่งจะใช้ในการวิเคราะห์ข้อมูล การค้นหา การทำกลุ่มหรืออธิบายมิติแห่งความสัมพันธ์ต่างๆ เช่นการขายสินค้า เช่น เราสามารถชี้ Attribute ที่เป็นไปได้ให้เกิดความชัดเจนในหลายมิติ เช่น

มิติผลิตภัณฑ์ เช่น รหัสสินค้า (Product-ID) ความหมายผลิตภัณฑ์อุตสาหกรรม

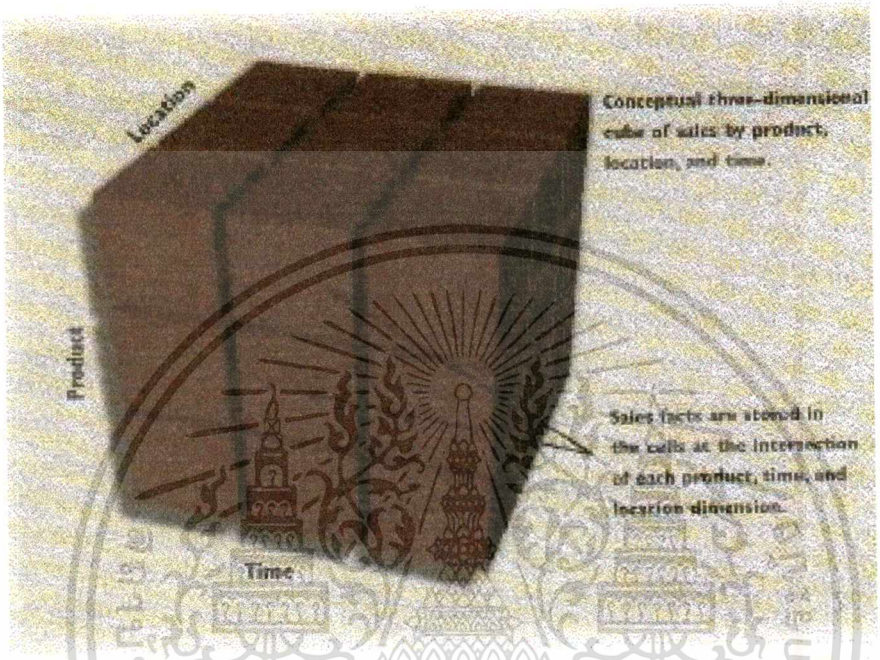
มิติสถานที่เช่น สถานที่ (Region) รัฐ (State) เมือง (City) และจำนวนที่จัดเก็บ

มิติเวลาเช่น ปี ไตรมาส เดือน อาทิตย์และวันที่

ผลิตภัณฑ์ สถานที่ และมิติเวลาแห่งจะแสดงรายละเอียดเกี่ยวกับการขาย การวิเคราะห์ข้อมูลสามารถหาความสัมพันธ์ของการขายกับผลิตภัณฑ์ สถานที่ และเวลา Star schemasสามารถจัดข้อมูลที่ต้องการหาความสัมพันธ์ผ่านมิติ โดยปราศจากการกำหนดขอบเขตของข้อมูลตัวอย่างเช่น จำนวนสั่งซื้อ จำนวนจัดซื้อ สถานะอื่น ๆ ในความเป็นจริงรูปแบบข้อมูลหลายมิติ เช่น ตัวอย่างการขายเป็นการแสดงที่ดีที่สุดโดยใช้รูป 3 มิติ เนื่องจากไม่มีหลักทางคณิตศาสตร์ที่จะใช้กำหนดจำนวนของมิติ ดังนั้นการใช้รูปแบบ 3 มิติจะทำได้ง่ายในการเห็นข้อมูลอย่างชัดเจนตัวอย่างของความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัมพันธ์ 3 มิติที่ใช้ในการวิเคราะห์ข้อมูลใน ภาพที่3 จะแสดงภาพของความสัมพันธ์ของมิติระหว่างการขายสัมพันธ์กับผลิตภัณฑ์ สถานที่และเวลา

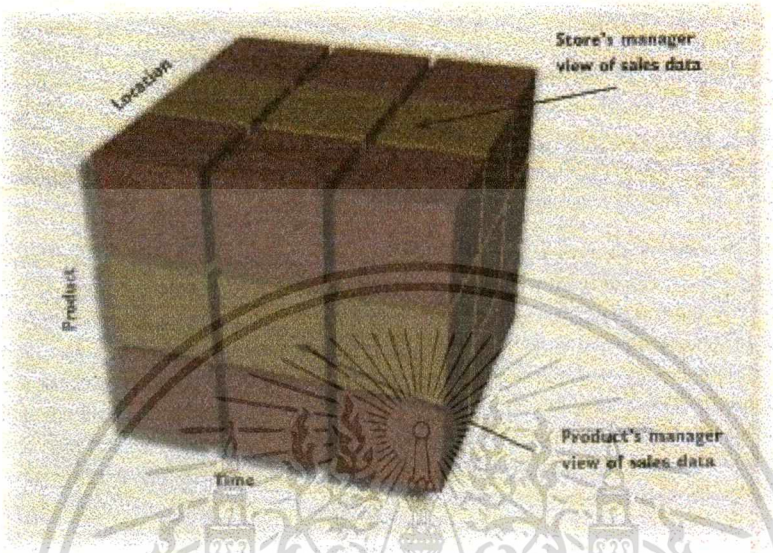


ภาพที่ 3 แสดงความสัมพันธ์ 3 มิติของข้อมูลการขายสินค้า (Peter Rob Carlos Coronel, 1997)

จากภาพจะสังเกตุว่าค่าของความสัมพันธ์ของการขายแต่ละอย่างถูกเก็บไว้ในสี่เหลี่ยมลูกบาศท์ซึ่งมีความสัมพันธ์กับสถานที่ ผลิตภัณฑ์และเวลาการค้นหาโดย ROLAP จะถูกใช้ในการวิเคราะห์ข้อมูลตรรกวิทยาและ End user สามารถนำมาวิเคราะห์ การวิเคราะห์ข้อมูลผ่านหลายมิติจะมีผลต่อระบบฐานข้อมูล การวัดความหนาแน่นของข้อมูลในสี่เหลี่ยมลูกบาศก์เรียกว่า Sparsity ซึ่งจะคำนวนโดยแบ่งจำนวนของค่าความจริงในลูกบาศก์สี่เหลี่ยมออกเป็นส่วนๆเมื่อมองเห็นสี่เหลี่ยมลูกบาศก์ไม่ใช่ทุกส่วนที่จะนำมาหาความสัมพันธ์ บางส่วนอาจจะไม่มีค่าเช่นมีผลิตภัณฑ์หลายอย่างที่ไม่ขายในระยะเวลาในสถานที่ที่ใช้ ในความจริงพบว่ามีมากกว่า 50% ของส่วนข้อมูลในสี่เหลี่ยมลูกบาศก์ที่จะนำมาใช้ ในกรณีของฐานข้อมูลหลายมิติจะใช้วิธีการของ Sparsity เพื่อลดขบวนการการใช้ทรัพยากรในการทำงาน (Overhead) ในการวิเคราะห์รูปแบบของมิตินี้จะเฉพาะเจาะจงลงไปที่ส่วน (Slice) ในสี่เหลี่ยมลูกบาศก์ตัวอย่างเช่น ผู้จัดการฝ่ายผลิตจะสนใจในการตรวจสอบ การขายผลิตภัณฑ์ขณะที่ผู้จัดการคลังสินค้าจะสนใจการจัดเก็บสินค้าที่ทำการขาย การใช้รูปแบบ 3 มิติที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไม่สามารถเข้าใจได้จะใช้วิธีการในการวิเคราะห์ โดยใช้ขบวนการหั่นเป็นสี่เหลี่ยมลูกบาศก์ออกเป็นลูกเต๋าส่วนๆเล็ก เรียกว่า Slice และ dice ตามภาพที่ 4

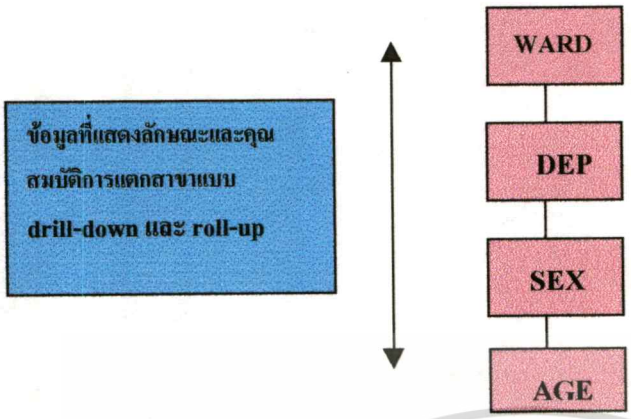


ภาพที่ 4 แสดงภาพของการ Slice และ diceหาความสัมพันธ์ของการขายสินค้า
(Peter Rob Carlos Coronel, 1997)

การหั่นสี่เหลี่ยมออกเป็นลูกเต๋าเล็กๆ (Slice and Dice) จะสามารถอธิบายส่วนต่างๆของสี่เหลี่ยมลูกบาศก์ โดยใช้ค่าของแต่ละ Attribute ในมิติที่ให้มา ตัวอย่างเช่น กำหนดเฉพาะไปที่ STORE-ID เพื่อที่จะมุ่งไปที่การจัดเก็บคลังสินค้าการกำหนดค่าของ Attribute โดยการ Slice และ Dice เช่น ภาพของการขายจะกำหนดขั้นตอนในการค้นหาวิธีการใหม่ ๆ การแบ่งประเภท การรวมข้อมูลที่เป็นไปได้ ความสำคัญของมิติต่างๆ คือเวลาซึ่งเวลาในมิติต่างๆนี้สามารถวิเคราะห์รูปแบบและทำนายได้ ดังนั้นเวลาในมิติจึงมีความสำคัญต่อบทบาทของนักวิเคราะห์ข้อมูลการให้ความสำคัญของเวลาในมิติต่างๆเพื่อวิเคราะห์ข้อมูลจะต้องมีการรวมข้อมูลในมิติแห่งเวลาใน Data Warehouse

4. การแตกสาขาของลักษณะและคุณสมบัติข้อมูล (Attribute hierarchies) มีการจัดลำดับขั้นของฟิลด์ในมิติต่าง ๆ โดยจัดข้อมูลแบบบนลงล่าง (Top-Down)โดยใช้ จุดประสงค์ 2อย่าง คือ การวิเคราะห์ข้อมูลแบบการรวมกัน (Aggregation) และการวิเคราะห์ข้อมูลแบบเจาะลึก (Drill-Down) หรือแบบ Roll-up การจัดลำดับขั้นแบบ Drill-down/Roll-up นี้ตามภาพที่ 5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 5
แสดงลำดับขั้นของฟิลด์ที่ทำการศึกษาปัจจัยที่เกี่ยวข้องกับการติดเชื้อในโรงพยาบาล

การนำเสนอ Star Schemas ค่าความจริง (Fact) และมิติแห่งความสัมพันธ์ (Dimension) ถูกแสดงโดยรูปแบบของตารางในฐานข้อมูลของ Data Warehouse ตารางความจริงมีความสัมพันธ์กับตารางในมิติแห่งความสัมพันธ์(ในรูปแบบความสัมพันธ์แบบ M:1 (many to one) ในอีกทางคือความจริงในแถวเดียวกันมีความสัมพันธ์กับมิติอื่นๆเช่น ตัวอย่างการขายจะสามารถสรุปได้ว่าในผลิตภัณฑ์ที่ปรากฏหลายครั้ง ๆ ในตารางความจริงของการขายได้

ตารางความจริงและมิติแห่งความสัมพันธ์มีความสัมพันธ์กับคีย์นอกและเป็นใจความหลักของ คีย์หลักและเงื่อนไขของคีย์นอก คีย์หลักจะอยู่ด้านหนึ่งของตารางมิติถูกเก็บเป็นส่วนหนึ่งของคีย์หลักในหลาย ๆ ด้าน ในตารางความจริง เพราะตารางความจริงจะมีความสัมพันธ์กับตารางหลาย ๆ มิติ คีย์หลักของตารางความเป็นจริง (Composite primary key)ในรูปแบบของ Star schema แต่ละ Record ของมิติ จะมีความสัมพันธ์กับจำนวน Record เป็นพัน ๆRecord คุณลักษณะของหน้าที่ในการค้นหาข้อมูลของ Star schema เพราะนักวิเคราะห์จะต้องมองความจริงผ่านฟิลด์ในมิติต่างๆ Data Warehouse มีตารางความจริงหลายตาราง และจะถูกออกแบบเพื่อตอบคำถามในการตัดสินใจเฉพาะ ตัวอย่างเช่น จะพัฒนาคำสั่งซื้อที่น่าสนใจใหม่ ๆ ขณะที่จะดำรงการขายดังเดิมที่น่าสนใจไว้ คำสั่งซื้อใดที่เป็นกุญแจสำคัญขององค์กรจะเป็นศูนย์กลางของ Star schema ซึ่งอาจจะเป็นบริษัทขายสินค้าผลิตภัณฑ์ระยะเวลาในกรณีนี้จะมีตารางฐานข้อมูลของบริษัทขายสินค้ามิติผลิตภัณฑ์จะถูกแสดงในตารางผลิตภัณฑ์ อย่างไรก็ตามการให้ความสนใจในการขายจะเฉพาะเจาะจงในช่วงเวลา ถ้าใช้เวลาที่แตกต่างกันของขอ้มูลจะต้องมีช่วงเวลาของตารางขึ้นมาเพื่อแสดงระยะเวลาของคำสั่งซื้อ เทคนิคในการออกแบบ Data Warehouse มี 4 เทคนิคคือ

1. การนอร์มัลไลซ์ตารางหลายมิติ (Normalization of dimensional tables)
2. ตารางความจริงนำเสนอการรวมความแตกต่าง (Multiple fact tables representing

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

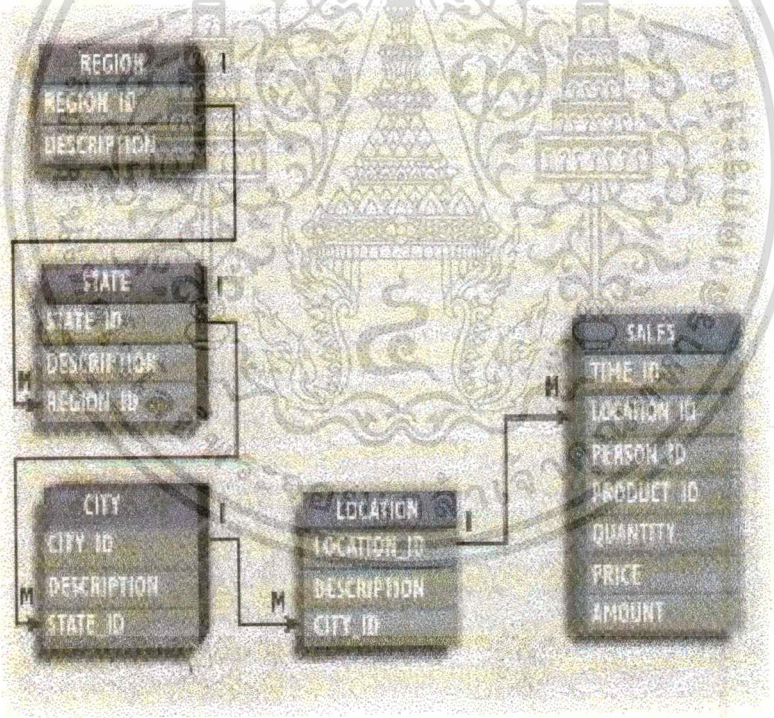
different aggregation levels)

3. ตารางความจริงของการดีนอร์มัลไลซ์ (Denormalization of fact tables)

4. ตารางแบ่งส่วนและการจำลอง (Table partitioning and replication)

1. ตารางนอร์มัลไลซ์หลายมิติ (Normalization of dimensional tables)

จะถูกสร้างแบบง่าย ๆ และยืดหยุ่นเพื่อให้ผู้ใช้ค้นหาผ่านมิติได้ตามภาพที่ 6 จะแสดงตาราง Location ซึ่งเป็นตารางมิติที่บรรจุ Transitive dependencies ผ่านสถานที่ รัฐและเมืองซึ่งสามารถสร้างความสัมพันธ์ เป็น 3 NF ได้ดังภาพที่ 6 การนอร์มัลไลซ์ตารางมิตินี้โดยการกรองข้อมูล (Filter) เพื่อสัมพันธ์กับมิติอื่น



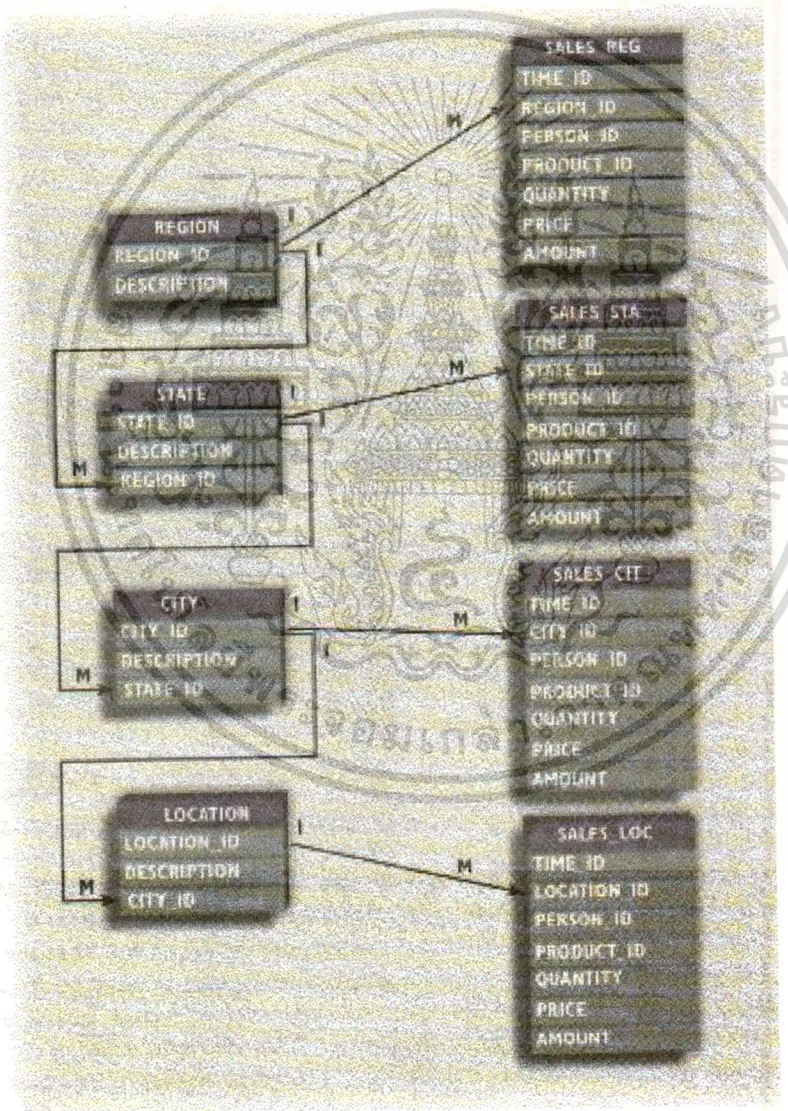
ภาพที่ 6 แสดงการนอร์มัลไลซ์ตารางมิติความสัมพันธ์

(Peter Rob Carlos Coronel, 1997)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ตารางความจริงนำเสนอการรวมความแตกต่าง (Multiple fact tables representing different aggregation levels)

วิธีนี้สามารถสร้างตารางที่มีความจริงหลายอย่างไว้ในตารางเดียวกันตามภาพที่ 7 จะแสดงตารางที่รวมกันนี้จะถูกคำนวนก่อนขั้นตอนการให้ข้อมูลจุดประสงค์ของเทคนิคนี้เพื่อรักษาเวลาในการดำเนินการช่วยให้เกิดความเร็วในการวิเคราะห์ข้อมูล ผู้ใช้เป็นผู้ปรับขบวนการในวิเคราะห์ โดยสรุปเป็นตารางความจริงแทนการหาตารางความจริงโดยการคำนวน



ภาพที่ 7 แสดงตาราง Multiple Fact Tables

(Peter Rob Carlos Coronel, 1997)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ตารางความจริงของดีนอร์มัลไลซ์ (Denormalization of fact tables)

ตารางนี้ทำให้ข้อมูลที่เข้าหาและเก็บรักษาข้อมูลในพื้นที่จัดเก็บดีขึ้นทำให้ค่าใช้จ่ายในการจัดเก็บข้อมูลลดลงและจำกัด ขนาดของตารางและฐานข้อมูลขนาดของ Record และจำนวน Record ที่สูงที่สุดในตารางเดียวกัน

4.การดีนอร์มัลไลซ์ (Table partitioning and replication)

การปรับปรุงประสิทธิภาพให้ดีขึ้นโดยการเก็บ Record เดียวจากหลายๆRecord เช่น ต้องการคำนวนการขายสินค้าชนิดหนึ่งในเขตต่างๆ โดยจะต้องเข้าไปดูข้อมูลการขายเฉพาะเขตนำมารวมและสรุปในตาราง ถ้ามีสินค้า 300,000 ชนิด จะต้องสรุปอย่างน้อย 300,000แถว ซึ่งเป็นการสิ้นเปลืองสำหรับการจัดทำ DBMS ดังนั้นเพื่อให้เกิดความเข้าใจง่ายและสามารถเปรียบเทียบข้อมูลในปีต่อปีในเดือนที่สนใจ การออกแบบโดยวิธีนี้จะเหมาะสมโดยกำหนดความถี่ของชนิดที่ต้องการไว้เพื่อการประเมิน เทคนิคที่นำมาใช้มีความสัมพันธ์กับ Data Warehouse มีวิธีดังนี้

1. การแยกตาราง (Partitioning) เป็นการแยกตารางออกเป็น Subset ของแถวหรือคอลัมภ์เพื่อช่วยให้การเข้าหาข้อมูลของลูกข่าย (Client) ทำได้ง่าย
2. การจำลอง (Replication) เป็นการจำลองตารางและใส่ข้อมูลใน Location ต่างๆเพื่อให้ง่ายต่อการเข้าถึงข้อมูล ซึ่งทำให้การวิเคราะห์ข้อมูลทางธุรกิจในมิติธรรมดามีประสิทธิภาพ
3. การกำหนดช่วงเวลา (Periodicity) จะแสดงข้อมูลปัจจุบันไว้ในตาราง หรือข้อมูลที่เกี่ยวข้องในช่วงเวลาของการเก็บข้อมูลในตาราง

3.5 เทคนิคของ Data Mining

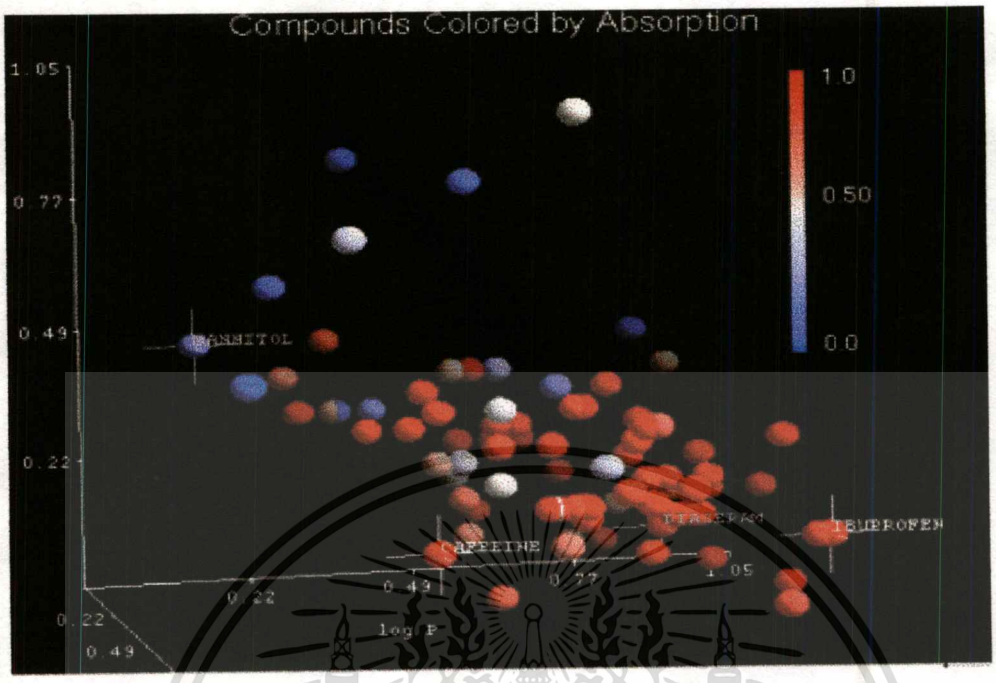
การสร้างรูปแบบ (Data Mining Model) เป็นการประมวลผลข้อมูลตาม Algorithm ที่ได้กำหนดไว้ในขั้นตอนนี้จะมีความสัมพันธ์กับการวิเคราะห์ผลข้อมูลและขั้นตอนที่ผ่านมา การพัฒนาที่แตกต่างกัน วิธีการคำนวนที่ใช้เป็นแบบอย่างมากที่สุดใน Data Mining นั้นมี 3 วิธีคือ

1. เครือข่ายระบบประสาท (Artificial Neural Networks)
2. การจำแนกพันธุ์กรรม (Genetic Algorithms)
3. การพิสูจน์กฎโครงสร้างต้นไม้ (Decision Trees)

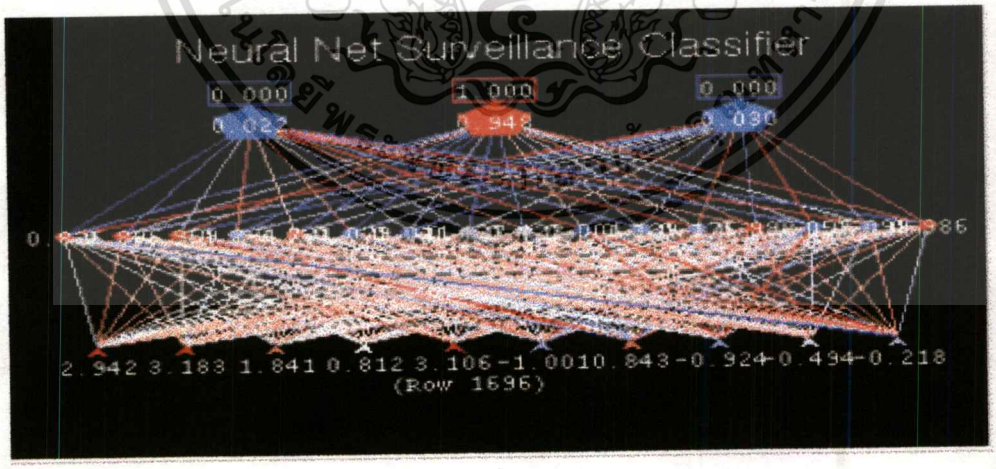
1.เครือข่ายระบบประสาท (Artificial Neural Networks)

การทำงานของระบบประสาท (Neural Networks)เป็นการทำงานเลียนแบบระบบการทำงานของระบบประสาทมนุษย์แต่ละ Process จะรับ Input เข้าไปคำนวนและสร้าง Output ออกมาในลักษณะที่ไม่ใช่การทำงานแบบเส้นตรงเพราะ ค่า Input ที่เข้าไปจะลำดับค่าความสำคัญไม่เท่ากัน ค่า Output ที่ได้จากการเชื่อมโยงกันนี้จะถูกนำมาเปรียบเทียบกับ Output ที่ตั้งไว้ ดังภาพที่ 8-9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

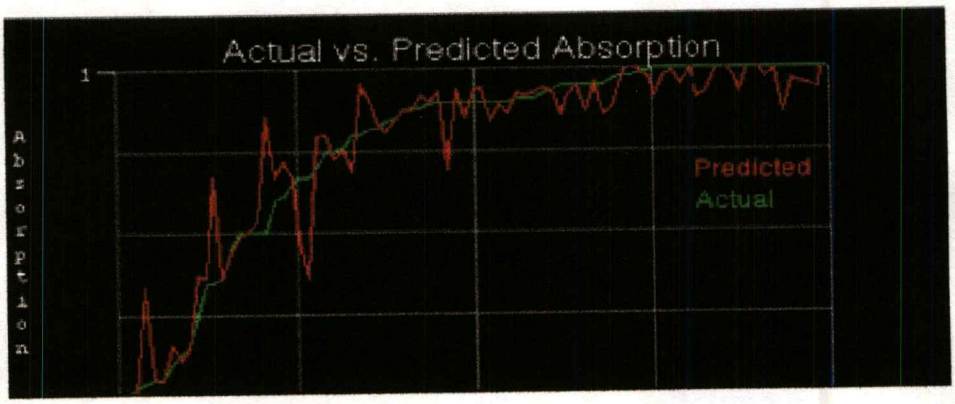


ภาพที่ 8 รูปแบบการวิเคราะห์ข้อมูลโดยใช้ขบวนการเครือข่ายระบบประสาท ในภาพเป็นการวิเคราะห์โมเลกุลของยา Manitol Caffein Diazepam (จาก www.patek.com)



ภาพที่ 9 แสดงรูปแบบเครือข่ายระบบประสาทวิเคราะห์โมเลกุลของยา Manitol Caffein Diazepam (จาก www.patek.com)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 10 ภาพแสดงความสัมพันธ์ของการศึกษาการดูดซึมของยาโดยมีการทำนายออกมาในลักษณะของกราฟเมื่อเปรียบเทียบกับค่าที่เป็นจริง (จาก www.patek.com)

2.การจำแนกพันธุ์กรรม (Genetic Algorithms)

เป็นความต้องการการสร้างพันธุ์กรรมที่ดีที่สุดบนขั้นตอนของวิวัฒนาการทางชีวภาพแนวความคิดนี้มีวิวัฒนาการในการพัฒนามาโดยตลอดมีความสามารถในการคัดเลือกสายพันธ์ที่ดีที่สุดได้และนำไปพัฒนาระบบการสร้างหรือจำลองพันธุ์กรรมที่มีคุณภาพมีการนำไปทดลองในพันธุ์พืชพบว่าได้พันธุ์ที่มีความสมบูรณ์ ขั้นตอนการทำงานแบบจำแนกพันธุ์กรรมนี้เริ่มจากการรวมกลุ่มข้อมูลหรือตัวแปรเข้าด้วยกันและมีความต้องการที่จะแบ่งกลุ่มหรือจับรวมกลุ่มของตัวแปร

ภาพที่ 11 แสดงรูปแบบของการจัดกลุ่มยีนส์ของเชื้อรา แบบ Cluster (จาก www.patek.com)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.การตัดสินใจโดยใช้ลักษณะของโครงสร้างแบบต้นไม้ (Decision Trees)

เป็นการสร้างโครงสร้างเพื่ออธิบายกลุ่มของข้อมูล ตัวแปรโดยมีการกำหนด Node ที่อยู่เหนือสุดของโครงสร้างต้นไม้ (Decision Trees)ให้เป็น Root Node และ Node ที่เรียงมาเป็น ChildNode การวิเคราะห์และการจำแนกตัวแปรเป็นการศึกษาตัวแปรเพื่อการบ่งชี้ถึงคุณลักษณะของข้อมูลธรรมดาทั่วๆไป หรือรูปแบบ (patterns) ในขั้นตอนนี้จะประยุกต์วิธีการคำนวนอย่างเฉพาะเจาะจงเพื่อหาการจัดกลุ่มข้อมูล การจำแนก การรวมกลุ่มหรือผลที่จะตามมา การพึ่งพา การเชื่อมโยงหรือความเกี่ยวพันกันของข้อมูล แบบแผน แนวโน้ม อัลกอลิทึม (Algorithm) ในการทำงานของโครงสร้างต้นไม้ (Decision Trees) ที่จะกล่าวถึงมี 3 แบบมีอัลกอลิทึม (Algorithm) หลักในการทำงานของโครงสร้างต้นไม้ (Decision Tree) อยู่ด้วยกันหลายชนิด

3.1 CART (Classification and Regression Trees) เป็นวิธีการสร้างโครงสร้างต้นไม้ที่ได้รับความนิยมมากที่สุดเป็นการทำงานในแบบ Binary Tree คือมีการแตกกิ่งออกเป็น 2 กิ่งเสมอ การทำงานจะต้องมีการกำหนดตัวแปรไม่อิสระ (Dependent Variable) ที่ต้องการจะศึกษาเสียก่อนแล้วจึงนำตัวแปรอิสระ (Independent Variable) มาแตกออกเป็นต้นไม้โดยเป็นการปลูกออกมาให้เต็มต้นเสร็จแล้วจึงทำการตัด Node โดยมีการตรวจสอบ Node ที่มีผิดพลาดออกไป แล้วจึงเลือก Node ที่นำเสนอความเกี่ยวข้องกับเรื่องที่จะศึกษามากที่สุด

3.2 C4.5 มีลักษณะคล้ายกับ CART คือมีการอธิบายถึงที่มาและที่ไปของ Node แต่ละ Node แต่สิ่งแตกต่างที่เห็นได้ชัดก็คือ CART ทำงานบน Binary Node ส่วน C4.5 ทำงานบนจำนวน Node หลาย ๆ Node ที่แตกกิ่งออกมาแต่ละครั้ง อีกความแตกต่างกันก็คือในการตัดกิ่งออกนั้นC4.5 ไม่จำเป็นต้องมีการตรวจสอบ Node ก่อน

3.3 CHAID (Chi-Square Automatic Interaction Detection) ลักษณะการทำงานของ CHAID แตกต่างจาก 2 ชนิดแรกเนื่องจาก CHAID พยายามที่จะหยุดการเจริญเติบโตของต้นไม้ก่อนที่มันจะโตเต็มที่ แล้วค่อย ๆ สร้าง Child Node ที่เหมาะสมจนต้นไม้เต็มที่ สิ่งที่แตกต่างจาก 2 ชนิด แรกก็คือ CHAID ทำงานบนข้อมูลที่เป็นประเภทเดียวกันเท่านั้น

3.6 กรณีศึกษาขบวนการทำงานของ Data Mining (Data Mining Process)

ขั้นตอนการทำ Mining นั้นมีขบวนการทั้งหมดหลายขั้นตอน ขั้นตอนการทำ Data Miningนั้นคือขั้นตอนการสร้างรูปแบบ โดยสร้างรูปแบบของกลุ่มข้อมูลเพื่อสร้างความเข้าใจในแนวโน้มจะเกิดขึ้นรูปแบบ และความเกี่ยวข้องกันของกลุ่มข้อมูลเพื่อใช้ในการทำนายแนวโน้มจากขั้นตอนการทำ Data Mining มีจุดประสงค์ทางธุรกิจเป็นตัวกำหนดทิศทางของขั้นตอนการทำ Data Mining ทั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หมดนอกจากนี้ขั้นตอนของการทำ Data Mining นั้นอาจดูเหมือนถูกเรียงลำดับกันแต่ว่าขั้นตอนจริงๆของการทำมีลักษณะเป็นขั้นตอนที่ต้องการย้อนกลับไปกลับมาตลอดเวลา โดยมีการทำซ้ำอยู่ตลอดเวลานอกจากนี้ในแต่ละขั้นตอนยังเป็นขั้นตอนที่ต้องใช้เวลาความพยายามและการใช้ทรัพยากรที่แตกต่างกัน

1.กำหนดวัตถุประสงค์ (Objective)

การกำหนดวัตถุประสงค์นั้นจะต้องมาจากการมองเห็นปัญหาที่มีอยู่และต้องการทำความเข้าใจปัญหาทางด้านทางธุรกิจนั้น ๆ ก่อน เพื่อดูถึงความจำเป็นที่จะนำระบบสารสนเทศเข้ามาใช้แก้ปัญหาที่มีอยู่ ให้เกิดความเหมาะสม การกำหนดปัญหาเท่ากับเป็นตัวกำหนดทิศทางในการทำระบบสารสนเทศในการทำ Data Mining การกำหนดปัญหาไม่ถูกต้องย่อมทำให้ไม่สามารถแก้ปัญหาที่มีอยู่ได้ การใช้ Data Mining ในการแก้ปัญหานี้จะประกอบ ด้วยดาวิเคราะห์ข้อมูลเบื้องต้นว่าเรามีข้อมูลอะไรอยู่บ้างและต้องการอะไรจากข้อมูลเพิ่มเติม จะทำให้สามารถมองเห็นข้อมูล (Database) และอัลกอลิทึม (Algorithm) ที่ใช้ในเบื้องต้นที่สัมพันธ์กับวัตถุประสงค์ในการทำระบบนั้น ๆ

2.การจัดการเตรียมข้อมูล (Data Preparation)

การจัดเตรียมข้อมูลมีความสำคัญมากในการทำระบบ Data Mining ในขั้นตอนนี้การทำงานจะต้องใช้เวลาในการเตรียมข้อมูลเป็นอย่างมาก โดยระบบกับคำถามบ่อยๆที่ว่า จะสร้างตารางข้อมูลอย่างไร และข้อมูลจะอยู่ในแบบฟอร์มอย่างไร และจะมีการดำเนินการอย่างไรให้ข้อมูลที่มีเป็นข้อมูลที่มีนัยสำคัญ เช่น การเก็บข้อมูลผู้ป่วยที่ได้รับความเจ็บปวดที่หลัง จะมีข้อมูล ของผู้ป่วยแพทย์ที่ทำการรักษา การประกันชีวิต หรือเกี่ยวกับโรงพยาบาล ซึ่งจะถูกเก็บในฐานข้อมูลที่แตกต่างกันและจำนวนมากถึงแม้ว่าข้อมูลจะเป็นอยู่ในรูปแบบของ Relational Database แต่การเข้าถึงข้อมูลค่อนข้างยุ่งยากดังนั้นการที่จะเข้าถึงข้อมูลมีหลายวิธี เช่น

1. การเข้าถึงข้อมูลของคลังสินค้า Data Warehouse วิธี Data Mining จะถูกพิจารณาเพราะใช้เวลาในการสร้างรูปแบบการตัดสินใจ (Decision Support Systems) ให้กับบริษัทในการที่จะทำให้ได้ข้อมูลที่แบ่งแยกความน่าสนใจของแนวโน้มของ Data Mining คือการรวมเข้ากับฐานข้อมูลของ Data Warehouse โดยตรงด้วย Data Mining Tool
2. การเข้าถึงข้อมูลโดยผ่านขบวนการ Relational Transaction- Based Database ข้อมูลที่ไม่อยู่ใน Data Warehouse สามารถที่จะ Access ข้อมูลจาก Relational หรือ Transaction–Based Database โดยใช้การเชื่อมต่อด้วย ODBC ซึ่งเป็นการเชื่อมต่อที่มีมาตราฐานสำหรับฐานข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่ละชนิด ปัญหาในการ Access ข้อมูลอาจแก้โดย Relational database โดยการสร้าง View Data

3. การเข้าถึงข้อมูลโดยการทำ Data Conversion Utility ถ้า Data Warehouse ไม่อยู่ในรูปที่ Data Mining Tool จะเข้าใจได้จะมีวิธีที่ใช้ข้อมูลได้หลายวิธีเช่นการทำ Data Conversion Utility โดยการ Merge ข้อมูลหรือการเปลี่ยน Format

4.การเข้าถึงข้อมูลโดยใช้ Query Tools การที่ใช้ query tool เพื่อร่วมใช้ Table ในDatabase ต้องสร้าง Data Source ที่เหมาะสม

ในขั้นตอนนี้จะแบ่งการทำงานของการจัดเตรียมข้อมูลแบ่งออกเป็น 3 ส่วน

2.1 การคัดเลือกข้อมูล (Data Selection) ที่มีอยู่ในแหล่งข้อมูลและคัดเลือกเฉพาะข้อมูลที่มีต้องการสำหรับการเริ่มต้นการวิเคราะห์ การเตรียมข้อมูลสำหรับการทำ Mining ในอนาคต การเลือกตัวแปรจำเป็นต้องเข้าใจและเข้าใจความหมาย กำหนดค่าที่เป็นไปได้ของข้อมูลนั้น ๆ ชนิดของตัวแปรแบ่งเป็น ตัวแปรแบบ Category แบบเป็นหมวดหมู่เช่นเพศ (หญิง ชาย) ระดับการศึกษา (ปริญญาตรี ปริญญาโท ปริญญาเอก) สถานะการแต่งงาน (โสดแต่งงานหม้าย) หมู่เลือด (A,B,AB,O) แผนกผู้ป่วย (ห้องเดี่ยว ห้องรวม ห้องคู่)เป็นต้นตัวแปรแบบ Quantitative แบบระบุจำนวนซึ่งตัวแปรจะแบ่งเป็น ตัวแปรที่มีค่าที่ต่อเนื่อง (Continuous) เช่น อายุ รายได้ น้ำหนัก ส่วนสูง เป็นต้น และตัวแปรที่มีค่าที่เป็นจำนวนเต็ม (Discrete) เช่น จำนวนวันที่นอนในโรงพยาบาลเดือนหรือ ในขั้นตอนนี้จะแบ่งการทำงานของการจัดเตรียมข้อมูลโดยทำการการเลือกผู้ป่วย (Case Selection) ซึ่งต้องมีเกณฑ์ที่แน่นอนในการเลือกผู้ป่วยมาศึกษา ควรระบุทั้งชนิด และความรุนแรงของโรคที่น่าสนใจ การเลือกผู้ป่วยควรมีการระบุเจาะจงให้ชัดเจน เพราะเป้าหมายของการศึกษาอยู่ที่การหาความสัมพันธ์ระหว่างปัจจัยเสี่ยงและโรค ให้ใกล้ความจริง (validity) มิใช่ต้องการขยายความสู่ประชากรเป้าหมาย (generalizability) เป็นสำคัญ เช่น เลือกศึกษาเฉพาะ lobular breast cancer ในหญิงก่อนหมดประจำเดือน แทนที่จะศึกษามะเร็งเต้านมทั้งหมดในสตรีทุกอายุ ทั้งนี้เพราะมะเร็งแต่ละชนิดอาจจะสัมพันธ์กับเหตุแต่ละอย่าง การเลือกผู้ป่วยที่กว้างเกินไป อาจทำให้หาความสัมพันธ์ที่จำเพาะไม่ได้ ในทำนองเดียวกัน ถ้าจะศึกษาปัจจัยเสี่ยงต่อความพิการแต่กำเนิด ถ้าเก็บข้อมูลทั้งโรคปากแหว่งโรคหัวใจ โรคเท้าพิการ รวมไปก็อาจจะหาความสัมพันธ์ที่จำเพาะไม่ได้ ต้องเลือกเฉพาะโรคใดโรคหนึ่งมาศึกษาหากเป็นไปได้ ควรเลือกผู้ป่วยทั้งเก่าและใหม่ หรือ incident case โดยเฉพาะโรคที่สนใจมีโอกาสหายหรือตาย การเลือกผู้ป่วยทั้งเก่าและใหม่ (prevalence case) ปนกันไปอาจมีอคติจากกลุ่มที่หายหรือตาย ซึ่งอาจจะสะท้อนถึงความรุนแรงของโรค กล่าวคือ คนที่หาย อาจจะมีพยาธิภาพน้อย และคนที่ตาย เป็นกลุ่มที่เป็นมาก พวกที่ไม่ตายและไม่หายเป็นกลุ่มที่อาจมีปัจจัยเสี่ยงต่างไปจากพวกที่ตาย หรือหาย ทำให้ผลสรุปผิดไป สำหรับโรคที่ไม่มีโอกาสหาย และธรรมชาติของโรค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไม่ยาวนัก ผู้ป่วยก็เสียชีวิต อาจใช้ Prevalence case แทน incident case ได้ ในกรณีเช่นนี้ ควรรวมผู้ป่วยที่เสียชีวิตแล้ว เข้าในการศึกษาด้วย เพื่อให้มีจำนวนผู้ป่วยมาก และได้ภาพที่ชัดเจน

ในการเลือกผู้ป่วย ควรพิจารณาเฉพาะผู้ที่ยังไม่มีประวัติของโรคขณะที่ได้รับปัจจัยเสี่ยง เช่น ถ้าอยากศึกษาว่า การใช้ยาบ้วนปากขจัดกลิ่นปาก ก่อนมีแผลเรื้องรังในปากเพราะผู้ป่วยที่มีแผลเรื้อรังในปาก อาจจะเป็นมะเร็งอยู่แล้ว และเพราะเขาเป็นมะเร็งปาก ทำให้มีกลิ่นปาก จึงมาเลือกใช้ยาบ้วนปาก เช่นนี้ก็ทำให้เกิดอคติในการศึกษา อคติที่เกิดเนื่องจากเลือกผู้ป่วยที่มีโรคก่อนได้รับปัจจัยเสี่ยง เรียกว่า Protopathic Bias (Proto = ตอนเริ่มเป็น : pathic = มีพยาธิสภาพ)
ในทำนองเดียงกัน ผู้ป่วยบางคนอาจจะได้รับปัจจัยเสี่ยงภายหลังเกิดพยาธิสภาพ ฉะนั้นปัจจัยเสี่ยงนั้นไม่น่าทำให้เกิดโรค จึงควรออกจากการศึกษา ตัวอย่างเช่น จะศึกษาเด็กที่ความพิการแต่กำเนิด ว่าเกิดจากแม่กินสมุนไพรชนิดหนึ่งหรือไม่ เราได้ตรวจพบเด็กพิการแต่กำเนิด จากแม่มีประวัติการรับประทานสมุนไพร แต่เริ่มรับประทานครั้งแรก หลังจากเด็กที่เป็นโรคคลอดมาแล้วเช่นยาสมุนไพรก็ไม่จะทำให้เด็กพิการ ไม่ควรจะนำเด็กนี้มาศึกษาแหล่งข้อมูลเกี่ยวกับโรคที่สนใจ อาจได้มาจากโรงพยาบาล ในต่างประเทศอาจจะหาข้อมูลได้จากมรณะบัตร บริษัทประกันชีวิต หรือบันทึกของแพทย์ผู้ให้การรักษา แต่ในประเทศไทย ข้อมูลเหล่านี้ยังไม่แม่นยำพอที่จะนำมาใช้ประโยชน์ได้

2.2 การกลั่นกรองความถูกต้องของข้อมูล (Data Processing) เป็นวิธีที่จะนำข้อมูลไปวิเคราะห์ว่าข้อมูลนั้นๆถูกต้องเหมาะสมหรือไม่ ข้อมูลที่สมบูรณ์เป็นตัวกำหนดความสำเร็จของการทำ Data Mining เพราะข้อมูลส่วนใหญ่ที่มีในองค์กรไม่ได้ถูกเตรียมไว้เพื่องานของ Data Mining โดยเฉพาะข้อมูลบางครั้งถูกจัดเก็บไว้ไม่ดี สิ่งที่มีความสำคัญในการกลั่นกรองข้อมูลคือคุณภาพของข้อมูล (Data Quality) การศึกษาตัวแปรที่มีผลต่อระบบ (Defining a Study) เป็นการศึกษาความชัดเจนของตัวแปรโดยต้องทำความเข้าใจกับตัวแปรที่พาดพิงต่อวัตถุประสงค์ของการทำ Data Mining เครื่องมือที่ใช้ในการกลั่นกรองข้อมูลมีดังต่อไปนี้ ตัวแปรแยกประเภท (Category) อาจใช้วิธีแจกแจงความถี่เพื่อให้เกิดความเข้าใจ เครื่องมือที่ใช้อาจใช้เป็นเครื่องมือทางด้านกราฟฟิค ตัวแปรเชิงปริมาณ (Quantitative) อาจใช้วิธีทางสถิติเช่นค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ในการทำการกลั่นกรองข้อมูลเราจะทำการเลือกข้อมูลที่ไม่สมบูรณ์ หรือข้อมูลที่ผิดพลาดออกไปแต่จะต้องไม่ละเลยข้อมูลบางตัวที่คิดว่าไม่สำคัญซึ่งอาจแสดงถึงแนวโน้มที่เราต้องการหาอยู่

3. การแปลงข้อมูล (Data Transformation)

เป็นการแปลงข้อมูลไปสู่โมเดล (Model) การวิเคราะห์โดยใช้วิธีการสถิติ ในการลดจำนวนข้อมูล หรือลดจำนวนตัวแปรในการ Process โดยการนำเอาตัวแปรตั้งแต่ 2 ตัวขึ้นไปมารวมกันแล้วทำการ Process รวมทั้งทำการกำหนดค่าของตัวแปรที่ทำการศึกษา เช่น อายุ กำหนดอายุ โดยจัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบ่งเป็นช่วงอายุต่างๆในแต่ละช่วงโดยกำหนดค่าแต่ละช่วงอายุไว้เพื่อสะดวกและง่ายต่อการทำ Database เพศกำหนดเป็นเพศหญิง,เพศชาย การศึกษาในแต่ละระดับการศึกษา เช่น ประถมศึกษา มัธยมศึกษา อนุปริญญา ปริญญาตรี ปริญญาโท ปริญญาเอก เป็นต้น ความสูง แบ่งตามช่วงความสูงของผู้ป่วยซึ่งมีหน่วยเป็นเซนติเมตรและใส่จำนวนตัวเลขในแต่ละค่าความสูง น้ำหนัก แบ่งเป็นช่วงน้ำหนักของผู้ป่วย ซึ่งมีหน่วยเป็นกิโลกรัมและจะใส่จำนวนตัวเลขในแต่ละค่าของน้ำหนัก รายได้ของผู้ป่วย โดยกำหนดช่วงของรายได้ เช่น >5,000 บาท/เดือน 5,000-10,000 บาท/เดือน 10,000-20,000 บาท/เดือน 20,000-35,000 บาท/เดือน และจะกำหนดค่าโดยใส่ตัวเลขแทนค่าในแต่ละตัวแปรแต่ละตัว การออกกำลังกาย จะแบ่งเป็นความถี่เช่นไม่เคยบางครั้ง สม่ำเสมอ เป็นต้น การสูบบุหรี่ จะแบ่งความถี่เช่น เคย ไม่เคย บางครั้งเป็นต้น การติดเชื้อในโรงพยาบาล(ติดเชื้อไม่ติดเชื้อ) ชนิดของเชื้อโรคโดยแยกตามประเภทของเชื้อโรคเช่นไวรัส แบคทีเรีย เชื้อรา เป็นต้น

4. ขั้นตอน Data Mining

ในขั้นตอนนี้จะเริ่มทำ Data Mining กันโดยมีวัตถุประสงค์คือการนำเอาวิธีการทำ Data Mining ที่ได้เลือกไว้มาเริ่มทำหรือวิธีที่จะทำการวิเคราะห์ข้อมูลจริงๆแล้วเราแยก Data Mining ออกเป็นอีกขั้นหนึ่ง แต่มองในแง่การทำงานแล้วค่อนข้างจะแยกออกได้ลำบากจากขั้นอื่นๆเนื่องจากต้องมีการทำงานที่เกี่ยวข้องกันและย้อนไปย้อนมา โดยทางกายภาพนั้นมีการใช้กรรมวิธี Data Miningจำนวนหนึ่งร่วมกัน ดังนั้นนักวิเคราะห์ข้อมูลจะต้องดูแลหลายๆเรื่องได้แก่ โครงสร้างข้อมูล ความหมายที่เกิดขึ้น โดยจะต้องดูแลข้อมูลที่ได้กลั่นกรองแล้วยังมีข้อมูลมากมาย รูปแบบของระบบงาน (Application) ที่หลากหลายจะเป็นตัวกำหนดว่าจะต้องใช้กรรมวิธี Data Mining กี่วิธีในการวิเคราะห์ปัญหานั้นๆ การพัฒนางาน Data Mining มีการใช้วิธีการทำงานหลายอย่าง นักวิเคราะห์จะต้องเลือก ระเบียบกฎเกณฑ์ที่เหมาะสมกับลักษณะของจุดประสงค์ ซึ่งแต่ละระเบียบกฎเกณฑ์มีทั้งข้อดีและข้อเสียในตัวมันเองซึ่งจะกล่าวถึงกรรมวิธีในการทำ Data Mining อีกส่วน

5. การวิเคราะห์ผล (Analysis Of Result)

เป็นการวิเคราะห์ผลของการทำการ Data Mining จะช่วยให้นักวิเคราะห์ข้อมูลสามารถวิเคราะห์ข้อมูลได้สะดวกและรวดเร็วขึ้น ซึ่งการวิเคราะห์โดยวิธีนี้จะแตกต่างจากวิธีทางสถิติเนื่องจากทางสถิติจะเป็นการพิสูจน์สมมุติฐานแต่ละผลของ Data Mining จะให้แนวทางกับการทำงานมากกว่าการวิเคราะห์โดยการทำ Data Mining เป็นขั้นตอนที่มีความสำคัญมากในขบวนการดำเนินการทั้งหมดเครื่องมือทางด้านกราฟิกจะช่วยทำให้นักวิเคราะห์ธุรกิจและนักวิเคราะห์ข้อมูลสามารถทำการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิเคราะห์ได้อย่างสะดวกรวดเร็วยิ่งขึ้นแม้ขั้นตอนของการวิเคราะห์ผลเองก็ไม่สามารถแยกออกจากขั้นตอนของการทำ Data Mining หลายๆตัวมีเครื่องมือประเภทนี้อยู่ด้วย

6. การพยากรณ์แนวโน้ม (Predictive Modeling)

เป็นขบวนการที่คล้ายกับการเรียนรู้ของมนุษย์คือมีความเข้าใจถึงลักษณะของตัวแปรที่จะศึกษาอย่างแท้จริงทั้งลักษณะของตัวแปรที่มีสำคัญเป็นพิเศษจะใช้โมเดลนี้ในการวิเคราะห์ฐานข้อมูลที่มีอยู่เพื่อกำหนดสาระสำคัญของข้อมูล ดังนั้นข้อมูลที่นำมาใช้จะต้องเป็นข้อมูลที่มีความสมบูรณ์จึงจะทำให้โมเดลสามารถให้คำทำนายได้อย่างเหมาะสมโดยการเริ่มต้นจะต้องให้คำตอบที่ถูกต้องเพื่อที่จะได้เห็นถึงข้อสังเกตใหม่ๆ โมเดลของการทำงานอาจอยู่ในรูปของตรรกวิทยาคือ ถ้า___แล้ว (IF THEN) ซึ่งเป็นการตรวจสอบข้อมูลในแต่ละกรณี ข้อมูลตัวแปรจะมีลักษณะที่สำคัญแบ่ง 2 แบบ คือ

1. การใช้การประเมินค่าความต่อเนื่องของตัวแปรที่มีความเกี่ยวเนื่องกับฐานข้อมูลเรียกว่า Value Prediction เช่น ถ้าต้องการทำนายช่วงเวลาที่ผู้ป่วยเกิดการติดเชื้อในโรงพยาบาล อายุของผู้ป่วย เพศ สถานที่หอผู้ป่วย ประเภทผู้ป่วยที่ทำการเฝ้าระวังการติดเชื้อ การได้รับยาปฏิชีวนะ เป็นต้น
2. การใช้หลักการประเมินค่าตัวแปรโดยการจัดกลุ่มตัวแปรที่มีเรียกว่า Database Classification เช่น แพทย์ที่ทำการศึกษา จำนวนวันที่ผู้ป่วยนอนในโรงพยาบาล เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
Decision Tree และขั้นตอนการดำเนินงาน

การศึกษาครั้งนี้เป็นการศึกษาแบบการสำรวจชุมชน (Multiserial cross-sectional study) เพื่อวัดปัจจัย เสี่ยงที่สนใจต่อการติดเชื้อในโรงพยาบาลเป็น ระยะ ๆ โดยที่ลักษณะประชากรที่ศึกษาเป็นประชากรที่รับการรักษาในโรงพยาบาลนั้น อาจเปลี่ยนแปลงอยู่เสมอ จึงต่างจากการศึกษาที่มีการติดตามคนกลุ่มหนึ่งเป็นเวลานาน (longitudinal cohort study) ข้อมูลจากการสำรวจแบบชุมชน (Multiserial cross-sectional study) จึงไม่เหมือน Longitudinal study เพราะปัญหาอคติที่เกิดจากความแตกต่างของประชากรที่เริ่มศึกษาและสภาวะแวดล้อมที่เปลี่ยนแปลงไป (environmental change) ตัวอย่าง Multiserial cross sectional study ได้แก่ การศึกษาของ Framingham ที่ศึกษาปัจจัยเสี่ยงต่อกล้ามเนื้อหัวใจตาย (Myocardial Infarction) และโรคเส้นเลือดในสมองตีบ แตก หรือตัน เป็นต้น

การศึกษานี้เป็นการศึกษาผลการใช้ระบบสารสนเทศเพื่อการควบคุมและป้องกันการติดเชื้อในโรงพยาบาลโดยระบบที่นำมาศึกษาคือระบบ Data Mining

4.1 ลักษณะของประชากรที่ศึกษา

ปัจจัยที่นำมาศึกษา ได้แก่ข้อมูลผู้ป่วยที่เข้ารับการรักษาในโรงพยาบาลเกษมราษฎร์ระหว่างวันที่ 1 มกราคม 2542 – 31 ธันวาคม 2542 และเป็นกลุ่มผู้ป่วยที่มีการแบ่งกลุ่มผู้ป่วยดังต่อไปนี้

1. กลุ่มผู้ป่วยหนักที่รับการรักษาในแผนกผู้ป่วยหนัก(ICU)
2. กลุ่มผู้ป่วยที่นอนพักรักษาตัวในโรงพยาบาลเกิน 3วัน
3. กลุ่มผู้ป่วยที่ได้รับการสอดใส่อุปกรณ์เครื่องมือแพทย์ (Instrument)
4. กลุ่มผู้ป่วยหลังผ่าตัดทุกราย/คลอดบุตร

4.2 เครื่องมือที่ใช้ในการศึกษา

โปรแกรมที่นำมาใช้ในการวิเคราะห์ปัจจัยที่มีผลต่อการติดเชื้อในโรงพยาบาลคือโปรแกรมระบบ Data Mining โดยมีขบวนการในการดำเนินงานดังต่อไปนี้

1.รวบรวมข้อมูลผู้ป่วยจากฐานข้อมูลผู้ป่วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.รวบรวมข้อมูลผู้ป่วยที่ได้รับการติดเชื้อในโรงพยาบาลเปรียบเทียบกับข้อมูลผู้ป่วยที่ไม่ติดเชื้อในโรงพยาบาล

3.พิจารณาปัจจัยที่คาดว่าจะมีผลต่อการติดเชื้อในโรงพยาบาล นำปัจจัยดังกล่าวมาวิเคราะห์ผ่านโปรแกรมที่ต้องการศึกษา ปัจจัยที่นำมาศึกษาประกอบด้วย

ส่วนที่ 1 ข้อมูลผู้ป่วย ได้แก่ HN AN อายุ เพศ ระยะเวลาที่พักรักษาตัวอยู่ในโรงพยาบาลชนิดของกลุ่มที่ทำการศึกษา ชนิดของการรับไว้รักษาในหอผู้ป่วย ประเภทผู้ป่วยที่เข้ารับบริการในโรงพยาบาลแยกตามแผนกที่รับผู้ป่วยพักรักษา

ส่วนที่ 2 การได้รับยาปฏิชีวนะ

ส่วนที่ 3 การวินิจฉัยการติดเชื้อในโรงพยาบาล

4.3 การรวบรวมข้อมูล

ในการศึกษาครั้งนี้ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลด้วยตนเอง โดยดำเนินการตามขั้นตอนดังต่อไปนี้

1. ผู้ศึกษาทำการเก็บข้อมูลผู้ป่วยตามที่ได้รับมอบหมายจากผู้บริหารให้ทำการเฝ้าระวังโดยได้รับอนุญาตให้เก็บรวบรวมข้อมูลผู้ป่วยในโรงพยาบาล
2. หลังจากได้รับอนุมัติจากผู้บริหารแล้ว ผู้ทำการศึกษาพบแพทย์ หัวหน้าหน่วยงาน หัวหน้าฝ่ายการพยาบาลเพื่อขอความร่วมมือในการทำการศึกษา
3. การเก็บข้อมูลของผู้ป่วยในช่วงระยะเวลา 1 ปีเป็นการศึกษาโดยเก็บข้อมูลในลักษณะแบบ Prospective และ Retrospective
4. การเก็บข้อมูลโดยการสำรวจเวชระเบียนผู้ป่วย
5. บันทึกรายละเอียดของผู้ป่วยในรายงานการเฝ้าระวังโรคต่อเนื่อง
6. ติดตามประเมินอาการและอาการแสดงของผู้ป่วยเพื่อประเมินการติดเชื้อในโรงพยาบาล
7. ผู้ศึกษาดำเนินการเฝ้าระวังการติดเชื้อและเก็บรวบรวมข้อมูลของปัจจัยเสี่ยงจนครบระยะเวลา 4 เดือนลักษณะข้อมูลเป็น Prospective และค้นหาข้อมูลแบบ Retrospective จำนวน 8 เดือน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 การวิเคราะห์ข้อมูลโดยใช้ Data Mining

การวิเคราะห์ข้อมูลโดยใช้ขบวนการของ Data Mining นี้ต้องเริ่มจากการจัดเตรียมข้อมูลที่จะนำไปใช้ในการวิเคราะห์เพื่อหารูปแบบของการค้นหาที่ต้องการและเป็นประโยชน์ต่อการศึกษา ขบวนการในการดำเนินการนี้ใช้ เครื่องมือได้ 3 รูปแบบคือ

1. การใช้โครงสร้างต้นไม้ (Tree Miner)
2. การหาความสัมพันธ์ของปัจจัยที่ทำการศึกษาเชิงจัดการ (Transaction Association)
3. การหาความสัมพันธ์ของปัจจัยที่ทำการศึกษาเชิงกรณี (Case Association)

และมีขบวนการในการจัดการข้อมูลโดยมีการจัดลำดับขบวนการดังต่อไปนี้

การจัดเตรียมข้อมูลที่ทำการศึกษาแล้วนำไปแปลงข้อมูลให้อยู่ในรูปต่างๆการจัดการข้อมูลให้อยู่ในรูปแบบการจัดการตารางขึ้นมามีขบวนการที่ช่วยในการสนับสนุนด้วยวิธีการและรูปแบบต่างๆ หลายวิธีการเช่นการรวมกันในตารางที่มีข้อมูลจากตารางที่แตกต่างกัน (Join) การรวมเข้าด้วยกันของตารางที่มีแถวข้อมูลที่มีความหมายเหมือนกัน (Merge) การกรองข้อมูลจากแถวข้อมูลของตารางตามที่มีข้อกำหนดไว้ (Filter) ตารางสามารถที่จะค้นหาได้ทั้งจากข้อมูลข้างล่างขึ้นไปหาข้อมูลด้านบน หรือสามารถที่จะค้นหาได้ทั้งจากข้อมูลด้านบนลงไปหาข้อมูลข้างล่าง (Sort) นอกจากนี้มีการจัดการข้อมูลที่ทำการศึกษาได้โดยใช้ขบวนการที่ช่วยในการสนับสนุนด้วยวิธีการและรูปแบบต่างๆหลายวิธีการเช่นการสร้างข้อมูลใหม่จากการคำนวณการจัดการซึ่งจะใช้ในการตรวจสอบความถูกต้องของข้อมูลและทำความหมายให้เกิดกับข้อมูลอย่างชัดเจน (Derive) หรือการสร้างข้อมูลใหม่โดยการดำเนินการสรุปข้อมูลให้อยู่ในรูปแบบของค่าต่างๆเช่นค่าโดยเฉลี่ย ค่าสูงสุด ค่าต่ำสุด การรวมจำนวน การนับจำนวนเป็น ขอบเขตของการทำงาน Data Mining เป็นกระบวนการหรือวิธีสำหรับการดำเนินการ Data Mining ในการวิเคราะห์ข้อมูลทางการแพทย์ เช่นเดียวกันกับการใช้เทคโนโลยีแก้ปัญหาทางด้านธุรกิจซึ่งแต่ละคนจะได้ประโยชน์ แตกต่างกันตามความรู้ความชำนาญในข้อมูลนั้นๆรวมทั้งความสำเร็จของกระบวนการขึ้นอยู่กับข้อมูลที่มีความแน่นอนและถูกต้องในการริเริ่มการพัฒนาข้อมูลซึ่งมีการกำหนดวิธีและกระบวนการที่เกี่ยวข้องกับรายละเอียดของงานในแต่ละขั้นตอน ดังนี้

1. กำหนดเป้าหมายที่ชัดเจน หรือวัตถุประสงค์สำหรับกระบวนการ ควรมีการดำเนินการหาความสัมพันธ์ของปัจจัยที่ทำการศึกษา เช่น ประเภทผู้ป่วย เพศ อายุ จำนวนวันที่นอนในโรงพยาบาล การได้รับยาปฏิชีวนะ แผนกที่รับผู้ป่วย ประเภทผู้ป่วยที่ทำการเฝ้าระวังโรค เป็นต้น
2. ทำการคัดเลือกเป็นกระบวนการที่คัดแยกปัจจัยที่จำเป็นสำหรับการทำ Data Mining
3. การจัดเตรียมข้อมูลเป็นการตรวจนับข้อมูลและการสร้างคอลัมภ์ใหม่ในข้อมูลโดยการรวมข้อมูล (Join) การรวมข้อมูลโดยการหาความสัมพันธ์เชิงซ้อน (Merze) การผ่านการคำนวน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การสรุปรวมในตารางที่พร้อมสำหรับใช้กับการทำ Data Mining ดังนั้นในตารางจะมีชุดข้อมูล 2 ชุดคือ ชุดสำหรับการค้นหารูปแบบ 1 ชุดและสำหรับการตรวจสอบรูปแบบ 1 ชุด

4. การตรวจสอบ การเตรียมข้อมูล เพื่อให้ได้รูปแบบที่ค้นหาและผลการเตรียมข้อมูลที่น่าเชื่อถือ โดยทั่วไปภาวะนี้รวมถึงการวิเคราะห์ทางสถิติซึ่งค่าที่ได้เช่นค่าต่ำสุด (Minimum) ค่าสูงสุด (Maximum) ค่าโดยเฉลี่ยเป็นต้น การแจกแจงความถี่ของแต่ละค่าของปัจจัยจึงรวมถึงฟิล์ม รูปภาพ เพื่อให้เกิดความเข้าใจระหว่างปัจจัยที่ทำการศึกษา
5. การค้นหารูปแบบ จะเป็นขั้นตอนที่มีประสิทธิภาพสูงสุด ถ้านำอัลกอลิทึมในการค้นหารูปแบบมาใช้ในขบวนการนี้ โดยนำความรู้พื้นฐานมาใช้ร่วมในการวิเคราะห์ ในกรณีที่จะใช้โครงสร้างต้นไม้ ผู้ใช้สามารถใช้จุดต่างๆ ในโครงสร้างต้นไม้ เพื่ออธิบาย หรือตรวจสอบเพื่อกรอกข้อมูลการอธิบายอัลกอลิทึมเพื่อพิจารณาข้อมูลต่อไปเพื่อใช้สำหรับการแตกกิ่ง การพิจารณาปัจจัยจะรวมถึงการวิเคราะห์ความสามารถของการค้นรูปแบบเพื่อทำนายเหตุการณ์และตรวจสอบความเป็นอิสระของข้อมูลรูปแบบขั้นตอนนี้จะใช้ตามวัตถุประสงค์ของการใช้วิเคราะห์ตัวปัจจัยซึ่งรูปแบบมีหลายแบบฟอร์ม
    1. รูปแบบ นำเสนอ ใช้ Data Mining tool สร้างรายงานในรูปแบบต่าง ๆ เช่น รูปแบบตาราง รูปแบบจุดความถี่ รูปแบบ 3 มิติ เป็นต้น แทนการใช้โครงสร้างต้นไม้
    2. รูปแบบรายงานเชิงธุรกิจ เพื่อทำ Business Intelligence Reports ซึ่งต้องการ Data Mining Tool เพื่อสร้าง SQL แทนโครงสร้างต้นไม้
    3. รูปแบบ การใช้คะแนน ข้อมูลและติดป้ายลงบันทึกในฐานข้อมูลกับรูปแบบนั้นซึ่งสามารถทำได้โดยตรง โดยใช้ Data Mining Tool หรือสร้าง SQL
    4. การใช้ระบบตัดสินใจ (Decision Support Systems)ใช้ Data Mining Tool สร้างC Model
    5. การตรวจสอบความผิดปกติ ใช้แบบอย่างของกระบวนการทางธุรกิจ การที่มีสถานการณ์ที่เบี่ยงเบนจากปกติจะตรวจจับได้โดยใช้ Data Mining Tool เป็นส่วนประกอบในการตรวจดูหรือผ่านการใช้ SQL

1. การใช้โครงสร้างต้นไม้ (Tree Miner)

เป็นวิธีการที่ออกแบบสำหรับการสร้างโครงสร้างต้นไม้จากชุดของตัวแปรโดยการแบ่งออกเป็นกิ่งหลายๆ ครั้ง เพื่อให้ได้ข้อมูลที่มีคุณสมบัติที่ต้องการจนกิ่งสุดท้าย การใช้โครงสร้างต้นไม้จะใช้ตัวเลข วันที่/เวลา และในการแตกกิ่งโครงสร้างต้นไม้จะทำได้ 2 กิ่งในแต่ละครั้ง ในขั้นตอนแรกการแตกกิ่งครั้งแรกค่าที่ใช้ควรเป็นค่าที่เป็นจำนวนเต็ม (Discrete) ขั้นต่อไป การจัดกลุ่มของตัวแปรอาจทำได้ 10 กลุ่ม หรือมากกว่า ซึ่งเป็นผลจากอัลกอลิทึม ที่สร้างขึ้นมาการดำเนินการ สามารถระบุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าของกลุ่มตัวแปรได้ชัดเจน โดยเชื่อมค่าของตัวแปรในแต่ละฟิลด์ จำนวนตัวเลข วันที่ เวลา การจัดกลุ่มค่าตัวเลขในช่วงค่าต่าง ๆจะทำให้การจัดทำโครงสร้างต้นไม้ทำได้ รวดเร็วขึ้นด้วย สำหรับค่าของผลลัพธ์ที่เป็นจำนวนเต็ม (Discrete) เมื่อใช้โครงสร้างต้นไม้ ผลที่ได้มีค่าหลายค่าควรพิจารณาวิเคราะห์ค่าที่มีจำนวนเต็ม < 4 ค่าจะช่วยให้เข้าใจและแปลผลได้ง่ายกรณีที่ผลลัพธ์มีจำนวน หลายค่า การจัดกลุ่มโดยการเลือกกลุ่มค่าของผลลัพธ์ที่น่าสนใจ แยกจากกลุ่มอื่นจะทำให้สามารถเห็นผลลัพธ์ค่อนข้างชัดเจนกว่า ตัวอย่างเช่น ถ้าผลลัพธ์ของอายุ เมื่อหากลุ่มของอายุ 21-30, 31-50, > 60 ความสนใจจะเป็นช่วงอายุซึ่ง ผลลัพธ์ตัวแปรจะเป็นจำนวน ถ้าไม่แบ่งเป็นช่วงของอายุ

2.การหาความสัมพันธ์เชิงจัดการ (Transaction Association)

การหาความสัมพันธ์เชิงการจัดการจะมีพารามิเตอร์ที่ช่วยจำกัดจำนวนรวมทั้งส่วนประกอบของการรวมกัน

1. ค่าสนับสนุน (Support) ค่าเริ่มต้นของค่าสนับสนุน = 0.20 หมายความว่า ความถี่ของชุดข้อมูลที่ทำการศึกษาต้อง >20 % ของจำนวน Transaction ทั้งหมดในชุดข้อมูลที่กำหนดนี้กระทบกับความสำเร็จของกระบวนการ Mining ที่จะขยายได้มากกว่าพารามิเตอร์ที่จะกล่าวตามมาค่าที่ต่ำกว่า Minimum Support หรือยาวกว่ากระบวนการอาจจะมีการรวมกันอาจจะมีประโยชน์อย่างยิ่งแม้ว่าจะนานๆ เกิดขึ้นทีและอาจนำไปสู่การค้นคว้าที่สำคัญที่สุด

2. ค่าความเชื่อมั่น (Minimum Confidence) ค่าเริ่มต้น = 0.7 หมายความว่าค่าความน่าเชื่อถือควรมากกว่าหรือเท่ากับ 70 %

3. ค่าสูงสุดในชุด (Maximum > Tems in set) หมายความว่าต้องใช้< จำนวน (N)_การรวมการจัดการจะมีศักยภาพเพิ่มขึ้นด้วยพารามิเตอร์นี้ นั่นคือเปลี่ยนจาก 3 เป็น 6 จะเพิ่มขึ้นเป็น 2 เท่า

4. จำนวนของความกว้าง (Number of numeric band) ค่าเริ่มต้น = 5 การแบ่งจำนวนข้อมูลเป็นวงกว้างโดยใช้วิธีของแคลคูลัสได้ค่าต่ำสุดถึงค่าสูงสุด (ค่าเฉลี่ย < 2 เท่าของค่าเบี่ยงเบนมาตราฐาน (Standard Deviation) ถึงค่าสูงสุด (ค่าเฉลี่ย +2 SD) โดยรายการจำนวน แล้วเรียงจากค่าต่ำสุดถึงค่าสูงสุด (Infinity) ความเร็วของขบวนการ Mining จะมีศักยภาพเพิ่มขึ้นด้วยพารามิเตอร์นี้นั้นคือเปลี่ยนจาก 5 เป็น 10 จะมากกว่า 2 เท่า

5. ค่าการแจกแจงความถี่ที่คาดหวัง (Surprise) มากกว่า/น้อยกว่าชุดของปัจจัยที่ทำการศึกษาต้องเกิดมากกว่าหรือน้อยกว่า การแจกแจงค่าความถี่ที่คาดหวังนั้นคือถ้ากำหนดค่า =2 เป็นค่าอย่างน้อยที่สุด 2 เท่าหรือครั้งหนึ่งของค่าที่คาดไว้ โดยสามารถเลือกความสัมพันธ์ของปัจจัยโดย คลิกบนคอลัมภ์เหนือกฎแห่งความสัมพันธ์ตัวอย่างเช่น ผู้ป่วย 10 % ของผู้ป่วยทั้งหมดได้รับการตัดเชื้อในโรงพยาบาลและ 10% ของทั้งหมดอายุ 31-50 ปี สามารถคาดได้ว่าเพียง 1% ของผู้ป่วยทั้งหมดที่ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รับการตัดเชื้อและอายุ 31-50 ปี ความสัมพันธ์เชิงการจัดการโดยพื้นฐานตารางข้อมูล Transaction ต้องมีอย่างน้อยที่สุด 2 ชุดข้อมูลหรือปัจจัยที่ทำการศึกษาคือชุดของปัจจัยหรือข้อมูลที่ผ่านการจัดการ (Transaction ID Field)โดยทั่วไปเป็นการสร้าง Transaction ของปัจจัยซึ่งจะหาความสัมพันธ์ ตัวอย่างเช่น ข้อมูลยอดขายของซุปเปอร์มาเก็ตจะมีข้อมูลที่เป็นจำนวนใบเสร็จของยอดขายเก็บไว้หรือรหัสของสินค้าที่ลูกค้าซื้อสินค้าไปถูก Check out เช่นเดียวกันกับยอดผู้ป่วยที่ได้รับการรักษาในโรงพยาบาลจะมีข้อมูลแพทย์ที่ทำให้บริการรักษากับผู้ป่วยแต่ละรายและชุดของปัจจัยหรือข้อมูลที่ทำการศึกษา (Item ID Field) โดยสร้างข้อมูลหรือปัจจัยซึ่งจะพาความสัมพันธ์ของข้อมูลจากปัจจัยที่ทำการศึกษาโดยกำหนดอยู่รูปของข้อความหรือรหัสในการวิเคราะห์บางอย่างนั้นต้องทำการพิจารณาค่าดังต่อไปนี้ซึ่งจะใช้ในขบวนการวิเคราะห์ค่าต่อไปนี้

1.ค่าความเชื่อมั่น ควรพิจารณาค่าที่มีความเชื่อมั่นอยู่เกณฑ์เฉลี่ยที่รับได้คือ 0.70 ดังนั้นปัจจัยที่ได้มาจากการวิเคราะห์โดยใช้ระบบนี้นั้นควรมากกว่าหรือเท่ากับค่า0.70ค่าความถี่ข้อชุดข้อมูลที่ทำการศึกษานี้ควรมากกว่าหรือเท่ากับ 0.2

2.เงื่อนไขในการหาความสัมพันธ์ของปัจจัยควรใช้น้อยกว่า n ชุด/ความถี่ของชุดข้อมูลโดยมีค่าเริ่มต้นเท่ากับ 3 การเปลี่ยนแปลงค่าจาก 3 – 6จะทำให้เกิดปัญหาของ overhead ในการคิดคำนวน

3.ค่าความกว้างของข้อมูลควรมีค่าเริ่มต้นเท่ากับ 5การเปลี่ยนแปลงค่าจาก5- 10จะทำให้เกิดปัญหาของoverheadในการคิดคำนวน

4.ชุดของปัจจัยที่ค้นหาความสัมพันธ์อาจเกิดการแจกแจงค่าความถี่ที่คาดหวังไว้มากกว่าหรือน้อยกว่าค่านี้เรียกว่า expected frequency distribution โดยค่านี้พิจารณาจากข้อมูลที่สนใจศึกษาเช่นถ้าตั้งไว้ 2.0หมายถึงค่าความคาดหวังนี้น้อยกว่า1/2หรือเท่ากับ 1 พิจารณาต่อว่าในการวิเคราะห์ผลปรากฏว่าค่าที่สนใจพบว่ามีแพทย์ประจำ (ที่มีรหัสน้อยกว่า 6) มีความสัมพันธ์กับ

การติดเชื้อในโรงพยาบาลโดยมีค่าความเชื่อมั่น 0.984

มีความสัมพันธ์ต่อการติดเชื้อแล้วยังสัมพันธ์กับการมีชีวิตของผู้ป่วยหลังได้รับการรักษาโดยมีค่า การแจกแจงความถี่ที่คาดหวัง = 0.96 เป็นต้น

3. ความสัมพันธ์เชิงกรณีเงื่อนไข (Case Association)

วิธีนี้ถูกออกแบบ เพื่อชี้ให้เห็นถึง ความสัมพันธ์ระหว่างตัวแปรที่ทำการศึกษา ตัวแปรที่มีไม่ชัดเจนจะถูกคัดออก การใช้กรณีเชิงความสัมพันธ์นี้ จะจัดกลุ่มของตัวแปร ตามที่ความถี่แจกแจงไว้ ขบวนการความสัมพันธ์เชิงกรณีนี้เป็นการสร้างความสัมพันธ์ของปัจจัยจะถูกสร้างขึ้นอยู่ในรูปแบบของเงื่อนไขซึ่งเงื่อนไขจะประกอบด้วย ตัวแปรเทียบกับค่าของตัวแปรนั้นๆความถี่ เป็นการนับจำนวนครั้งของการจัดกลุ่ม ที่เกิดขึ้นของตัวแปร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าความเชื่อมั่น (Confidence) เป็นค่าความเชื่อถือในเงื่อนไขที่เปรียบเทียบค่าของตัวแปร โดยค่าสูงสุด = 1.0 และ ต่ำสุด = 0.01 ค่าการแจกแจงความถี่ของค่าที่คาดหวัง (Surprise) เป็นชุดของตัวแปรที่ทำการศึกษาโดยกำหนดค่าไว้ให้ค่าสูงสุด=2,ค่าต่ำสุด=0.5ซึ่งหมายความว่าชุดของตัวแปรที่ทำการศึกษาจะต้องมีค่ามากกว่า (0.5) 2 เท่า หรือ น้อยกว่า (2) จำนวนครึ่งหนึ่ง ดังนั้นจึงเท่ากับ 1>45 ความสนใจจะเป็นช่วงอายุซึ่ง ผลลัพธ์ตัวแปรจะเป็นจำนวน ถ้าไม่แบ่งเป็นช่องของอายุ

4.5 ขอบเขตของระบบที่นำมาศึกษา

ระบบ Data Mining ที่นำมาศึกษานี้ มีขอบเขตของระบบดังนี้

1. จำนวน Record ในตารางข้อมูลใส่ข้อมูลได้ 2,000 ล้าน Record
2. จำนวนของคอลัมภ์ในตารางข้อมูลใส่ได้ 1,000 คอลัมภ์
3. มีการจัดการได้ 100 operation ต่อการสั่ง 1 รายการ
4. มีการเชื่อมต่อ 20 การเชื่อมโยงต่อการสั่ง 1 รายการ
5. ชื่อของตัวแปร ที่ทำการศึกษาใส่ได้ 40 ตัวอักษร

4.6 การอภิปรายผลจากขบวนการ Data Mining

การศึกษาผลจากการศึกษาข้อมูลที่น่าจะเป็นปัจจัยต่อความเสี่ยงในการที่ผู้ป่วยจะได้รับติดเชื้อในโรงพยาบาลมีหลายกรณีที่ทำการศึกษา เช่น

1.การศึกษาความสัมพันธ์ระหว่างการติดเชื้อในโรงพยาบาลกับบุคลากรทางการแพทย์

2.การศึกษาความสัมพันธ์ระหว่างการติดเชื้อในโรงพยาบาลกับหอผู้ป่วยที่ผู้ป่วยพักทำการรักษา

3.การศึกษาความสัมพันธ์ระหว่างการติดเชื้อในโรงพยาบาลกับการได้รับยาปฏิชีวนะของผู้ป่วย

4.การศึกษาความสัมพันธ์ระหว่างการติดเชื้อในโรงพยาบาลกับระยะเวลาที่ผู้ป่วยพักรักษาตัวในโรงพยาบาล

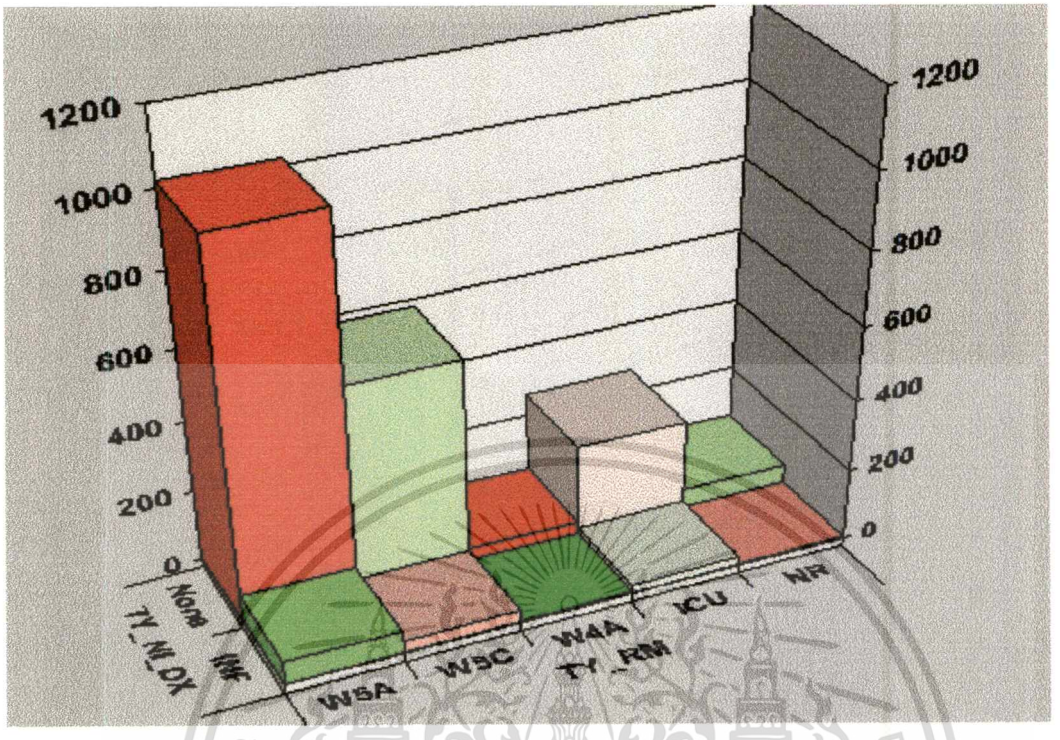
5.การศึกษาความสัมพันธ์ระหว่างการติดเชื้อในโรงพยาบาลกับการจัดกลุ่มแบ่งตามการเฝ้าระวังโรคในโรงพยาบาล

6.การศึกษาความสัมพันธ์ระหว่างการติดเชื้อในโรงพยาบาลกับช่วงอายุผู้ป่วย

7.การศึกษาความสัมพันธ์ระหว่างการติดเชื้อในโรงพยาบาลกับเพศของผู้ป่วย

โดยทั้งหมดที่ดำเนินการศึกษาพบว่ามีความเป็นไปได้ในการที่ผู้ป่วยจะได้รับการติดเชื้อดังรายละเอียดที่จะอธิบายต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

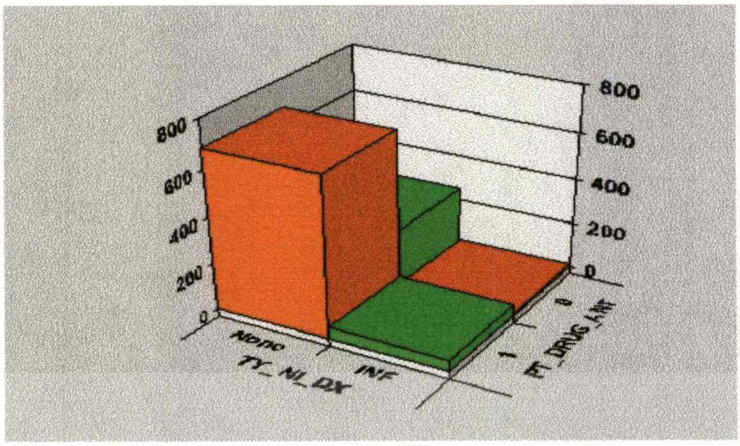


ภาพที่ 12 แสดงรายงานความสัมพันธ์เปรียบเทียบปัจจัยที่ทำการศึกษา โดยนำเสนอในรูปแบบของ 3 มิติ (Field Versus Field) (TY_NI_DX) และ (PT_RM)

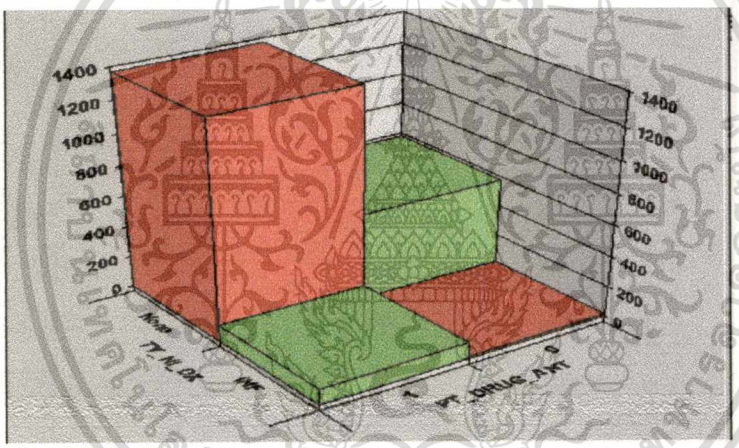
จากภาพที่ 12 เป็นการแสดงความสัมพันธ์ของผู้ป่วยโดยนำเสนอในรูปรายงาน (รายงานผู้ป่วยที่ทำการศึกษามีประมาณ 2,075 record) แบบเปรียบเทียบในรูป 3 มิติระหว่างกลุ่มผู้ป่วยที่ได้รับการติดเชื้อในโรงพยาบาลและไม่ได้รับการติดเชื้อในโรงพยาบาลโดยแยกตามหอพักที่ผู้ป่วยเข้ารับการรักษาทำให้ทราบข้อมูลว่าใน หอผู้ป่วย 5A มีผู้ป่วยรับการบริการที่เป็นกลุ่มเป้าหมายในการศึกษาจำนวนสูงสุดเมื่อเปรียบเทียบหับหอผู้ป่วยอื่นที่รวมทั้งมีจำนวนผู้ป่วยที่ได้รับการติดเชื้อสูงกว่าแผนกอื่นโดยเฉพาะเมื่อเปรียบเทียบกับแผนกผู้ป่วยหนักซึ่งถ้าให้พิจารณาแล้วได้ 2 แนวทางคือ

1. จำนวนผู้ป่วยที่เข้ารับบริการเป็นปัจจัยที่มีผลต่อการติดเชื้อในโรงพยาบาล
2. ควรพิจารณาหาความสัมพันธ์ของการติดเชื้อในโรงพยาบาลกับ W5A

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(TY_NI_DX) และ (PT_DRUG_ANT) (จำนวนrecord = 1,000record)

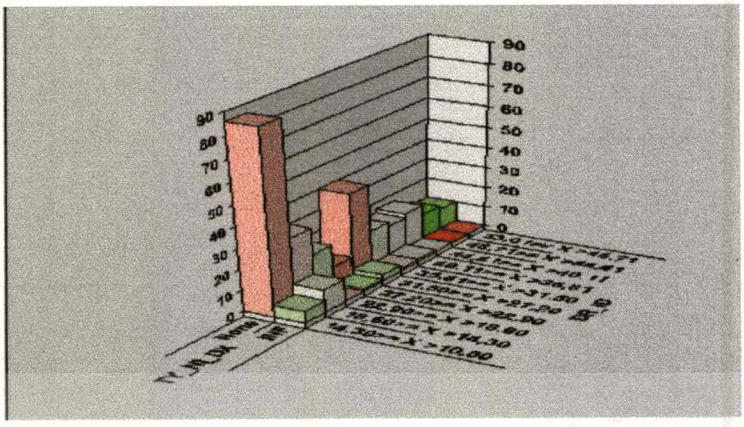


(TY_NI_DX) และ(PT_DRUG_ANT)(จำนวนrecord= 2,075 record)

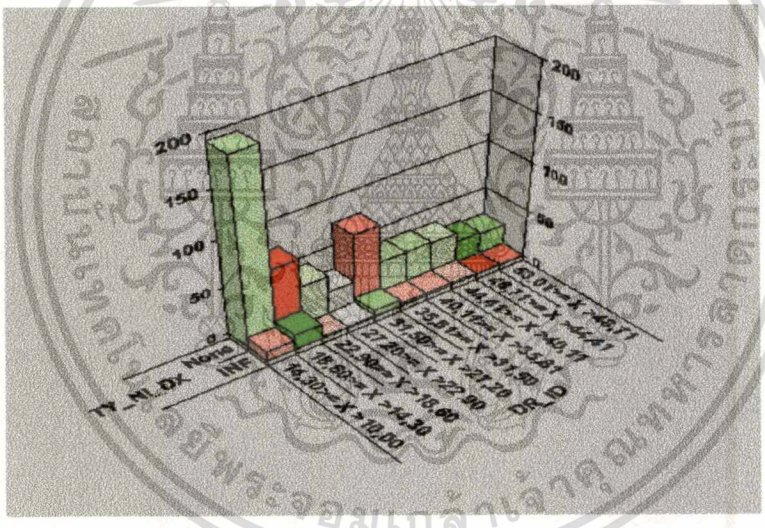
ภาพที่ 14 แสดงรายงานความสัมพันธ์โดยการเปรียบเทียบปัจจัยที่ทำการศึกษา โดยนำเสนอในรูปแบบของ (Field Versus Field)(จำนวนผู้ป่วย 1000 คน/2075คน)

จากภาพที่14 แสดงรูปแบบรายงาน 3 มิติโดยนำเสนอเปรียบเทียบปัจจัยอิสระที่ทำการศึกษาระหว่างกลุ่มผู้ป่วยที่รับบริการ ความสัมพันธ์ระหว่างกลุ่มผู้ป่วยที่ได้รับการติดเชื้อในโรงพยาบาลกับกลุ่มผู้ป่วยที่ได้รับยาปฏิชีวนะในโรงพยาบาลพบว่ากลุ่มผู้ป่วยที่ได้รับยาปฏิชีวนะและเป็นกลุ่มผู้ป่วยที่ไม่ได้รับการติดเชื้อในโรงพยาบาลมีจำนวนสูงกว่าผู้ป่วยที่ได้รับยาปฏิชีวนะและได้รับการติดเชื้อในโรงพยาบาลจึงเป็นข้อสังเกตที่สำคัญและพิจารณาเข้าไปตรวจสอบว่าผู้ป่วยที่ได้รับยาปฏิชีวนะนั้นป่วยด้วยโรคอะไรจึงทำให้มีการใช้ยาปฏิชีวนะสูงในโรงพยาบาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(จำนวนrecord= 1,000 record)



(จำนวน record= 2,075 record)

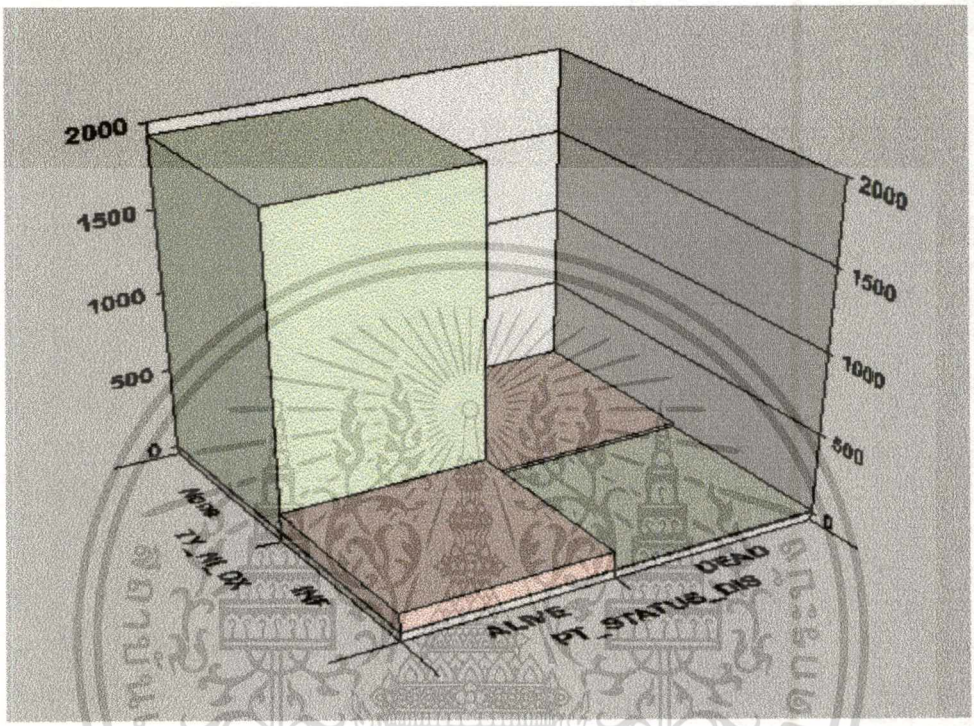
ภาพที่ 15 แสดงรูปแบบรายงาน 3 มิติ โดยนำเสนอเปรียบเทียบปัจจัยที่ทำการศึกษาในบุคลากรทางการแพทย์แต่ละท่านกับการติดเชื้อในโรงพยาบาล (TY_NI_DX) และ (DR_ID)

(เปรียบเทียบจำนวน record 1,000/ 2,075 record)

จากภาพที่15 แสดงรายงานแสดงความความสัมพันธ์การวินิจฉัยการติดเชื้อในโรงพยาบาลสัมพันธ์ปัจจัยด้านบุคลากรที่ให้การรักษาโดยเปรียบเทียบกับจำนวน record 1,000/2,075 record โดยจะพบว่ามีบุคลากรทางการแพทย์ได้ให้การรักษาผู้ป่วยที่ได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาลใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

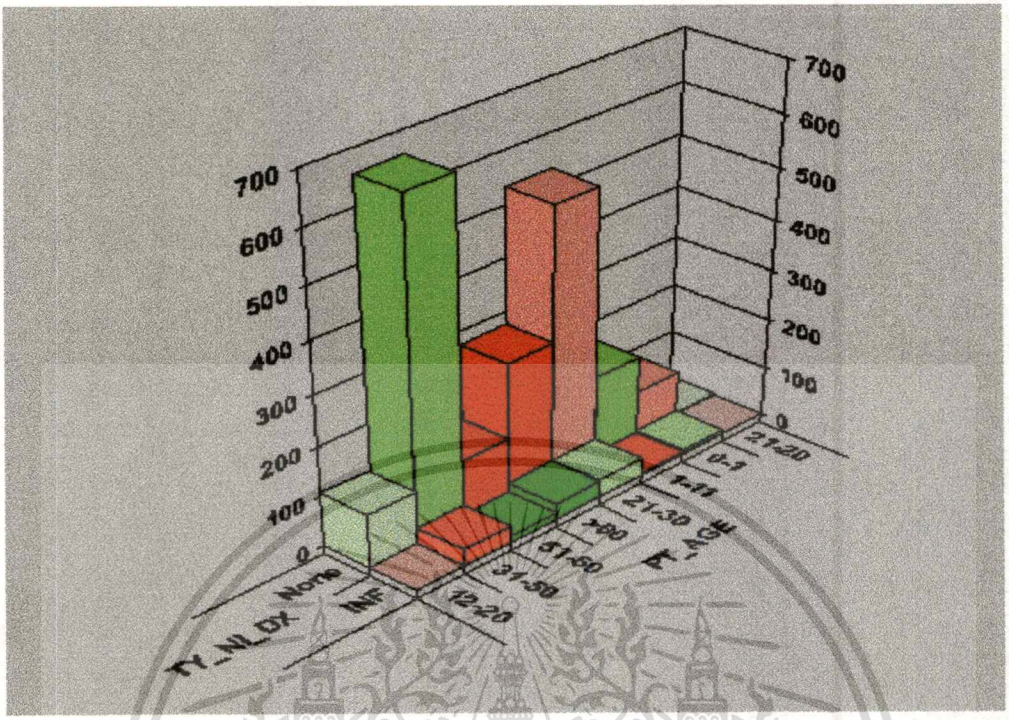
record พบว่าเป็นแพทย์ที่มีรหัสอยู่ระหว่าง 14>x>18 ซึ่งควรพิจารณาโดยใช้เครื่องมือศึกษาดูความสัมพันธ์เชิงกรณีต่อไป



ภาพที่ 16 แสดงรายงานความสัมพันธ์โดยการเปรียบเทียบปัจจัยที่ทำการศึกษา โดยนำเสนอในรูปแบบของ 3 มิติ (Field Versus Field) (TY_NI_DX และ PT_STATUS_DIS)

จากภาพที่ 16 เป็นภาพแสดงความสัมพันธ์ของผู้ป่วยโดยนำเสนอในรูปรายงานแบบอยู่ในรูป 3 มิติ ระหว่างกลุ่มผู้ป่วยที่ได้รับการติดเชื้อในโรงพยาบาลและไม่ได้รับการติดเชื้อในโรงพยาบาล โดยเปรียบเทียบความสัมพันธ์ผลการรักษาผู้ป่วยต่อการวินิจฉัยการติดเชื้อในโรงพยาบาลพบว่าอัตราการตายของผู้ป่วยที่รับไว้รักษาในโรงพยาบาล/อัตราการติดเชื้อในโรงพยาบาลมีน้อยกว่าดังนั้นขั้นตอนต่อไปคือการหาค่าของความสัมพันธ์ดังกล่าวโดยใช้เครื่องมือของระบบ Data Mining

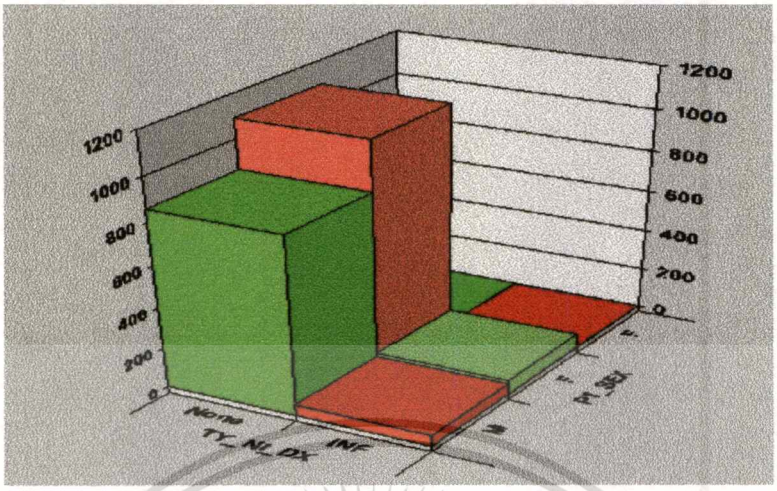
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 17 แสดงรายงานความสัมพันธ์โดยการเปรียบเทียบปัจจัยที่ทำการศึกษา โดยนำเสนอในรูปแบบของ 3 มิติ (Field Versus Field) (PT_NI_DX และ PT-AGE)

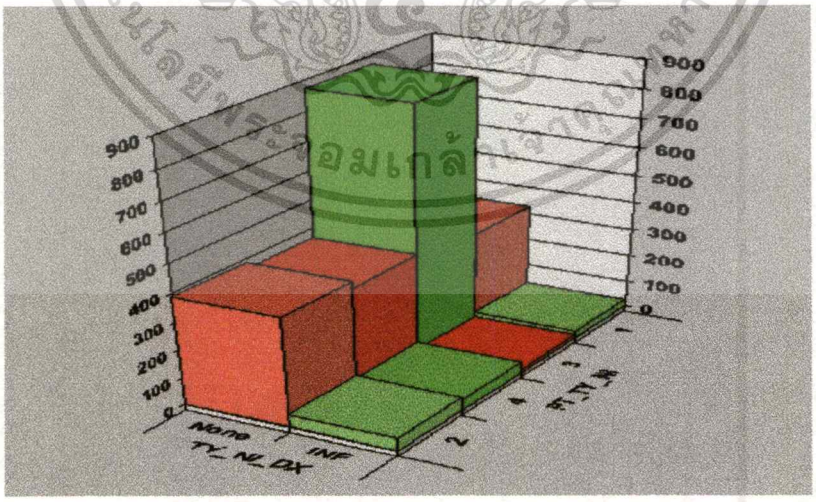
ภาพที่ 17 เป็นภาพแสดงรายงานรูปแบบ 3 มิติเปรียบเทียบให้เห็นถึงความสัมพันธ์ของกลุ่มผู้ป่วยที่ทำการศึกษาการติดเชื้อในโรงพยาบาลกับอายุของผู้ป่วยที่ทำการศึกษาซึ่งจะพบว่าจากการศึกษากลุ่มผู้ป่วยที่ทำการเฝ้าระวังโรคที่มีอายุระหว่าง (21-30)(31-50)(>60) ปีและมีการติดเชื้อในโรงพยาบาลมีอัตราสูงกว่ากลุ่มอายุผู้ป่วยอื่นที่ทำการศึกษาซึ่งตามทฤษฎีกล่าวว่ากลุ่มผู้ป่วยที่มีโอกาสติดเชื้อในโรงพยาบาลสูงจะอยู่ในกลุ่มผู้สูงอายุและเด็กแรกเกิดเนื่องจากมีภูมิคุ้มกันต่ำกว่ากลุ่มผู้ป่วยกลุ่มอื่น จึงเป็นข้อสังเกตที่จะนำไปสู่การวิเคราะห์ค่าความสัมพันธ์เชิงกรณีต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 18 แสดงรายงานความสัมพันธ์โดยการเปรียบเทียบปัจจัยที่ทำการศึกษา โดยนำเสนอในรูปแบบของ 3 มิติ (Field Versus Field) (TY_NI_DX และ PT_SEX)

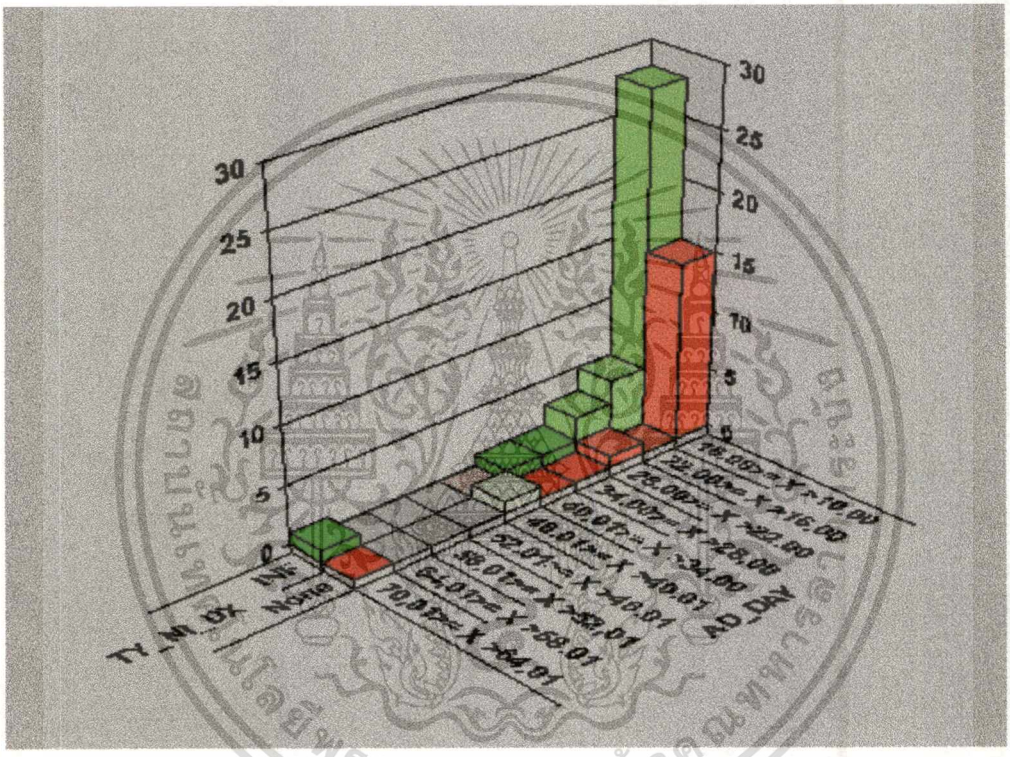
ภาพที่ 18 เป็นภาพแสดงรายงานรูปแบบ 3 มิติโดยเปรียบเทียบจำนวนกลุ่มผู้ป่วยที่ทำการศึกษาเกี่ยวกับการวินิฉัยการติดเชื้อในโรงพยาบาล กับเพศของกลุ่มผู้ป่วยที่ทำการศึกษา ซึ่งจะพบว่ากลุ่มผู้ป่วยเพศหญิงที่ทำการศึกษา ได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาลมากกว่าเพศชาย



ภาพที่ 19 แสดงรายงานความสัมพันธ์โดยการเปรียบเทียบปัจจัยที่ทำการศึกษา โดยนำเสนอในรูปแบบของ 3 มิติ (Field Versus Field)(TY-NI_DX)(PT_TY_NI)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

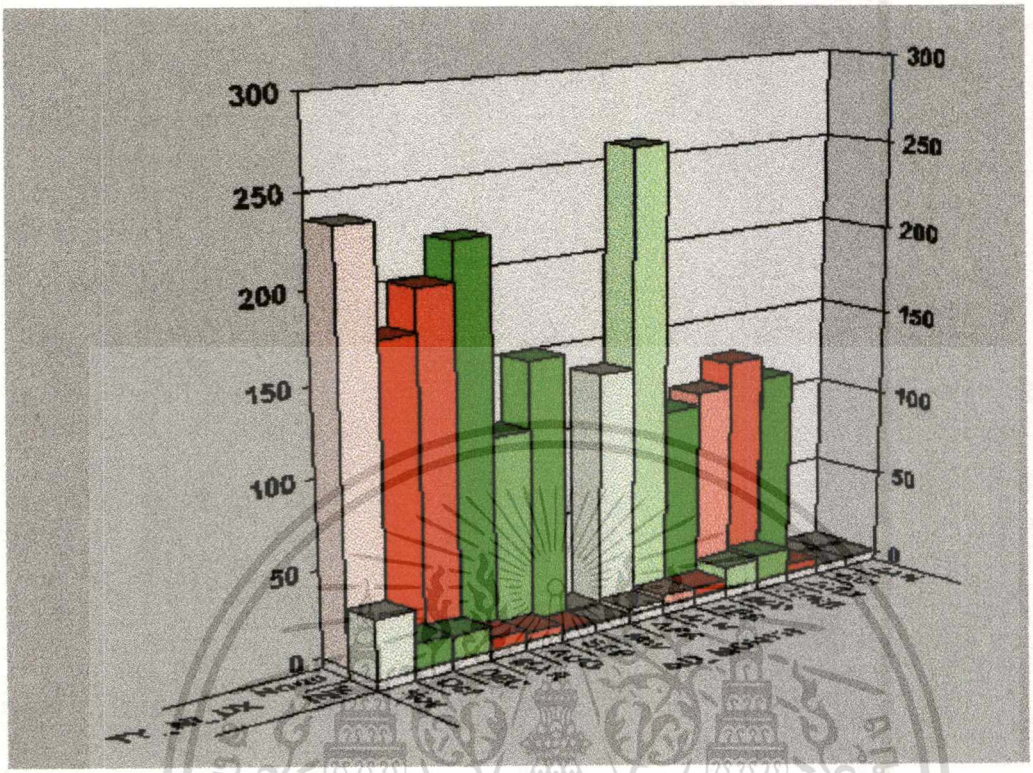
จากภาพที่ 19 เป็นภาพแสดงรายงานรูปแบบ 3 มิติแสดงให้เห็นถึงความสัมพันธ์เมื่อเทียบกันในกลุ่มผู้ป่วยที่ทำการเฝ้าระวังโรคติดเชื้อในโรงพยาบาลซึ่งจะสังเกตว่ากลุ่มผู้ป่วยที่มีการติดเชื้อในโรงพยาบาลในกลุ่มที่ 1 กลุ่มที่2 กลุ่มที่ 4 มีความใกล้เคียงกันในด้านการติดเชื้อในโรงพยาบาลโดยกลุ่มที่3 พบว่ามีการติดเชื้อน้อยกว่ากลุ่มอื่นซึ่งควรมีการพิจารณาค้นหาว่ามีปัจจัยอะไรที่สัมพันธ์กันแล้วทำให้มีการติดเชื้อในโรงพยาบาลน้อยเนื่องจากในทางทฤษฎีแล้วควรที่จะมีการติดเชื้อสูงกว่า



ภาพที่ 20 แสดงรายงานความสัมพันธ์โดยการเปรียบเทียบปัจจัยที่ทำการศึกษา โดยนำเสนอในรูปแบบของ 3 มิติ (Field Versus Field)

ภาพที่ 20 แสดงรูปแบบ รายงาน 3 มิติให้เห็นถึงความสัมพันธ์เมื่อเทียบกับจำนวนของกลุ่มผู้ป่วยที่ทำการศึกษาการติดเชื้อในโรงพยาบาลกับจำนวนวันที่ผู้ป่วยแต่ละรายเข้ารับการรักษาในโรงพยาบาลพบว่าจำนวนวันที่พักรักษาตัวระหว่างมีอัตราการติดเชื้อในโรงพยาบาลสูงจะอยู่ในช่วงระหว่างวันที่นอนในโรงพยาบาล คือช่วง $16 > x > 10$

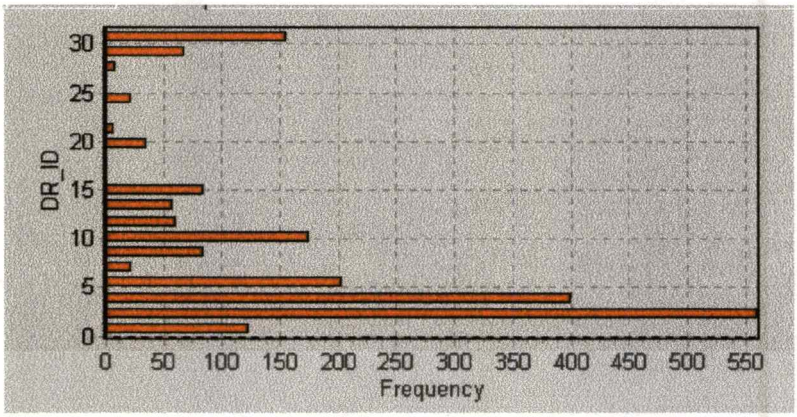
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 21 แสดงรายงานความสัมพันธ์เปรียบเทียบปัจจัยที่ทำการศึกษา
โดยนำเสนอในรูปแบบของ 3 มิติ (Field Versus Field)
(TY_NI_DXและAD_MONTH)

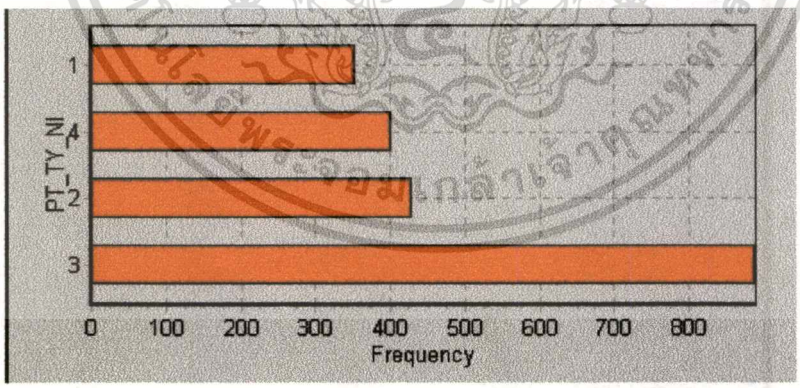
ภาพที่ 21 แสดงรูปแบบรายงาน3 มิติให้เห็นถึงความสัมพันธ์การเปรียบเทียบ กลุ่มผู้ป่วยที่มีการติดเชื้อในโรงพยาบาลกับเดือนที่ผู้ป่วยรับไว้ในโรงพยาบาลจะสังเกตว่าเดือนที่ผู้ป่วยมีการติดเชื้อในโรงพยาบาลสูงคือเดือนพฤศจิกายนเมื่อเปรียบเทียบกับเดือนมกราคมที่มีจำนวนผู้ป่วยที่ทำการเฝ้าระวังใกล้เคียงกันซึ่งควรมีการพิจารณาหาความสัมพันธ์เชิงกรณีต่อเพื่อเป็นแนวทางในการทำการวิเคราะห์การติดเชื้อในโรงพยาบาล การพิจารณากราฟ 3 มิติจะช่วยให้เกิดความเข้าใจโดยง่าย และในเดือนพฤศจิกายนปีถัดไปควรมีการเข้าไปสอบสวนโรคในเดือนนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 22 แสดงรายงานความถี่ (Frequency) ของแพทย์ที่รับรักษาผู้ป่วย (DR_ID)

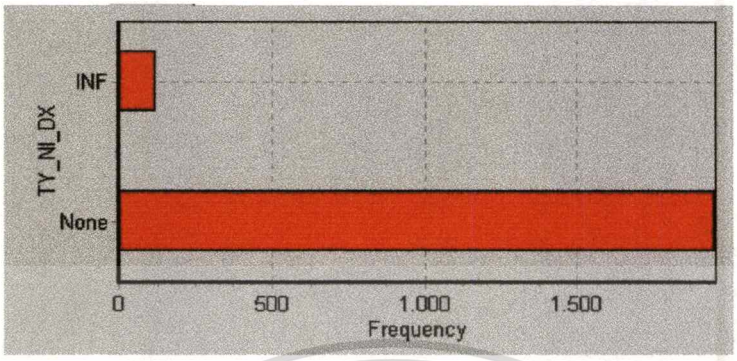
จากภาพที่ 22 แสดงกราฟแท่งแสดงรายงานความถี่ของแพทย์ที่ทำการรักษาผู้ป่วยจากจำนวน Record ของผู้ป่วย 2,075 Record จะพบความถี่ของแพทย์แต่ละท่านแตกต่างกันซึ่งควรพิจารณาความสัมพันธ์ของการติดเชื้อในโรงพยาบาลเพื่อค้นหาความสัมพันธ์พบว่าแพทย์บางท่านมีผู้ป่วยที่ให้การรักษาในจำนวนที่มากกว่าแพทย์ท่านอื่นโดยเฉพาะแพทย์ประจำโรงพยาบาล (รหัส < 6) ที่เป็นแพทย์เฉพาะทางอายุรกรรมและศัลย์กรรม



ภาพที่ 23 แสดงความถี่ของกลุ่มผู้ป่วยที่ทำการศึกษาการเฝ้าระวังโรคในโรงพยาบาล

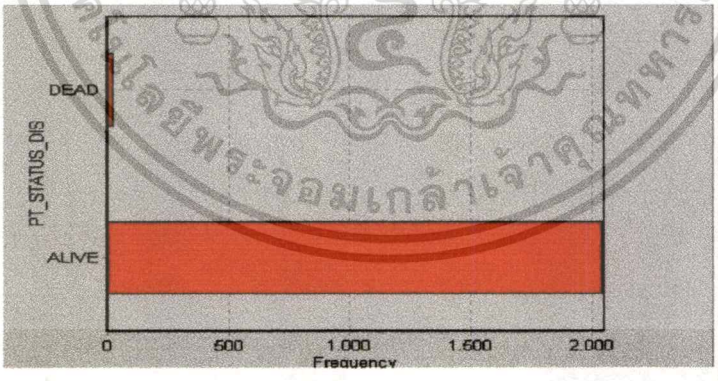
จากภาพที่ 23 แสดงรายงานความถี่ของกลุ่มผู้ป่วยที่ทำการศึกษาการเฝ้าระวังโรคในโรงพยาบาล ตั้งแต่ วันที่ 1 ม.ค. 42 –31 ธ.ค.42 จะพบว่ามีกลุ่มผู้ป่วยที่ทำการเฝ้าระวังโรคประเภท ผู้ป่วยได้รับการใส่ Instrument (3)มีความถี่สูงกว่ากลุ่มผู้ป่วยประเภทอื่นซึ่งมีความถี่ใกล้เคียงกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 24 แสดงความถี่ของการวินิจฉัยการติดเชื้อในโรงพยาบาล(TY_NI_DX) ในกลุ่มผู้ป่วยที่ทำการศึกษาจากจำนวน Recordของผู้ป่วย 2,075 Record

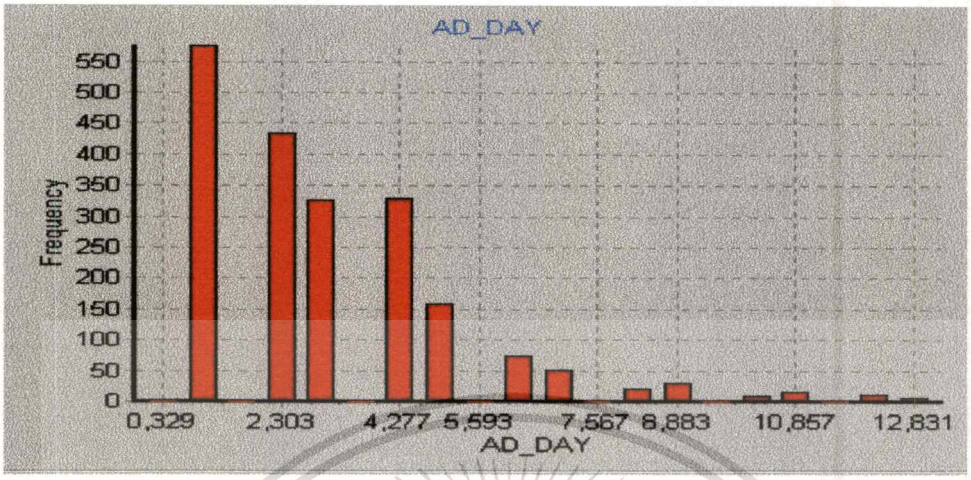
จากภาพที่ 24 เป็นรายงานแสดงความถี่การวินิจฉัยการติดเชื้อในโรงพยาบาลตั้งแต่วันที่ 1ม.ค.42–31ธ.ค.42 จะพบว่าความถี่ของการติดเชื้อในโรงพยาบาลเปรียบเทียบกับความถี่ของการวินิจฉัยว่าไม่การติดเชื้อในโรงพยาบาลมีความถี่น้อยกว่า



ภาพที่ 25 แสดงรายงานความถี่ (Frequency) ของผลของการรักษากลุ่มผู้ป่วย ที่ทำการศึกษาจำนวนผู้ป่วย 2,075 คน (PT_STATUS_DIS)

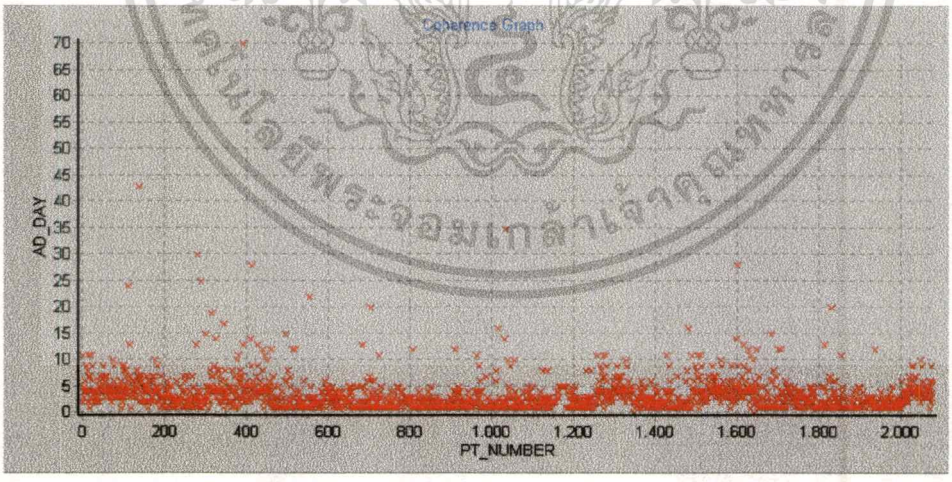
จากภาพที่ 25 แสดงความถี่ของกลุ่มผู้ป่วยที่ทำการศึกษาการเฝ้าระวังโรคติดเชื้อในโรงพยาบาล พบว่ามีความถี่ของผู้ป่วยที่ถึงแก่กรรม =29 Record (จากจำนวนRecord2,075Record) เปรียบเทียบกับจำนวนกลุ่มผู้ป่วยที่มีชีวิต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 26 แสดงรายงานความถี่ของจำนวนผู้ป่วยต่อจำนวนวันที่เข้ารับการรักษาในโรงพยาบาลระหว่าง 1 ม.ค. 42 – 31 ม.ค. 42

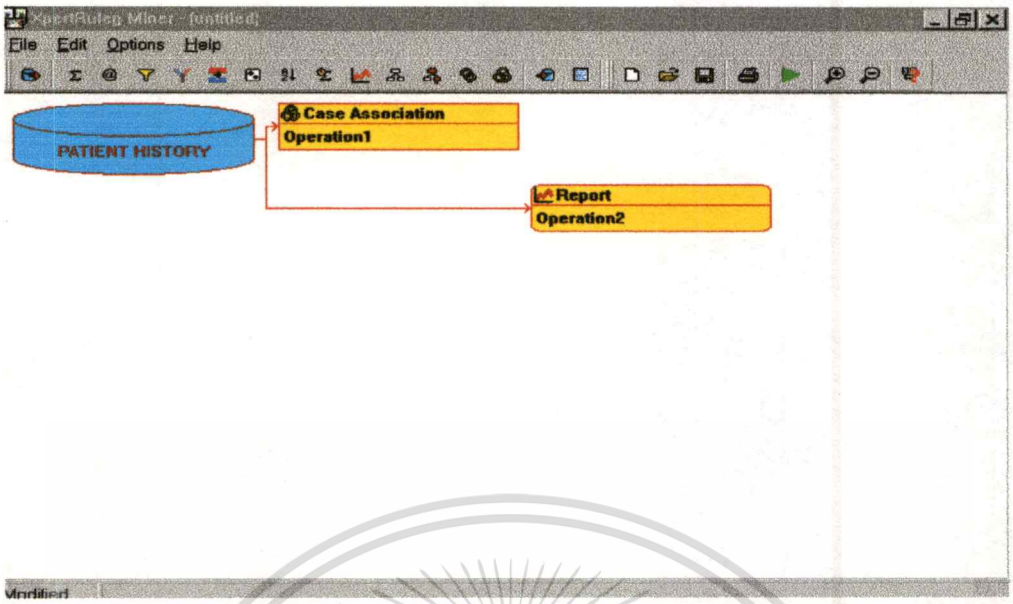
จากภาพที่ 26 แสดงความถี่ของผู้ป่วยต่อจำนวนวันที่นอนในโรงพยาบาลซึ่งมีข้อสังเกตว่าผู้ป่วยที่นอนในโรงพยาบาลพิจารณาจำนวนวันที่นอนที่มีความถี่สูงจะอยู่ในช่วง 0.329 – 4.277 วัน



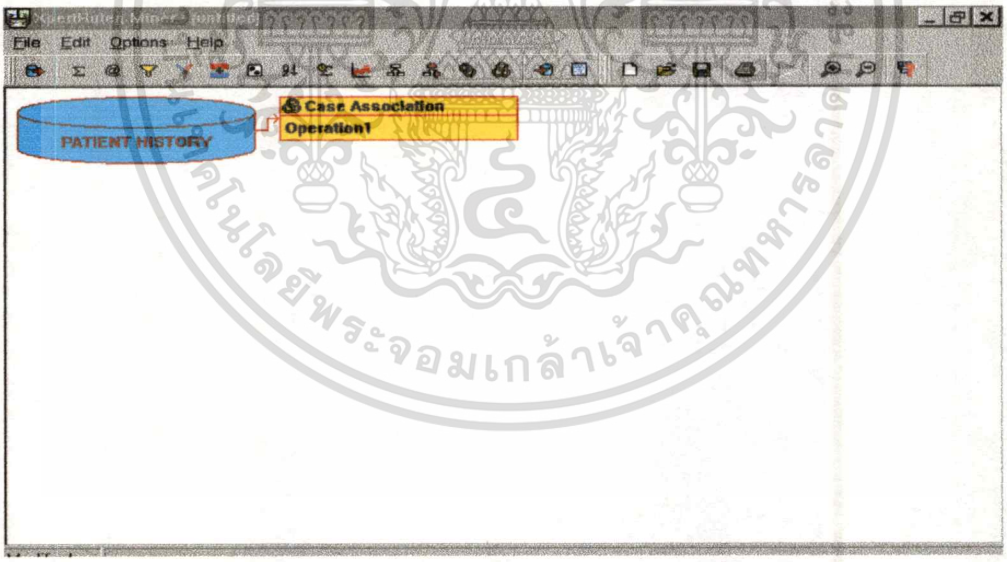
ภาพที่ 27 ภาพแสดงรายงานการจุดความถี่ระหว่างผู้ป่วยแต่ละรายกับจำนวนวันที่นอนในโรงพยาบาล (จำนวน Record 2,075 Record)

จากภาพที่ 27 พบว่าความหนาแน่นของจำนวนผู้ป่วยที่นอนในโรงพยาบาลอยู่ในช่วงเวลา 1-5 วันซึ่งสอดคล้องกับภาพที่ 26 ที่แสดงความถี่ของจำนวนผู้ป่วยต่อจำนวนวันที่ผู้ป่วยนอนในโรงพยาบาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

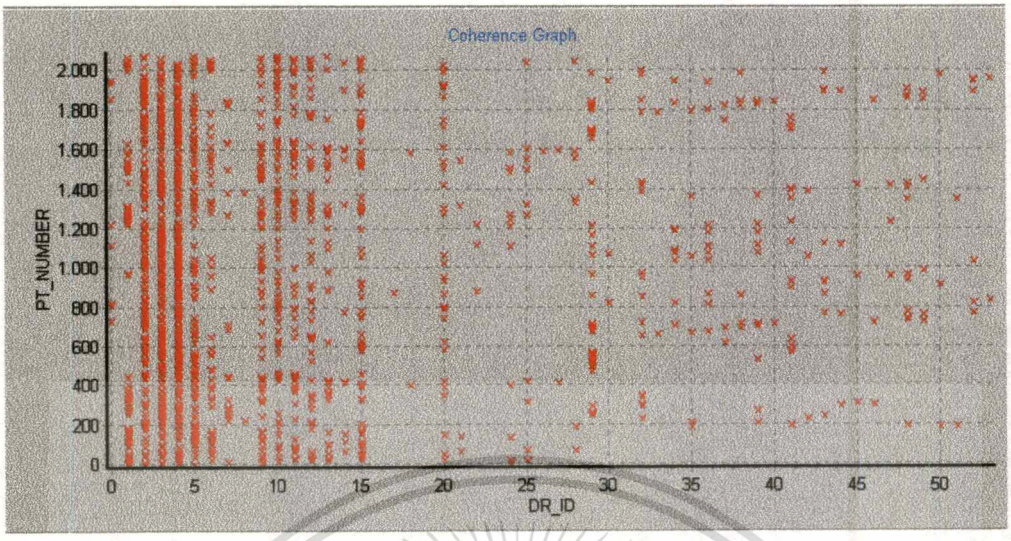


ภาพที่28 แสดงรูปแบบของระบบในการออกรายงานการวิเคราะห์ข้อมูล



ภาพที่ 29 ภาพแสดงรูปแบบของระบบในการสร้างความสัมพันธ์เชิงกรณีที่ใช้ในการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



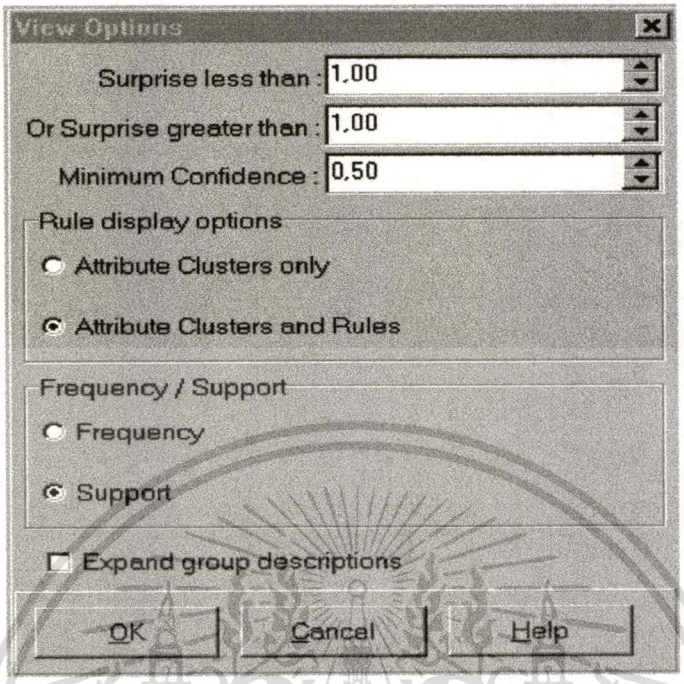
ภาพที่ 30 ภาพแสดงรายงานการจุดความถี่จำนวนผู้ป่วยต่อแพทย์ที่ทำการรักษาผู้ป่วย (จำนวน Record 2,075 Record)

จากภาพที่ 30 เป็นภาพแสดงรายงานการจุดความถี่จำนวนผู้ป่วยต่อแพทย์ที่ทำการรักษาผู้ป่วยมีข้อสังเกตว่าแพทย์ที่ทำการรักษาผู้ป่วยที่เฝ้าระวังโรคมีความหนาแน่นสูงจะอยู่ในช่วงรหัสระหว่าง1- 6 ซึ่งนำไปสู่การวิเคราะห์ความสัมพันธ์ระหว่างการติดเชื้อของผู้ป่วยและแพทย์ที่ทำการรักษา

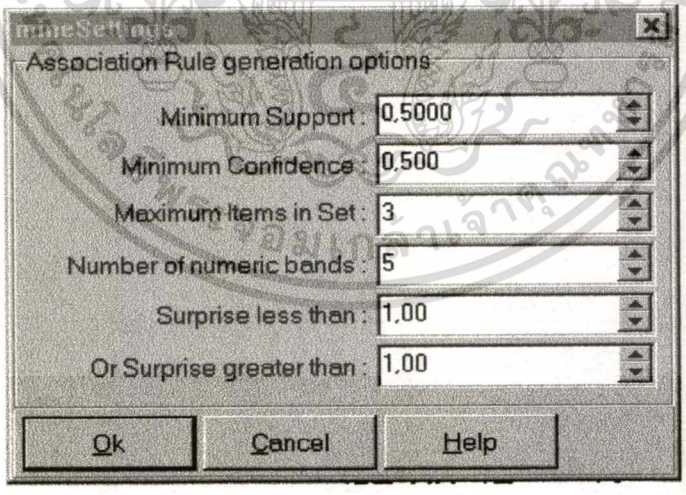
	PT_NUMBER	PT_HN	PT_AN	PT_AGE	PT_SEX	TY_RM	PT_DEP	DR_ID	AD_MONTH	PT_TY_NI	TY_NI_DX
1	1	42-35182	42-4863	12-20	M	W5A	Orthro	24	Nov	2	None
2	2	42-34413	42-4698	31-50	F	W5A	Surg	5	Nov	4	None
3	3	42-25449	42-4686	31-50	M	W5C	Med	2	Nov	4	None
4	4	42-33040	42-4558	12-20	M	W5A	Med	1	Oct	2	None
5	5	42-33928	42-4589	31-50	M	W5A	Orthro	9	Oct	4	None
6	6	42-33875	42-4586	51-60	M	W5C	Med	1	Oct	2	INF
7	7	42-34088	42-4627	12-20	F	W5C	Orthro	9	Nov	4	None
8	8	42-33761	42-4563	51-60	M	W5A	Med	1	Oct	2	None
9	9	42-34120	42-4633	>60	F	W5A	Med	3	Nov	2	None
10	10	40-02794	42-4625	>60	M	W5A	Med	4	Nov	2	None
11	11	42-17841	42-4702	21-30	F	W5A	Med	2	Nov	2	INF
12	12	42-21834	42-4676	21-30	F	W5A	Med	1	Nov	2	None
13	13	42-34107	42-4629	1-11	M	W5A	Ped	10	Nov	2	INF
14	14	42-34497	42-4707	21-30	F	W5C	Orthro	9	Nov	3	None
15	15	42-3134	42-4684	31-50	M	W5C	Med	3	Nov	2	None
16	16	42-32739	42-4593	31-50	F	W5C	NS	7	Oct	4	INF
17	17	42-34546	42-4723	21-30	M	W5C	Orthro	9	Nov	4	None
18	18	40-03073	42-4682	21-30	M	W5A	Med	1	Nov	2	None
19	19	42-34387	42-4692	>60	M	W5A	Surg	25	Nov	4	None
20	20	42-16525	42-4731	21-30	M	W5C	Orthro	24	Nov	2	None

ภาพที่ 31 ภาพแสดงรายงานตัวอย่างตารางข้อมูลผู้ป่วยที่ทำการวิเคราะห์ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

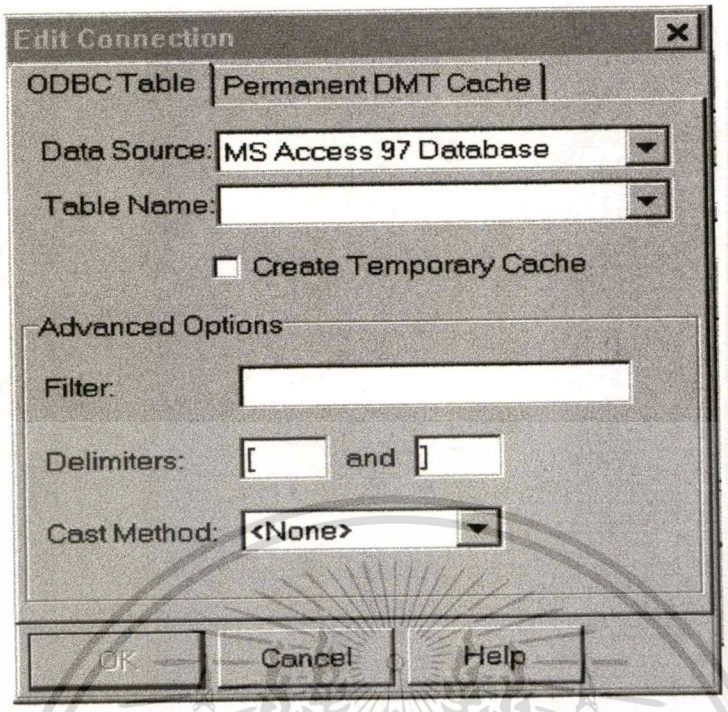


ภาพที่ 32 ภาพแสดงการเลือกขอบเขตค่าในการวิเคราะห์ข้อมูลแบบไม่จัดกลุ่ม

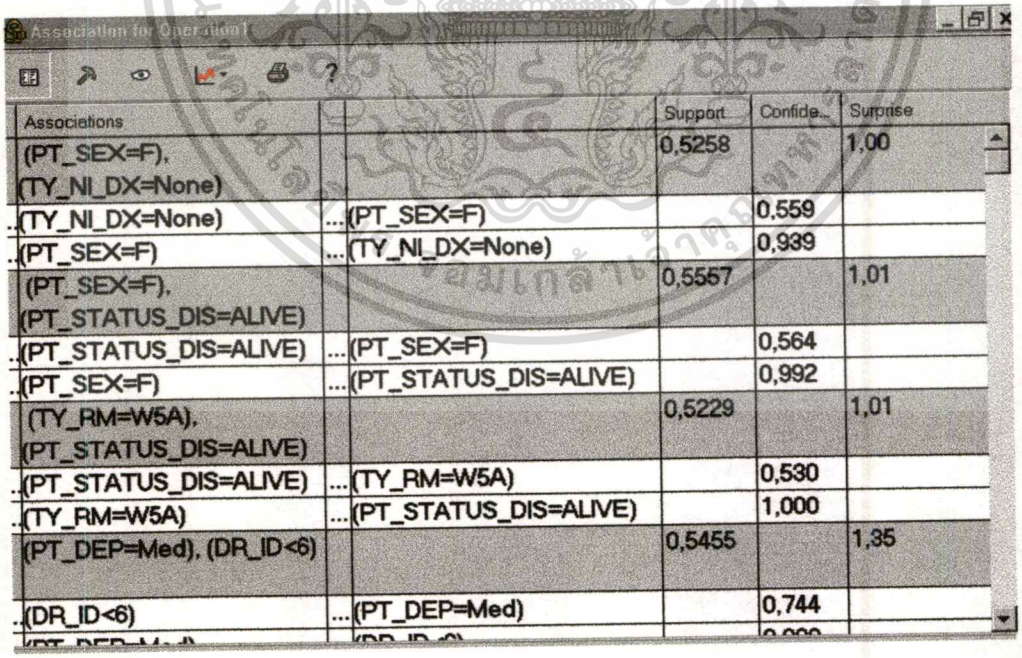


ภาพที่ 33 ภาพแสดงการกำหนดค่าของการเลือกกฎในการสร้างความสัมพันธ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



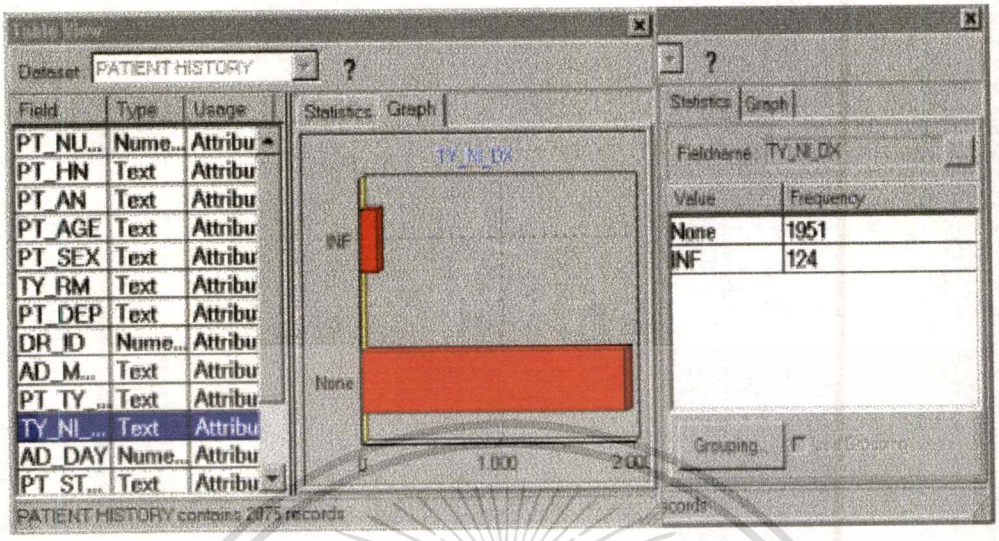
ภาพที่ 34 ภาพขยายการเชื่อมโยงข้อมูลโดยใช้ Microsoft Access เป็นแหล่งข้อมูลในการวิเคราะห์



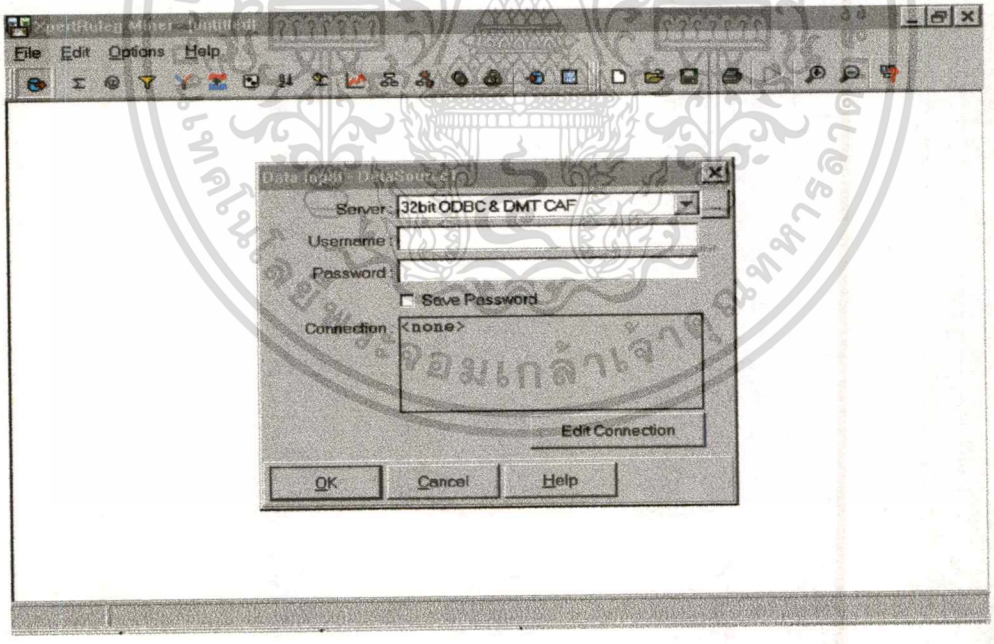
Associations			Support	Confide...	Surprise
(PT_SEX=F), (TY_NI_DX=None)			0,5258		1,00
(TY_NI_DX=None)	...	(PT_SEX=F)		0,559	
(PT_SEX=F)	...	(TY_NI_DX=None)		0,939	
(PT_SEX=F), (PT_STATUS_DIS=ALIVE)			0,5567		1,01
(PT_STATUS_DIS=ALIVE)	...	(PT_SEX=F)		0,564	
(PT_SEX=F)	...	(PT_STATUS_DIS=ALIVE)		0,992	
(TY_RM=W5A), (PT_STATUS_DIS=ALIVE)			0,5229		1,01
(PT_STATUS_DIS=ALIVE)	...	(TY_RM=W5A)		0,530	
(TY_RM=W5A)	...	(PT_STATUS_DIS=ALIVE)		1,000	
(PT_DEP=Med), (DR_ID<6)			0,5455		1,35
(DR_ID<6)	...	(PT_DEP=Med)		0,744	

ภาพที่ 35 แสดงผลการวิเคราะห์โดยใช้ความสัมพันธ์เชิงกรณี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

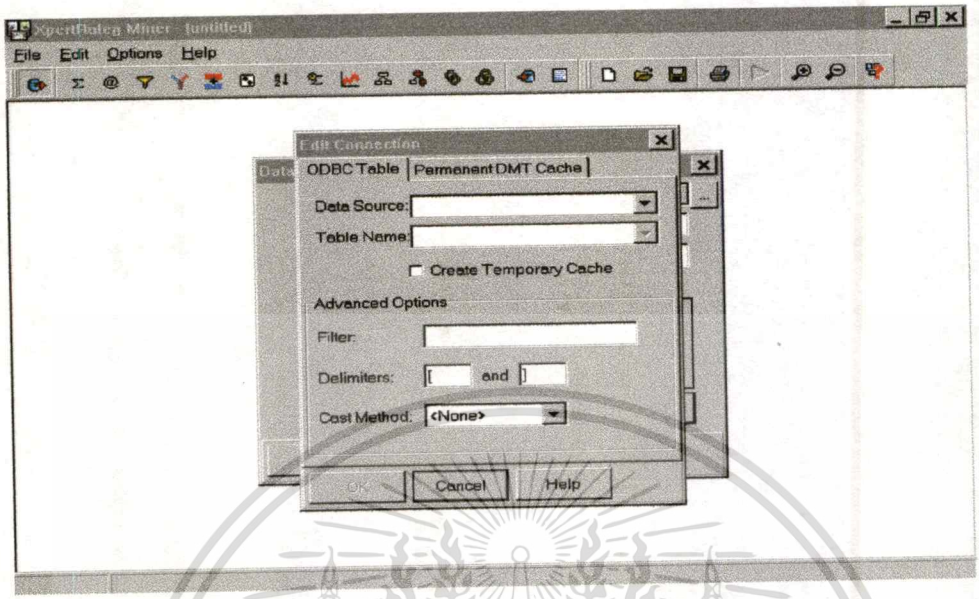


ภาพที่ 36 ภาพแสดง Table View และ กราฟของ TY_NI_DX ที่ทำการวิเคราะห์

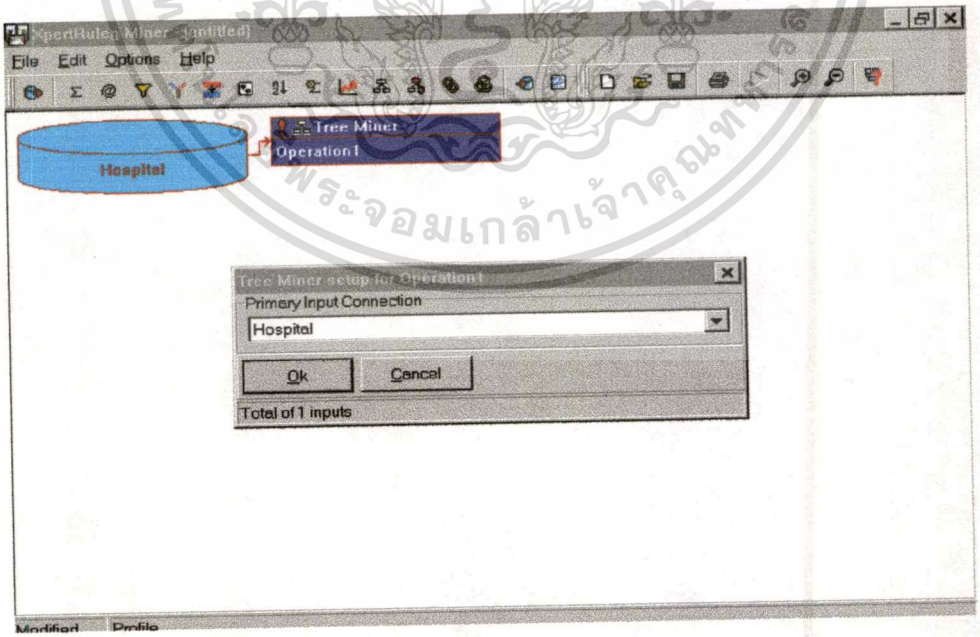


ภาพที่ 37 ภาพแสดงขั้นตอนในการเข้าระบบเพื่อเชื่อมต่อกับฐานข้อมูลผู้ป่วย ก่อนจะเข้าระบบการวิเคราะห์ข้อมูลขั้นตอนที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

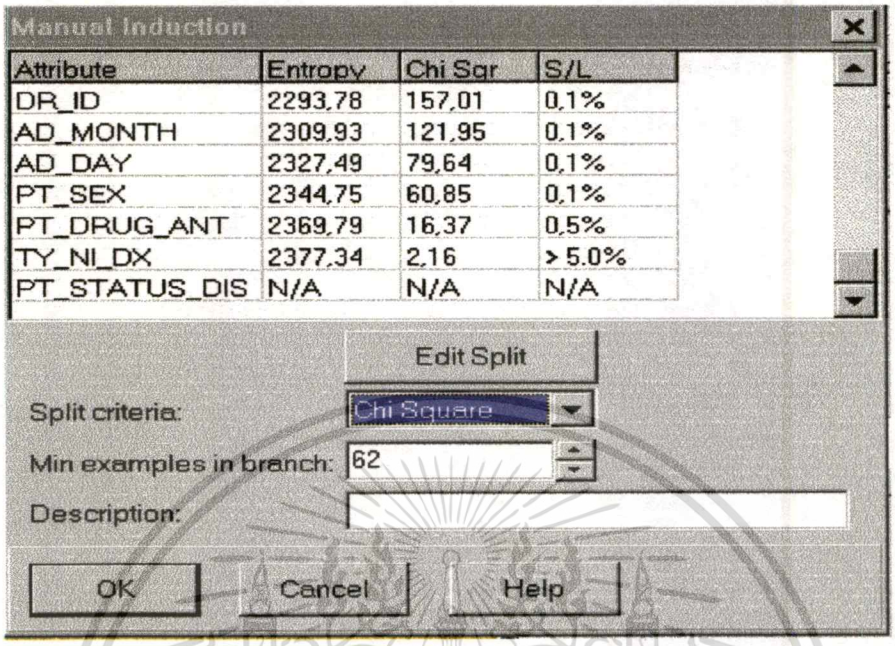


ภาพที่ 38 ภาพแสดงขั้นตอนในการเข้าระบบเพื่อเชื่อมต่อกับฐานข้อมูลผู้ป่วย ก่อนจะเข้าระบบการวิเคราะห์ข้อมูลขั้นตอนที่ 2 โดยใช้ ODBC

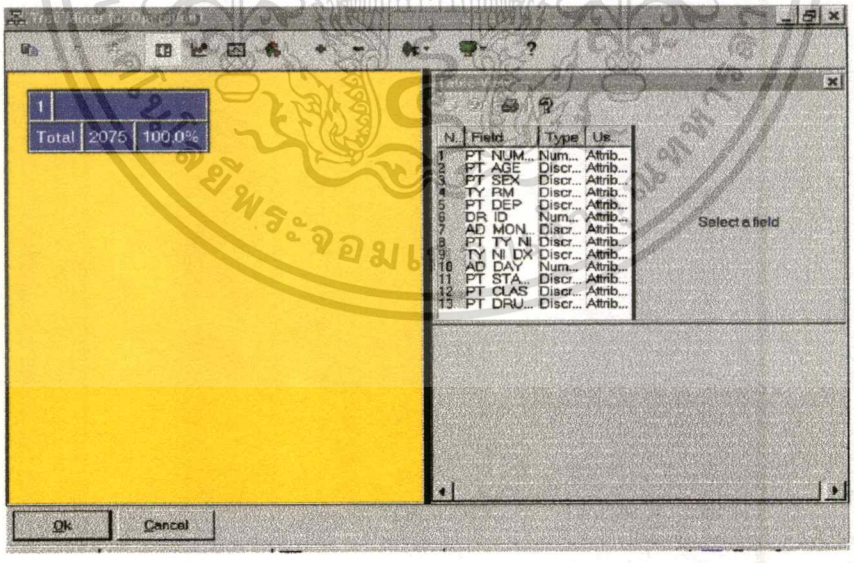


ภาพที่ 39 ภาพแสดงการเข้าระบบเพื่อการวิเคราะห์แบบโครงสร้างต้นไม้โดย (Decision Trees) โดยใช้ Tree Miner

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

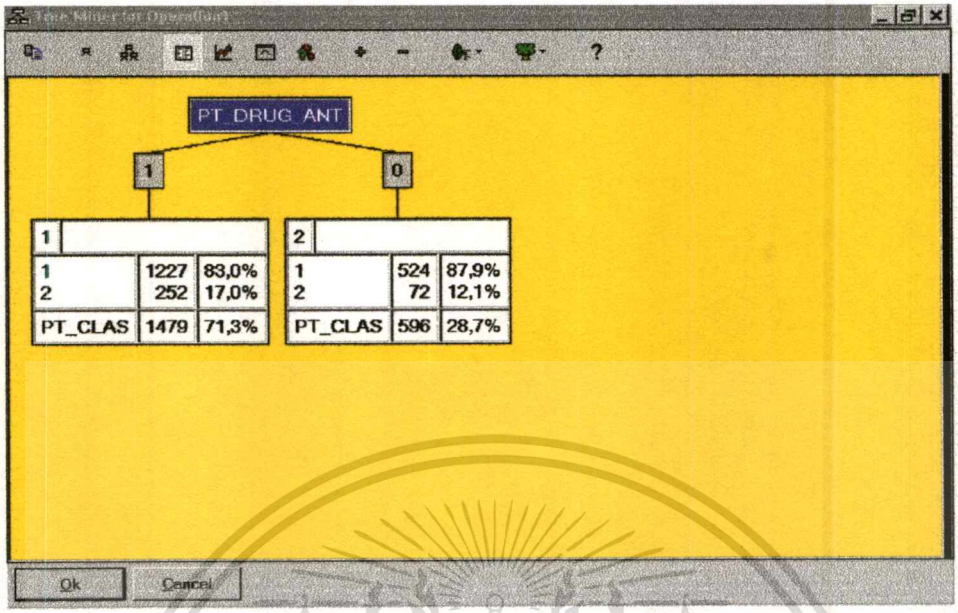


ภาพที่ 40 ภาพแสดงการคัดเลือกกิ่งที่จะทำการศึกษาโดยแสดงหัวข้อให้เลือก

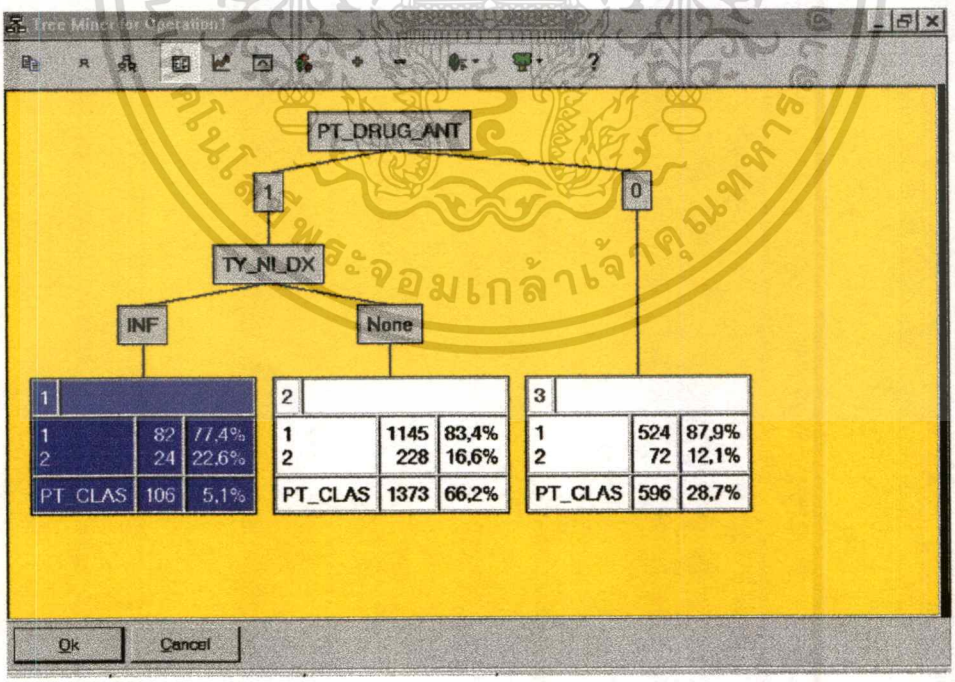


ภาพที่ 41 ภาพแสดง Node แรกที่ทำคัดเลือกเพื่อการแตกกิ่งของโครงสร้างต้นไม้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

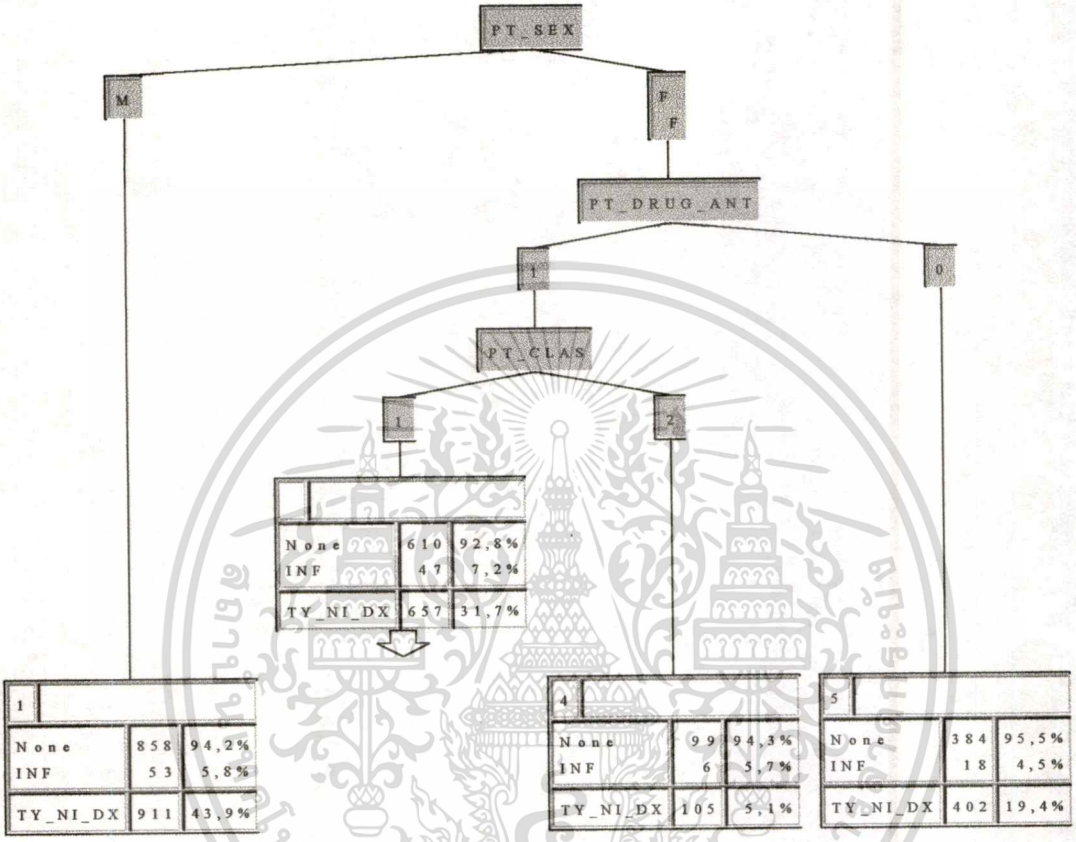


ภาพที่ 42 ภาพแสดงการแตกกิ่งโครงสร้างต้นไม้คัดเลือกประเภทของผู้ป่วย และกลุ่มผู้ป่วยที่ได้รับยาปฏิชีวนะ



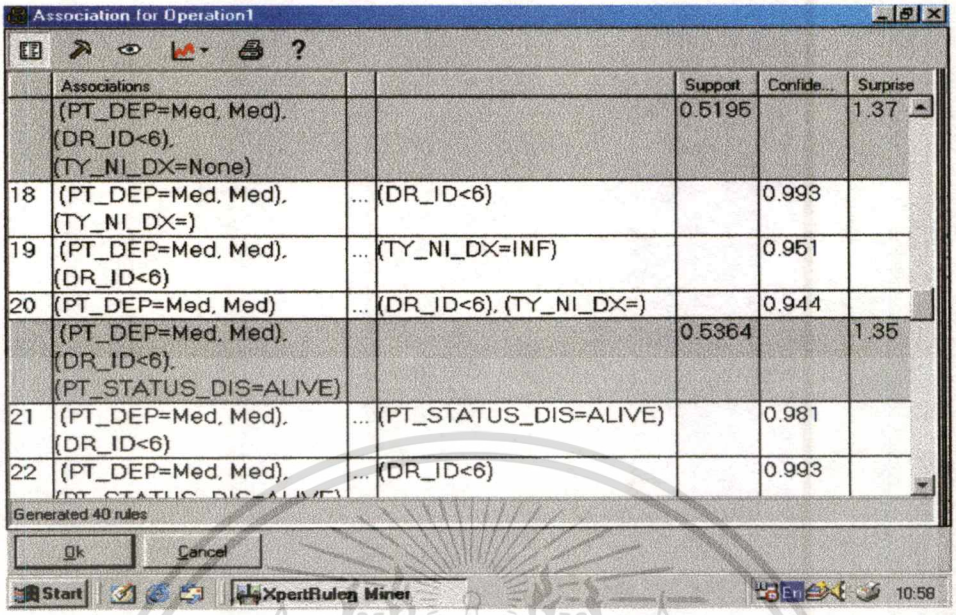
ภาพที่ 43 ภาพแสดงการภาพแตกกิ่งโครงสร้างต้นไม้คัดเลือกประเภทของผู้ป่วย การวินิจฉัยการติดเชื้อในโรงพยาบาลของผู้ป่วย และกลุ่มผู้ป่วยที่ได้รับยาปฏิชีวนะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

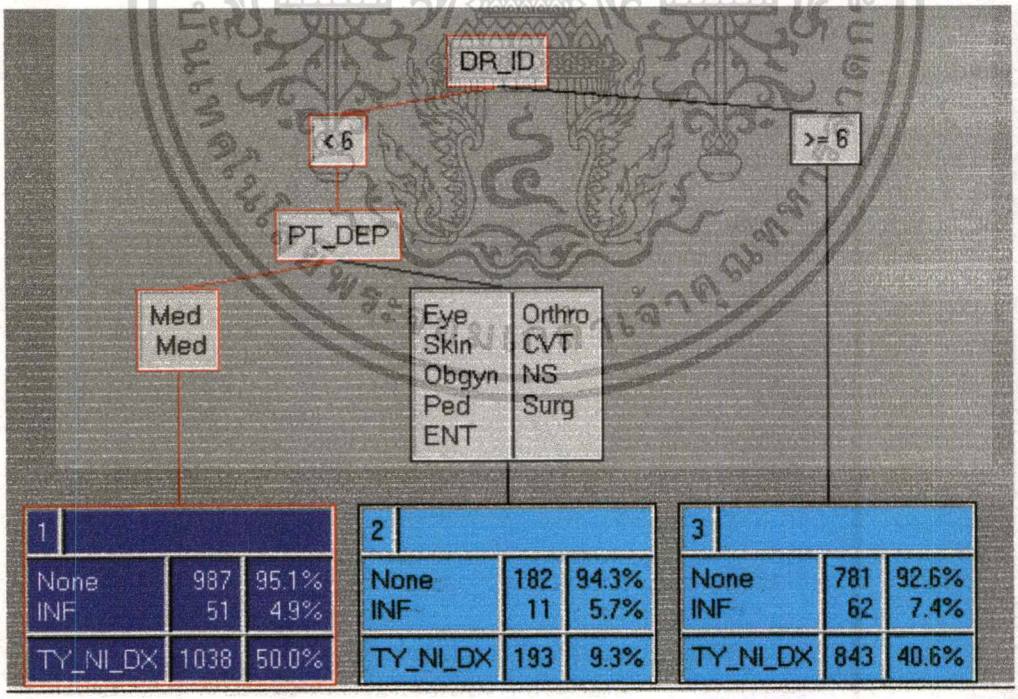


ภาพที่ 44 ภาพแสดงการแตกกิ่งของกลุ่มผู้ป่วยที่ทำการศึกษาการเฝ้าระวังโรคโดยทำการวิเคราะห์ข้อมูลเกี่ยว เพศ การได้รับยาปฏิชีวนะ การแบ่งกลุ่มผู้ป่วยที่เข้ารับบริการ การวินิจฉัยการติดเชื้อในโรงพยาบาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



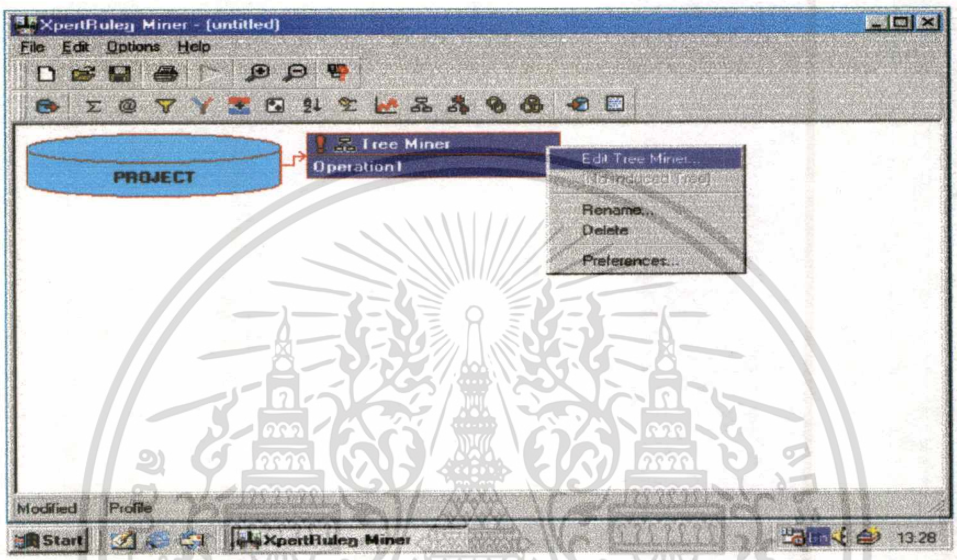
ภาพที่ 45 แสดงค่าความสัมพันธ์เชิงเงื่อนไขที่วิเคราะห์ (Case Association)



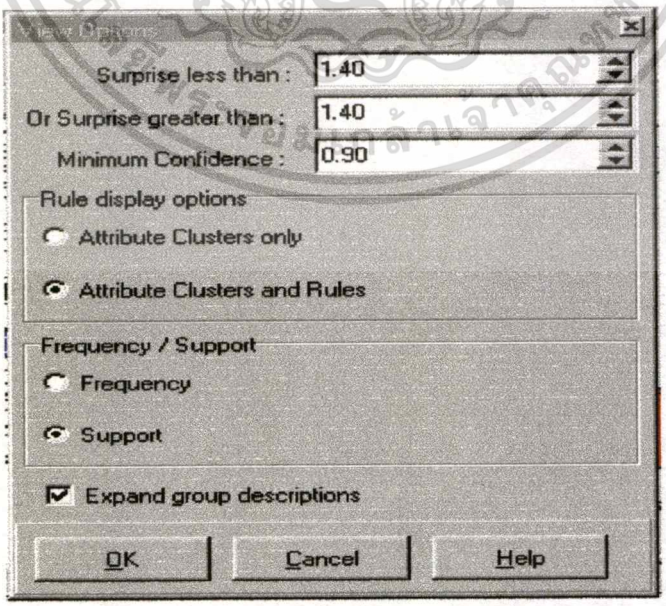
ภาพที่ 46 แสดงโครงสร้างต้นไม้พิจารณาความสัมพันธ์ของเงื่อนไขของแพทย์ที่ทำการตรวจรักษาเป็นแพทย์ประจำ (DR_ID) การวินิจฉัยการติดเชื้อในโรงพยาบาล (TY_NI_DX) ผู้ป่วยแผนกอายุรกรรม (PT_DEP)

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากภาพที่ 46 เป็นภาพแสดงการวิเคราะห์พบว่าถ้าผู้ป่วยแผนกอายุรกรรมและแพทย์ที่ทำการตรวจรักษาเป็นแพทย์ประจำแล้วผู้ป่วยได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาลจะมีค่าเท่ากับ 4.9%จากการวิเคราะห์พบว่ามีค่าความเชื่อมั่นของข้อมูลดังกล่าว = 0.951 ค่าความถี่ของชุดข้อมูล = 0.5195 และการแจกแจงความถี่ของค่าที่คาดหวัง = 1.37

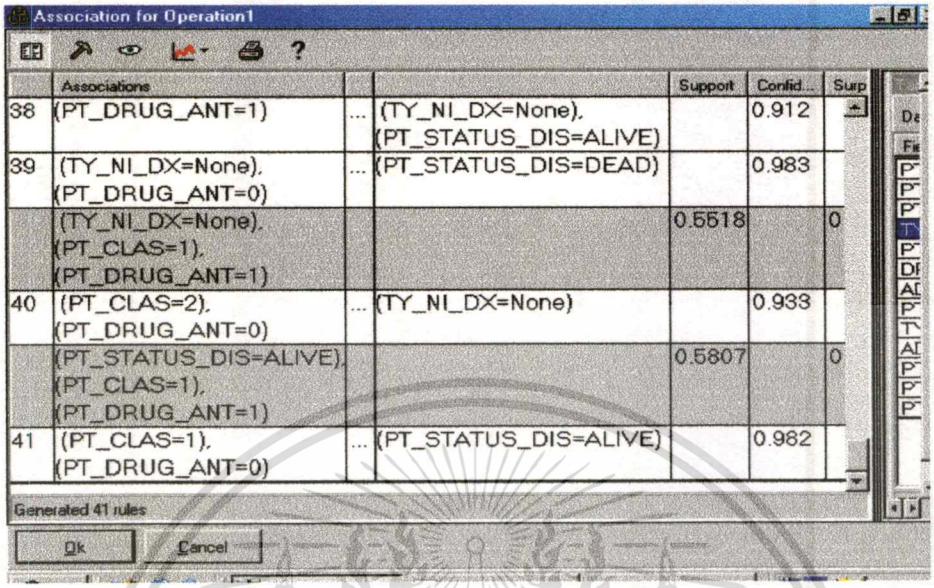


ภาพที่ 47 แสดงการเข้าระบบเพื่อสร้างโครงสร้างต้นไม้โดยใช้Tree Miner



ภาพที่ 48 แสดง การกำหนดขอบเขตค่าในการวิเคราะห์ความสัมพันธ์แบบจัดแบ่งกลุ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



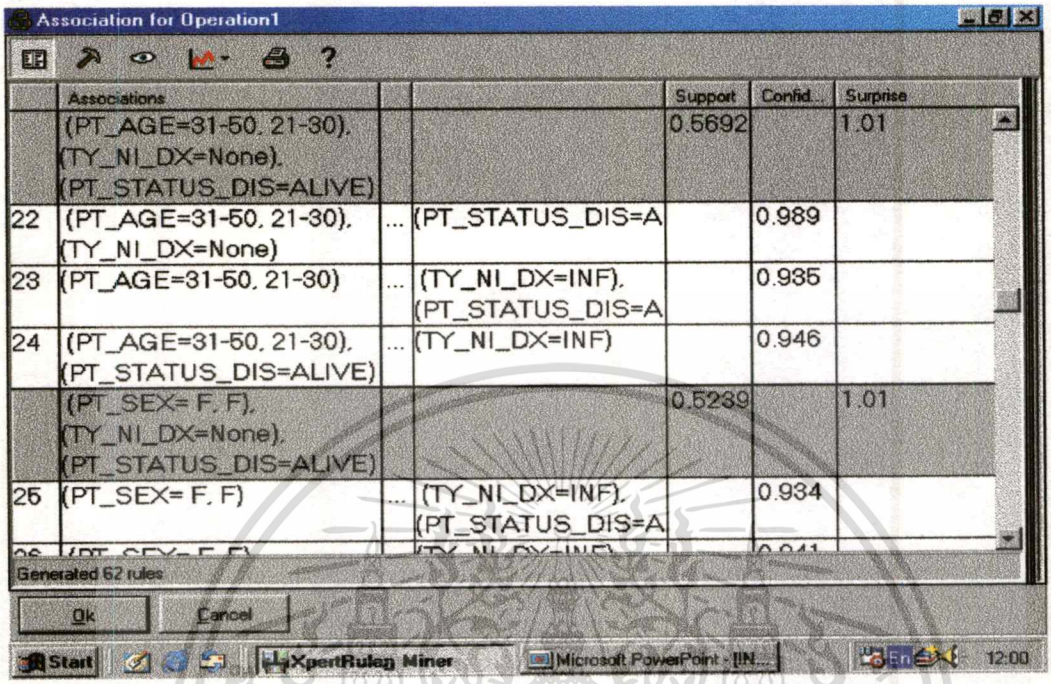
ภาพที่ 49 แสดงความสัมพันธ์ระหว่างเงื่อนไขที่ทำการศึกษาการติดเชื้อในโรงพยาบาลของผู้ป่วย โดยวิเคราะห์ผู้ป่วยประกันสังคม (PT_CLASS=2) ที่ไม่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT=0) และได้รับการวินิจฉัยว่าไม่ติดเชื้อในโรงพยาบาล (TY_NI_DX =None)

จากภาพที่ 49 เป็นภาพแสดงการวิเคราะห์ข้อมูลเชิงเงื่อนไขหาความสัมพันธ์ของปัจจัยเสี่ยงต่อการติดเชื้อในโรงพยาบาลพบว่า

(1) เมื่อผู้ป่วยประกันสังคม (PT_CLASS =2) และผู้ป่วยที่ไม่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT =0) แล้วได้รับการวินิจฉัยว่าไม่ติดเชื้อในโรงพยาบาล (TY_NI_DX =None) ได้ค่าความเชื่อมั่น = 0.933 ค่าความถี่ของชุดข้อมูล = 0.5518

(2) เมื่อผู้ป่วยแผนกอายุรกรรม (PT_DEP=MED) และได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาล (TY-NI_DX=INF) แล้วผู้ป่วยเสียชีวิตภายหลังได้รับการรักษา (PT_STATUS_DIS=DEAD) ได้ค่าความเชื่อมั่น= 0.937 ค่าความถี่ของชุดข้อมูล = 0.5157 ค่าแจกแจงความถี่ของค่าคาดหวัง = 1.01

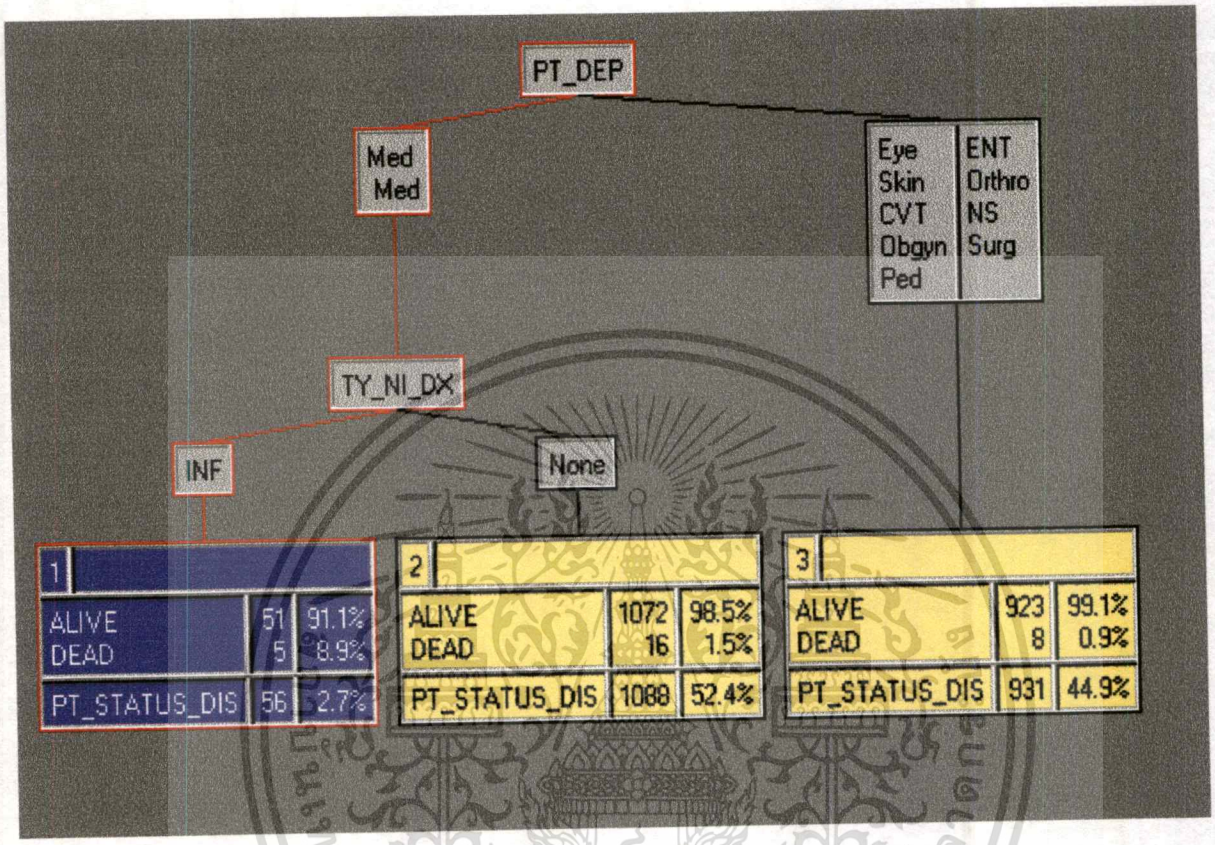
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 50แสดงความสัมพันธ์ระหว่างเงื่อนไขที่ทำการศึกษาการติดเชื้อในโรงพยาบาลของผู้ป่วย (PT_AGE=31-50,21-30 แล้ว TY_NI_DX =INFแล้วPT_STATUS_DIS=ALIVE)

จากภาพที่ 50 เป็นภาพแสดงความสัมพันธ์ระหว่างเงื่อนไขที่ทำการศึกษาการติดเชื้อในโรงพยาบาลของผู้ป่วยโดยพิจารณาผู้ป่วยที่มีอายุในช่วง 31-50ปี และ 21-30 ปีเนื่องจากเป็นวัยทำงานพบว่ามีการติดเชื้อในโรงพยาบาลแล้วภายหลังได้รับการรักษามีการรอดชีวิตเท่ากับ 98.6% ได้ค่าความเชื่อมั่น =0.989 ค่าแจกแจงความถี่ที่คาดหวัง =1.01 ค่าความถี่ของชุดข้อมูล = 0.5692

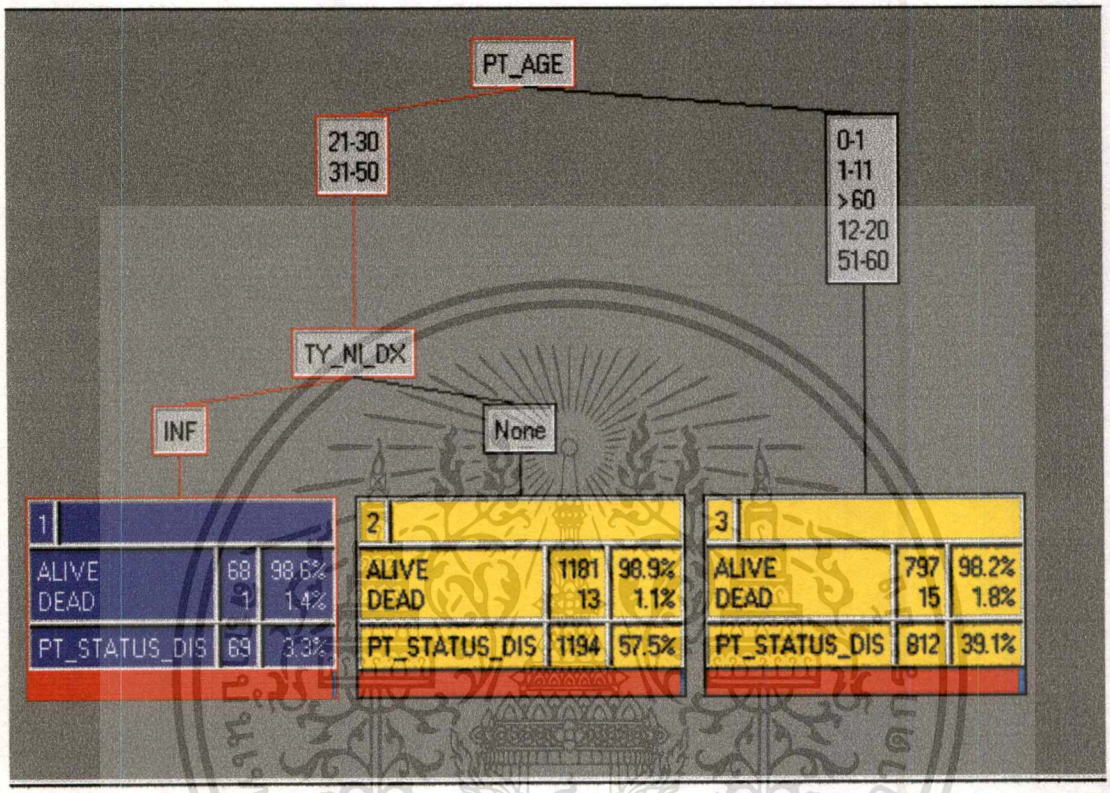
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 51 ภาพแสดงโครงสร้างต้นไม้โดยพิจารณาสร้างความสัมพันธ์ของเงื่อนไข (PT_DEP=MED, TY_NI_DX=INF, PT_STATUS_DIS=DEAD)

จากภาพที่ 51 เป็นภาพแสดงความสัมพันธ์ของเงื่อนไข เมื่อผู้ป่วยอายุรกรรม (PT_DEP= MED) ที่ได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) แล้วผู้ป่วยเสียชีวิตภายหลังการรักษา (PT_STATUS_DIS=DEAD) เท่ากับ 8.9% โดยมีค่าความเชื่อมั่น=0.955 ค่าความถี่ของชุดข้อมูล = 0.5157 ค่าการแจกแจงความถี่ของค่าคาดหวัง =1.01%

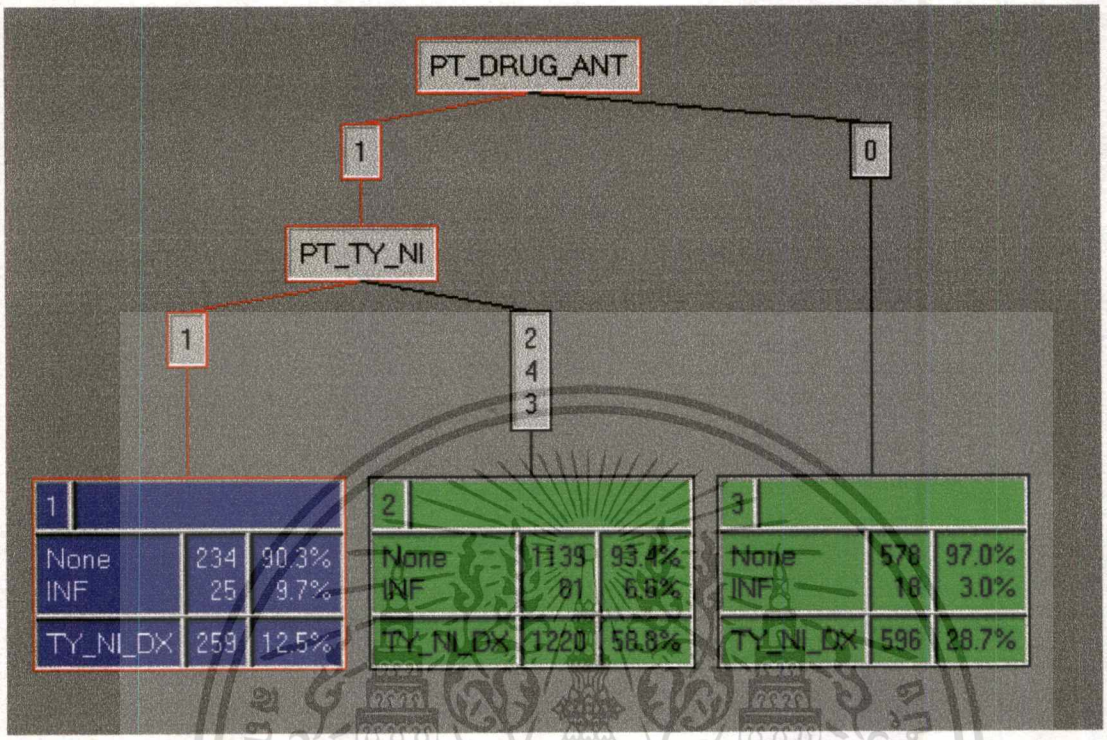
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 52 แสดงโครงสร้างต้นไม้โดยพิจารณาสร้างความสัมพันธ์ของเงื่อนไข (PT_AGE),(TY_NI_DX),(PT_STATUS_DIS)

จากภาพที่ 52 เป็นภาพแสดงความสัมพันธ์ของเงื่อนไขของปัจจัยที่ทำการศึกษาพบว่า เมื่อผู้ป่วยที่มีอายุระหว่าง 21-30 และ 31-50 (PT_AGE = 21-30, 31-50) ที่ได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาล(TY_NI_DX=INF)แล้วผู้ป่วยเสียชีวิตภายหลังการรักษา (PT_STATUS_DIS=DEAD) เท่ากับ1.4% โดยมีค่าความเชื่อมั่น=0.955

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 53 แสดงโครงสร้างต้นไม้โดยพิจารณาสร้างความสัมพันธ์ของเงื่อนไข (PT_DRUG_ANT), (PT_TY_NI), (TY_NI_DX)

จากภาพที่ 53 เป็นภาพแสดงความสัมพันธ์ของเงื่อนไขของปัจจัยที่ทำการศึกษาพบว่า

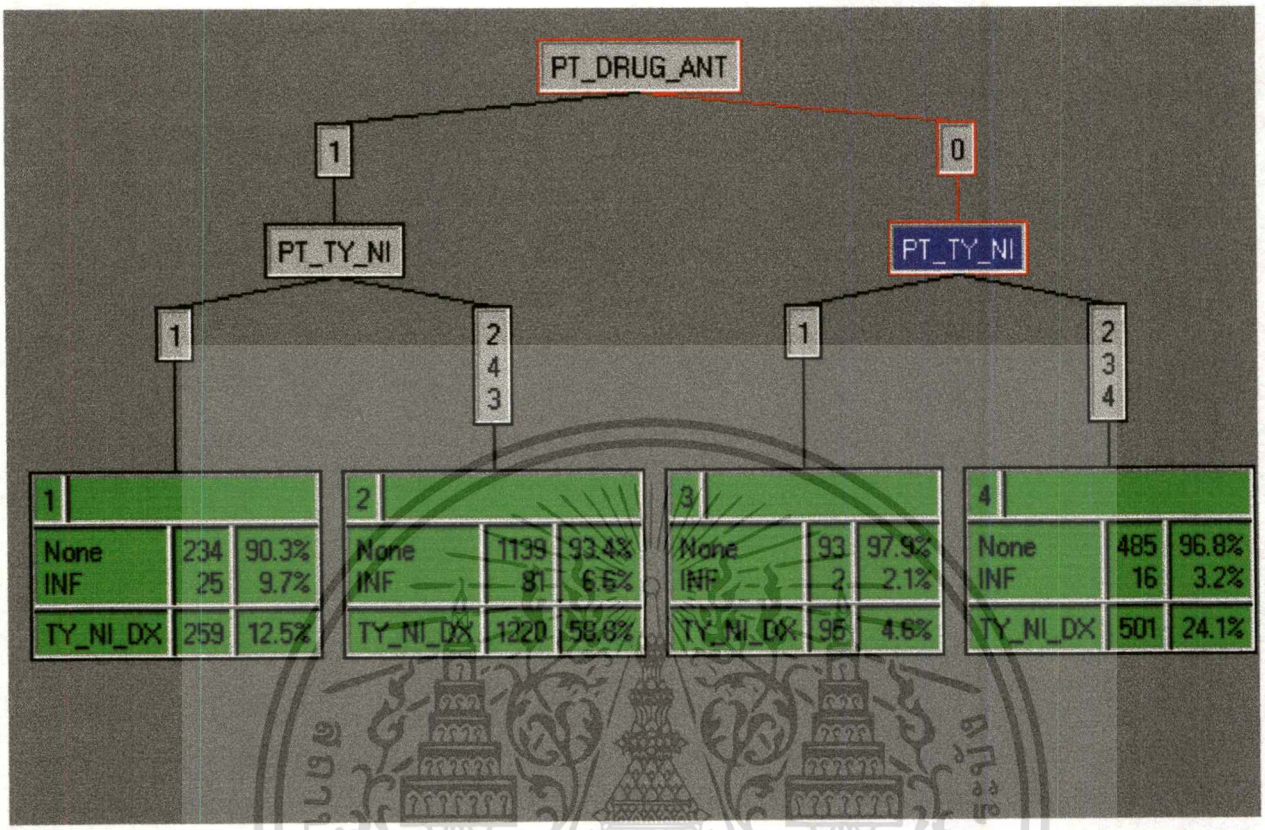
(1) กลุ่มผู้ป่วยหนัก (PT_TY_NI=1) ที่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT=1) และได้รับการวินิจฉัยว่าไม่ติดเชื้อในโรงพยาบาล (TY_NI_DX=None) พบว่ามีค่าเท่ากับ 90.3%

(2) กลุ่มผู้ป่วยหนัก (PT_TY_NI=1) ที่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT=1)และได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) พบว่ามีค่าเท่ากับ 90.7%

(3) กลุ่มผู้ป่วยที่มีไม่ได้รับยาปฏิชีวนะ(PT_DRUG_ANT=0)และได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) พบว่ามีค่าเท่ากับ 3.0%

(4) กลุ่มผู้ป่วยที่มีไม่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT=0) และได้รับการวินิจฉัยว่าไม่ได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX= None) มีค่าเท่ากับ97.0%

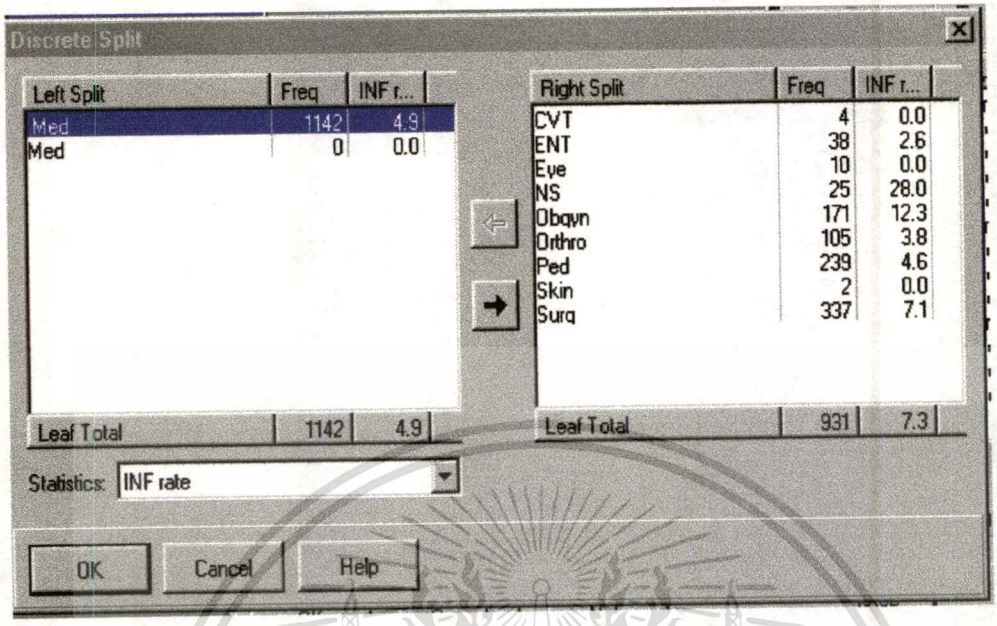
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 54 แสดงโครงสร้างต้นไม้โดยพิจารณาความสัมพันธ์เชิงเงื่อนไข (PT_DRUG_ANT), (PT_TY_NI), (TY_NI_DX)

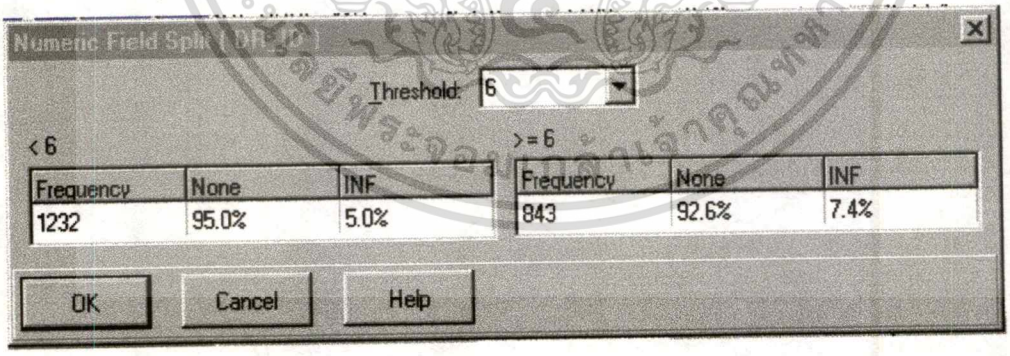
ภาพที่ 54 เป็นภาพแสดงความสัมพันธ์เชิงเงื่อนไขเปรียบเทียบข้อมูลในกลุ่มผู้ป่วยหนักที่ได้รับยาปฏิชีวนะกับกลุ่มที่ไม่ได้รับยาปฏิชีวนะพบว่า ในผู้ป่วยกลุ่มผู้ป่วยหนักที่ทำการเฝ้าระวังโรคที่ได้รับยาปฏิชีวนะพบว่ามีการติดเชื้อในโรงพยาบาลสูงกว่ากลุ่มผู้ป่วยหนักที่ไม่ได้รับปฏิชีวนะและพบว่าได้รับการวินิจฉัยติดเชื้อในโรงพยาบาล =7.6%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 55 แสดงการจัดแยกกลุ่มผู้ป่วยที่ทำการศึกษาที่มีค่าแบบ Discreteออกเป็น 2 กิ่งสาขา

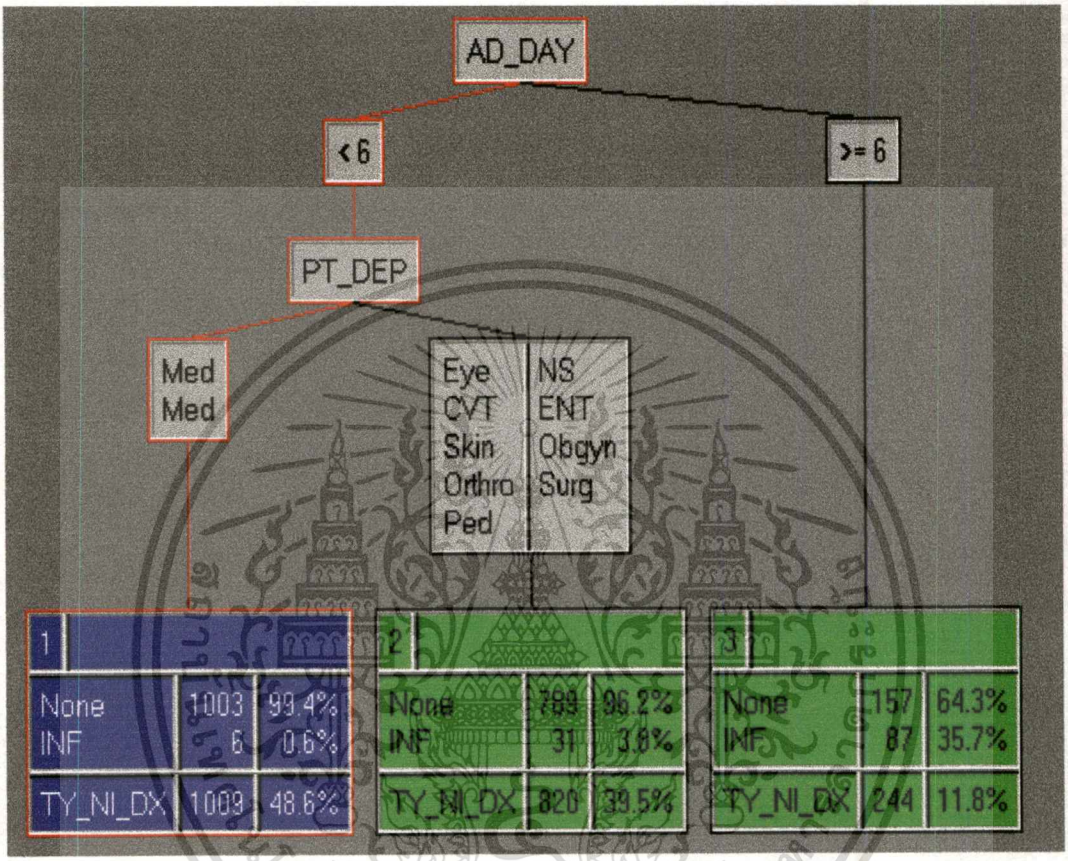
จากภาพที่ 55 เป็นภาพแสดงการแยกกลุ่มที่มีการแจกแจงค่าแบบ Discrete โดยกำหนดการแยกกลุ่มผู้ป่วยอายุรกรรมที่สนใจออกจากกลุ่มผู้ป่วยแผนกอื่นโดยใช้วิธีการของ Discrete Split



ภาพที่ 56 แสดงการแยกกลุ่มผู้ป่วยที่ทำการศึกษาเป็น 2 กิ่งสาขาจากภาพเป็นการแยกแบบจัดกลุ่ม

จากภาพที่ 56 แสดงการแยกกลุ่มแพทย์ที่ทำการศึกษาแบบจัดกลุ่มโดยการกำหนดรหัส มีการแจกแจงค่าให้อยู่ในช่วงที่กำหนดแบบ Numeric Field Split ในรูปค่าที่ระบุคือ การกำหนดรหัสแพทย์ที่ทำการศึกษาให้แพทย์ที่มีรหัส <6 เป็นแพทย์ประจำโรงพยาบาลและแพทย์ที่มีรหัสมากกว่า 6เป็นแพทย์ part time

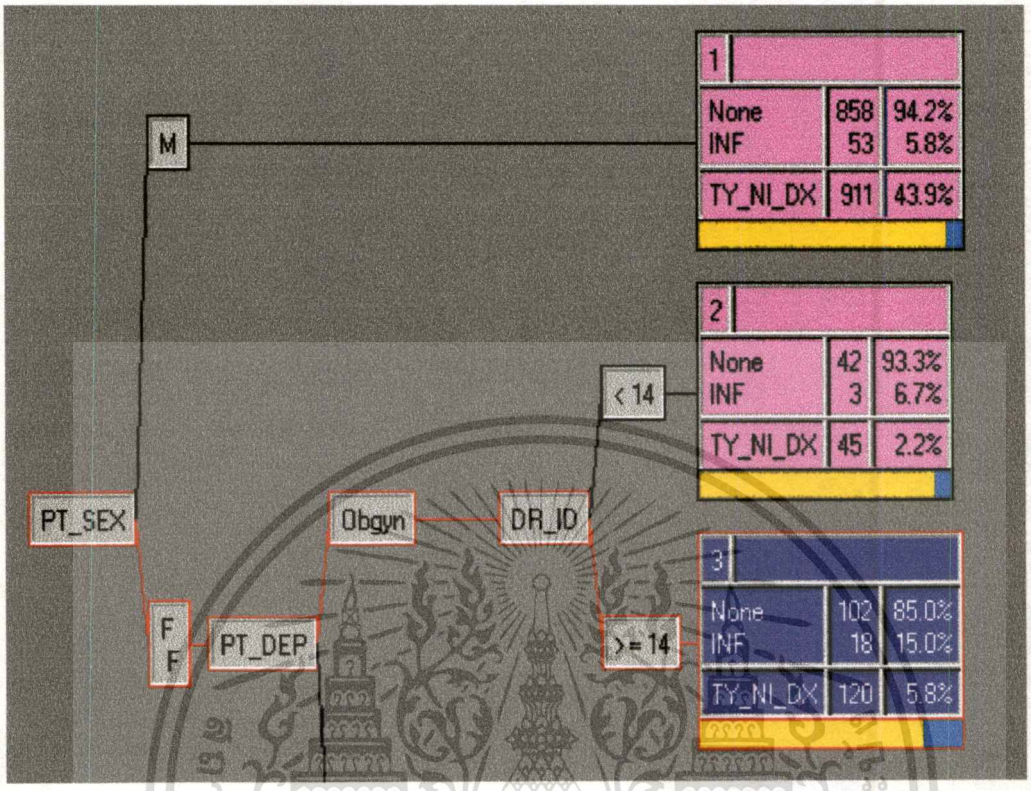
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 57 แสดงโครงสร้างต้นไม้โดยพิจารณาความสัมพันธ์เชิงเงื่อนไข (AD_DAY), (PT_DEP), (TY_NI_DX)

ภาพที่ 57 แสดงความสัมพันธ์เชิงเงื่อนไขเปรียบเทียบข้อมูลในกลุ่มผู้ป่วยที่นอนพักรักษาตัวในโรงพยาบาลโดยกำหนดช่วงระหว่าง < 6 วัน และ>6วันพบว่ากลุ่มผู้ป่วยที่นอนพักรักษาตัวในโรงพยาบาล > 6วันได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาลสูงกว่ากลุ่มผู้ป่วยที่นอนพักรักษาตัวใน< 6วัน ถึง 31.3%

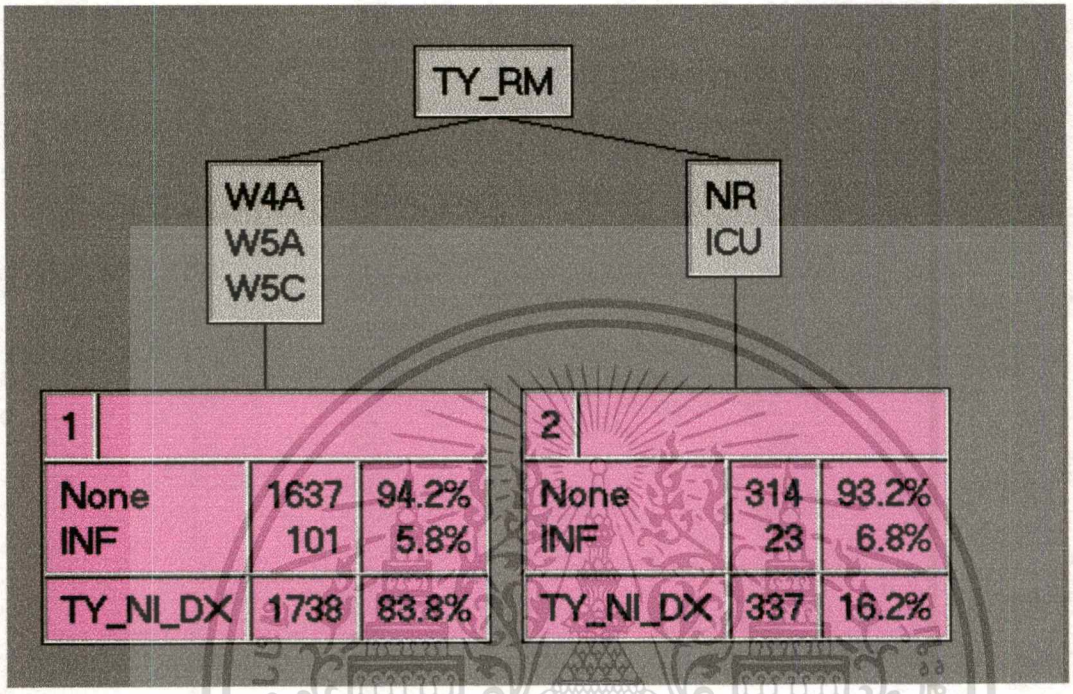
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 58 แสดงโครงสร้างต้นไม้โดยพิจารณาความสัมพันธ์เชิงเงื่อนไข (PT_SEX) (DR_ID), (PT_DEP), (TY_NI_DX)

จากภาพที่ 58 แสดงการหาความสัมพันธ์เชิงเงื่อนไขเปรียบเทียบข้อมูลในกลุ่มผู้ป่วยเพศหญิงที่เข้ารับการรักษาในโรงพยาบาลเนื่องจากพบความถี่ของการติดเชื้อในเพศหญิงสูงกว่าเพศชายจึงพิจารณาที่แผนกสูติ-นรีเวชเป็นประเด็นแรกพบว่าผู้ป่วยหญิงแผนกสูติ-นรีเวชที่ได้รับการดูแลจากแพทย์ part time (รหัส >6) ได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาลโดยมีอัตราการติดเชื้อสูงถึง 15 %

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



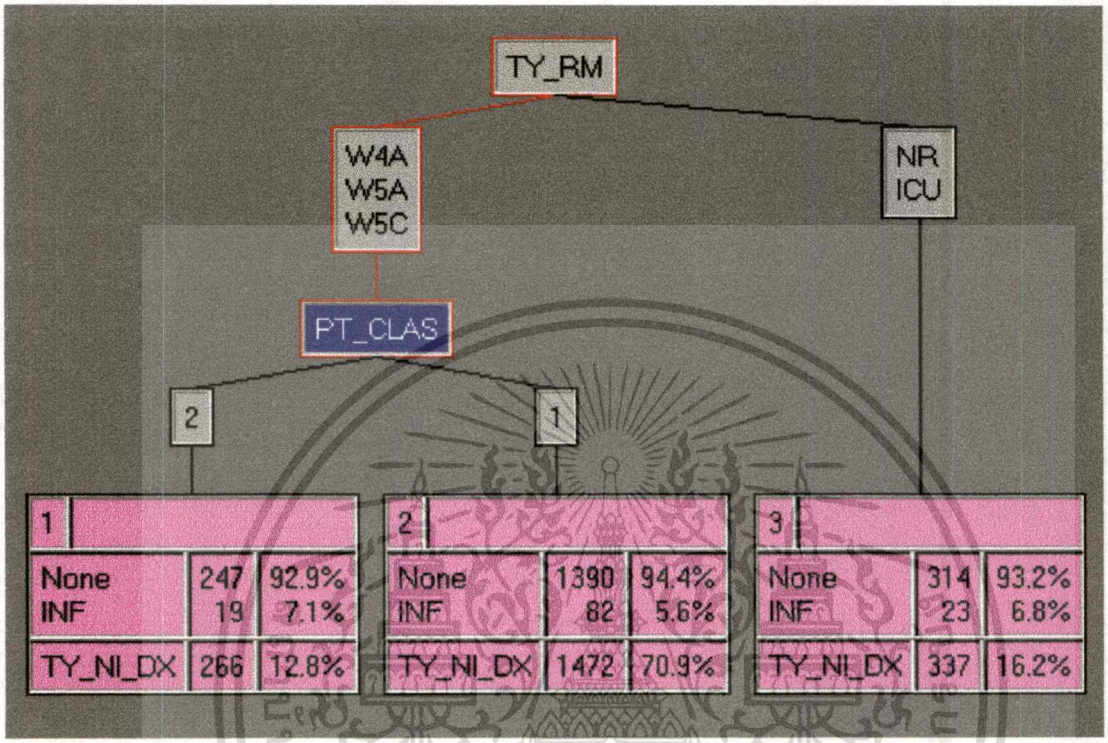
ภาพที่ 59 แสดงโครงสร้างต้นไม้พิจารณาผู้ป่วยที่ได้รับการติดเชื้อในโรงพยาบาลแยกตามประเภทหอผู้ป่วย (TY_NI_DX) และ (TY_RM)

จากภาพที่ 59 เป็นภาพแสดงการพิจารณากลุ่มผู้ป่วยที่รับบริการในหอผู้ป่วยทั่วไปที่ไม่ใช่หอผู้ป่วยเฉพาะโรค (TY_RM =W5C ,W5A, W4A) แล้วพบว่าได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) มีอัตราการติดเชื้อใกล้เคียงกันดังนี้

(1) เมื่อ (TY_RM = W5C, W5A, W4A) และ (TY_NI_DX=INF) เท่ากับ 5.8%

(2) เมื่อ (TY_RM = ICU,NR) และ(TY_NI_DX=INF) เท่ากับ 6.8%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 60 แสดงโครงสร้างต้นไม้พิจารณาความสัมพันธ์เชิงกรณี (TY_RM) (PT_CLASS) (TY_NI_DX)

จากภาพที่ 60 เป็นภาพแสดงการพิจารณากลุ่มผู้ป่วยที่รับบริการในหอผู้ป่วยทั่วไปที่ไม่ใช่หอผู้ป่วยเฉพาะโรค (TY_RM=W5C , W5A, W4A) และเป็นผู้ป่วยประเภททั่วไป (PT_CLASS=1) แล้วได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาล (TY-NI_DX=INF) มีค่าเท่ากับ 5.6%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
บทสรุป

จากการศึกษาพบว่าองค์กรที่ต้องการนำเทคนิคของ Data Mining เข้ามาช่วยในการทำงานด้านการบริการสุขภาพนั้น มีความจำเป็นและเป็นแนวทางที่ทำให้องค์กรเกิดความรู้ที่จะนำมาใช้ประโยชน์โดยเฉพาะองค์กรด้านสุขภาพที่มีปัจจัย ที่ทำให้มีผลต่อการบริการรักษาให้ด้านคุณภาพเพื่อให้ตรงตามเป้าหมายขององค์กร ในการพัฒนาคุณภาพการให้บริการ ในขั้นตอนของการปฏิบัติงานจริงองค์กรใช้อัลกอลิทึม (Algorithm) ของโครงสร้างต้นไม้ (Decision Trees) เป็นเครื่องมือในการแยกแยะกลุ่มและหาความสัมพันธ์ของปัจจัยผู้รับบริการที่ได้รับการติดเชื้อในโรงพยาบาลมีวิธีการศึกษาการทำงานของโปรแกรมระบบซึ่งดึงเอาตัวแปรที่เกี่ยวข้องและตัวแปรอื่นๆเข้ามาศึกษาสรุปและนำผลดังกล่าวไปใช้ในประโยชน์การพัฒนาทางธุรกิจการรักษาพยาบาลและจากการศึกษานั้นจะเห็นได้ว่ามีตัวแปรหลายตัวที่เลือกมาวิเคราะห์ซึ่งผู้ใช้งานจะพบว่าตัวแปรที่ถูกดึงมานั้นอาจจะไม่ใช่ตัวแปรต้นที่ผู้ศึกษาคาดไว้ในการศึกษางานในลักษณะอื่นๆ จะพบว่าอัลกอลิทึม (Algorithm) อาจดึงตัวแปรจำนวนมากออกมาทำให้ผู้วิเคราะห์ที่มีวิจารณญาณและประสพการณ์ในงานนั้นๆในการวิเคราะห์ว่าตัวแปรที่เลือกมานั้นมีความเหมาะสมและความเกี่ยวข้องกับตัวแปรต่างๆมากเพียงใดมีความเกี่ยวข้องกับตัวแปรที่ต้องการศึกษามากเพียงใดถ้าหากพบว่าตัวแปรที่เลือกมาไม่เหมาะสมก็จะต้องใช้ประสพการณ์ของผู้ใช้งานในการตัดตัวแปรที่ทำการศึกษาบางตัวออกหรือเรียกว่าการกำจัด (Pruning) เพื่อหาผลสรุปจากกรรมวิธีการศึกษา และเพิ่มประสิทธิภาพให้มากยิ่งขึ้น

5.1 สรุปผลการศึกษา

การศึกษาครั้งนี้เป็นการศึกษาระบบสารสนเทศเพื่อนำมาใช้ในงานทางด้านการแพทย์เพื่อศึกษาปัจจัยที่มีผลต่อการติดเชื้อในโรงพยาบาล ซึ่งแบ่งออกเป็น 4 กลุ่ม ผู้ป่วยคือ

1. ผู้ป่วยหนักที่พักรักษาตัวในห้องผู้ป่วยหนัก (ICU)
2. ผู้ป่วยที่นอนในโรงพยาบาล เกิน 3 วัน
3. ผู้ป่วยที่ได้รับการสอดใส่อุปกรณ์เครื่องมือแพทย์ (Instrument)
4. ผู้ป่วยที่ได้รับการผ่าตัด หรือคลอดบุตร

โดยทำการศึกษา ผู้ป่วยในลักษณะของการย้อนดูจากประวัติในอดีต (Retrospective) โดยดูจากแบบฟอร์มรายงานการเฝ้าระวังโรคแบบต่อเนื่องและจากประวัติเวชระเบียนของผู้ป่วย

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตั้งแต่ 1 มกราคม 42 – 31 สิงหาคม 2542 และทำการศึกษาประวัติในอนาคต (Prospective) ตั้งแต่ 1 กันยายน42–31ธันวาคม42 ซึ่งจำนวนผู้ป่วยที่ทำการศึกษาและนำมาวิเคราะห์มีจำนวนข้อมูลที่ทำการวิเคราะห์ 2,075 record เครื่องมือที่นำมาใช้ในการวิเคราะห์คือโปรแกรม XpertRule เป็นโปรแกรมระบบที่ใช้ในการวิเคราะห์ให้เกิดความรู้และสามารถนำมาทำนายผลในอนาคตได้เรียกว่า Data Mining จากข้อมูลที่ ทำการศึกษาได้ทำการวิเคราะห์รวมทั้งพิจารณาค่าความถี่ของชุดข้อมูล (Support) การแจกแจงความถี่ของค่าคาดหวัง (Surprise) ค่าความเชื่อมั่นของข้อมูลเปรียบเทียบกัน จากข้อมูลที่ทำการวิเคราะห์ ผลการศึกษาพบว่า

1. ผู้ป่วยพบว่าปัจจัยที่สำคัญของโรงพยาบาลที่ทำให้เกิดการติดเชื้อในโรงพยาบาล ผู้ป่วยแผนกอายุรกรรมที่มีแพทย์ประจำเป็นผู้ทำการรักษา (รหัสที่น้อยกว่า6) แล้วผู้ป่วยได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาลพบว่ามีค่าความเชื่อมั่น=0.951และมีค่าความถี่ของชุดข้อมูลที่ทำการศึกษา (Support) =0.5195(หมายความว่า ชุดความถี่ของรายการที่นำมาวิเคราะห์มีค่า $\geq$ 51% รวมทั้งค่าแจกแจงค่าความถี่ที่คาดหวังไว้ (Surprise) = 1.37 หมายความว่าจากจำนวนผู้ป่วยที่เป็นผู้ป่วยอายุรกรรมได้รับการรักษาจากแพทย์ประจำที่มีรหัส <6 แล้วผู้ป่วยได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาลมีจำนวน 1.37% ของจำนวนผู้ป่วยดังกล่าว

2. จากการวิเคราะห์พบความสัมพันธ์ของเงื่อนไข เมื่อผู้ป่วยอายุรกรรม (PT_DEP=MED)ที่ได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) แล้วผู้ป่วยเสียชีวิตภายหลังการรักษา (PT_STATUS_DIS=DEAD) เท่ากับ 8.9%โดยมีค่าความเชื่อมั่น=0.955 ค่าความถี่ของชุดข้อมูล =0.5157 ค่าการแจกแจงค่าความถี่ของการคาดหวัง =1.01

3. จากการวิเคราะห์พบว่าถ้าผู้ป่วยแผนกอายุรกรรมและแพทย์ที่ทำการตรวจรักษาเป็นแพทย์ประจำแล้วผู้ป่วยได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาลจะมีค่าเท่ากับ 4.9% จากการวิเคราะห์พบว่ามีค่าความเชื่อมั่นของข้อมูลดังกล่าว = 0.951 ค่าความถี่ของชุดข้อมูล = 0.5195 และการแจกแจงค่าความถี่ของการคาดหวัง = 1.37

4. จากการวิเคราะห์ข้อมูลเชิงเงื่อนไขความสัมพันธ์ของปัจจัยเสี่ยงต่อการติดเชื้อในโรงพยาบาลพบว่าเมื่อผู้ป่วยประกันสังคม (PT_CLASS =2) และผู้ป่วยที่ไม่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT =0) แล้วได้รับการวินิจฉัยว่าไม่ติดเชื้อในโรงพยาบาล (TY_NI_DX =None) ได้ค่าความเชื่อมั่น = 0.933 ค่าความถี่ของชุดข้อมูล = 0.5518

5. พบว่ากลุ่มผู้ป่วยที่ทำการศึกษาที่มีอายุ (31 –50) (21–30) ปี ได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาลและมีอัตราการรอดชีวิตหลังได้รับการรักษา พบว่าผลการวิเคราะห์ได้ค่าความเชื่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มั่นเท่ากับ 0.946 ค่าการแจกแจงความถี่ของค่าคาดหวัง=1.01ในความถี่ของชุดรายการ=0.5692 นอกจากนั้นทำการศึกษาเงื่อนไขการติดเชื้อในโรงพยาบาลของกลุ่มผู้ป่วยโดยพิจารณาผู้ป่วยที่มีอายุในช่วง 31-50ปี และ 21-30 ปี เนื่องจากเป็นวัยทำงานพบว่ามีการติดเชื้อในโรงพยาบาลแล้วภายหลังได้รับการรักษามีอัตราการรอดชีวิตเท่ากับ 98.6% ได้ค่าความเชื่อมั่น =0.989 ค่าแจกแจงความถี่ที่คาดหวัง =1.01 ค่าความถี่ของชุดข้อมูล = 0.5692เมื่อผู้ป่วยที่มีอายุระหว่าง 21-30และ31-50 (PT_AGE =21-30,31-50ปี) ที่ได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) แล้วผู้ป่วยเสียชีวิตภายหลังการรักษา (PT_STATUS_DIS=DEAD) เท่ากับ 1.4% โดยมีค่าความเชื่อมั่น = 0.955

6. พิจารณากลุ่มผู้ป่วยที่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT=1) และเป็นผู้ป่วยหนัก (PT_TY_NI=1) ได้รับการวินิจฉัยว่าไม่ติดเชื้อในโรงพยาบาล (TY_NI_DX=None) มีค่าเท่ากับ 90.3%และผู้ป่วยที่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT=1) และเป็นผู้ป่วยหนัก (PT_TY_NI=1) ที่ได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) มีค่าเท่ากับ 90.7% แสดงความสัมพันธ์เชิงเงื่อนไขเปรียบเทียบข้อมูลในกลุ่มผู้ป่วยหนักที่ได้รับยาปฏิชีวนะกับกลุ่มที่ไม่ได้รับยาปฏิชีวนะพบว่า ในผู้ป่วยกลุ่มผู้ป่วยหนักที่ทำการเฝ้าระวังโรคที่ได้รับยาปฏิชีวนะพบว่ามีการติดเชื้อในโรงพยาบาลสูงกว่ากลุ่มผู้ป่วยหนักที่ไม่ได้รับปฏิชีวนะและพบว่าได้รับการวินิจฉัยติดเชื้อในโรงพยาบาล = 7.6%

7. พิจารณากลุ่มผู้ป่วยที่ไม่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT=0) และได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) มีค่าเท่ากับ3.0%

8. พิจารณากลุ่มผู้ป่วยที่ไม่ได้รับยาปฏิชีวนะ (PT_DRUG_ANT=0) และได้รับการวินิจฉัยว่าไม่ได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX= None) มีค่าเท่ากับ97.0%

9. ความสัมพันธ์เชิงเงื่อนไขเปรียบเทียบข้อมูลในกลุ่มผู้ป่วยที่นอนพักรักษาตัวในโรงพยาบาลโดยกำหนดช่วงระหว่างมากกว่า 6 วัน และน้อยกว่า 6วัน พบว่ากลุ่มผู้ป่วยที่นอนพักรักษาตัวในโรงพยาบาล น้อยกว่า 6วันได้รับการวินิจฉัยว่าได้รับการติดเชื้อในโรงพยาบาลสูงกว่ากลุ่มผู้ป่วยที่นอนพักรักษาตัวในระยะเวลาน้อยกว่า6วัน ถึง 31.3%

10. พิจารณากลุ่มผู้ป่วยที่รับบริการในหอผู้ป่วยทั่วไปที่ไม่ใช่หอผู้ป่วยเฉพาะทาง (TY_RM=W5C , W5A, W4A) และเป็นผู้ป่วยประเภททั่วไป (PT_CLASS=1) แล้วได้รับการติดเชื้อในโรงพยาบาล (TY_NI_DX=INF) ได้ผล=5.6%

11.แสดงการหาความสัมพันธ์เชิงเงื่อนไขเปรียบเทียบข้อมูลในกลุ่มผู้ป่วยเพศหญิงที่เข้ารับการรักษาในโรงพยาบาลโดยได้รับการดูแลจากแพทย์ part time แผนกสูตินรีเวชพบว่าได้รับการวินิจฉัยว่าติดเชื้อในโรงพยาบาลเป็นอัตราการติดเชื้อที่สูงถึง 15 %

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ประโยชน์ที่ได้รับจากการศึกษา

1. ได้ข้อมูลพื้นฐานของผู้ป่วยที่รับบริการในโรงพยาบาลสามารถนำข้อมูลที่ได้เป็นแนวทางในการเฝ้าระวัง ปัจจัย ตัวแปร ที่มีความสัมพันธ์กับการติดเชื้อในโรงพยาบาล
2. ผู้บริหารสามารถนำผลจากการศึกษาเป็นแนวทางในการวางแผนเพื่อการปรับปรุงการให้บริการพยาบาลโดยสามารถกำหนดค่าชี้วัดในการควบคุมและป้องกันการติดเชื้อในโรงพยาบาล
3. สามารถนำผลการศึกษานำเสนอผู้บริหารเพื่อพิจารณาร่วมกับการกำหนดตัวชี้วัดในการควบคุมคุณภาพการให้ บริการทางการแพทย์
4. สามารถคาดเดาการติดเชื้อ หรือแนวโน้มของการติดเชื้อในโรงพยาบาลได้

5.3 ข้อเสนอแนะในการนำผลการศึกษาไปใช้

จากการศึกษาครั้งนี้แสดงให้เห็นว่า การติดเชื้อในโรงพยาบาลมีความสัมพันธ์กับแพทย์ที่ทำการรักษา การแบ่งแยกแผนกในการทำการรักษาผู้ป่วย รวมทั้งอัตราการติดเชื้อในโรงพยาบาลแล้วผู้ป่วยเสียชีวิตดังนั้นผู้ศึกษามีข้อเสนอแนะในการนำผลการศึกษาไปใช้ดังต่อไปนี้

1. นำผลการศึกษาไปใช้ในการกำหนดค่าดัชนีชี้วัดคุณภาพการให้การบริการการรักษาพยาบาล เช่น ค่าดัชนีชี้วัดการติดเชื้อในโรงพยาบาล อัตราการเสียชีวิตของผู้ป่วย
2. นำผลการศึกษาเป็นแนวทางในการจัดตั้งเกณฑ์มาตราฐานเพื่อตรวจสอบมาตราฐานคุณภาพบริการพยาบาล
3. เพื่อประเมินและติดตามการให้บริการพยาบาล
4. นำผลการศึกษาเป็นพื้นฐานในการจัดหลักสูตรอบรมเจ้าหน้าที่ทางการพยาบาลเพื่อรับทราบปัญหาและวิธีปฏิบัติในการป้องกันการติดเชื้อในโรงพยาบาล

5.4 ข้อเสนอแนะในการศึกษาครั้งต่อไป

1. ควรมีการศึกษารายละเอียดเฉพาะรายผู้ป่วยที่ได้รับการติดเชื้อในโรงพยาบาลเพิ่มเพื่อวิเคราะห์ปัจจัยส่วนบุคคลที่มีความสัมพันธ์กับการติดเชื้อในโรงพยาบาล เช่นโรคที่ผู้ป่วยได้รับการวินิจฉัยก่อนเกิดการติดเชื้อในโรงพยาบาล โรคประจำตัวของผู้ป่วย ระยะเวลาในการรักษาโรคติดเชื้อในโรงพยาบาล จำนวนวันที่ค้นพบการติดเชื้อในโรงพยาบาล ชนิดของการติดเชื้อในโรงพยาบาลเป็นต้น
2. ควรศึกษาผลการรักษาโดยชนิดของยาปฏิชีวนะกับแพทย์ที่ทำการรักษาผู้ป่วยเพื่อทราบลักษณะนิสัยของแพทย์ในการใช้ยาปฏิชีวนะและเพื่อพยากรณ์แนวโน้มของการใช้ยาปฏิชีวนะใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อนาคตรวมทั้งควรวิเคราะห์ระยะเวลาที่เหมาะสมในการใช้ยาปฏิชีวนะสัมพันธ์กับชนิดของโรคติดเชื้อในโรงพยาบาล

3. ควรศึกษารายละเอียดเฉพาะในปัจจัยด้านบุคลากรทางการแพทย์ที่ให้การรักษาเพื่อวิเคราะห์เช่น ชนิดของอุปกรณ์ที่แพทย์ใส่ให้ผู้ป่วย ชนิดการผ่าตัด จำนวนชั่วโมงที่ทำการผ่าตัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

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

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

"Knowledge Seeker." available: http:/www.angoss.com/

Meers, P.,Jacobsen, W and Mcphersan, M.(1992). Hospital Infection Control for Nurse. California: Chapman & Hall.

Rob, Peter. Carlos Coronel.(1997). Database systems, A Division of International Thomson Publishing

Wenzel, R. P. (1993). Prevention and Control of Nosocomial Infection (2nd ed.) Baltimore: Williams & Wilkins.

" XpertRule Miner" available: http:/www.attar.com

เติมศรี ชำนิจารกิจ. (2531). สถิติประยุกต์ทางการแพทย์. กรุงเทพฯ จุฬาลงกรณ์มหาวิทยาลัย.

มหิดล,มหาวิทยาลัย.(2535). รายงานการสำรวจความชุกของโรคติดเชื้อในโรงพยาบาล กรุงเทพฯ: ชมรมควบคุมโรคติดเชื้อในโรงพยาบาลแห่งประเทศ.

สมพนธ์ บุณยคุปต์, สมหวัง ด่านชัยวิจิตร, และสมศักดิ์ โลห์เลขา. (2532).การวินิจฉัยและการรักษาโรคติดเชื้อที่พบบ่อย. กรุงเทพฯ :กรุงเทพเวชสาร.

สมหวัง ด่านชัยวิจิตร. (2538) วิธีป้องกันโรคติดเชื้อในโรงพยาบาล กรุงเทพฯ:คณะแพทย์ศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

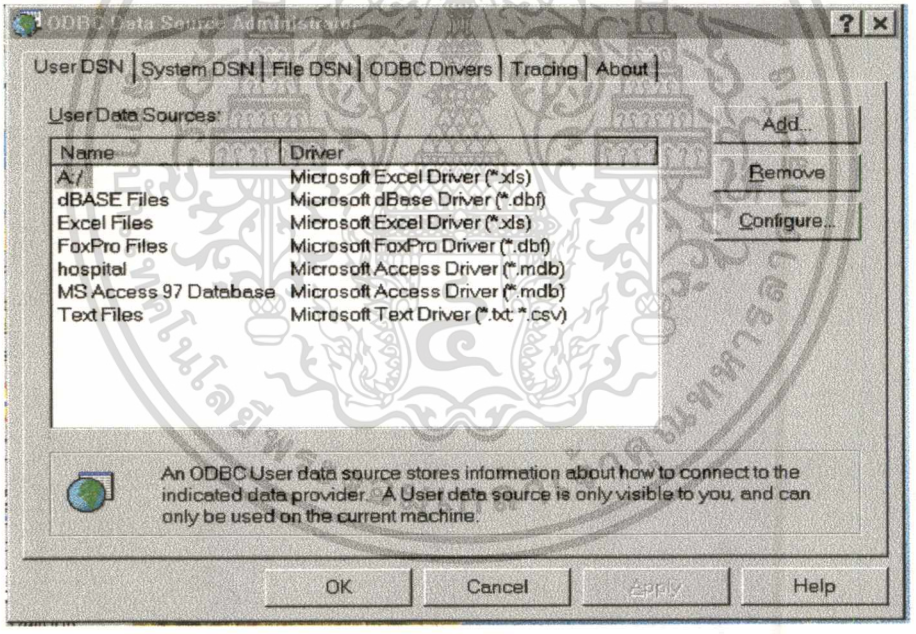
ข้อมูลที่นำมาใช้ในการวิเคราะห์ด้วยขบวนการ Data Mining เพื่อหาปัจจัยที่มีผลต่อการติดเชื้อในโรงพยาบาลจะต้องใช้การเรียกฐานข้อมูลโดยโปรแกรม ODBC ในการเชื่อมโยงข้อมูลเพื่อนำมาวิเคราะห์ดังภาคผนวก ก นอกจากนั้นมีขบวนการทางสถิติที่ช่วยในการคำนวนเพื่อหาความสัมพันธ์ของปัจจัย ในภาคผนวก ข

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ODBC

เพื่อเป็นการเรียกใช้ฐานข้อมูลที่สร้างมาจากโปรแกมแอบพลิเคชั่นต่างโปรแกรมจะอาศัยโปรแกรมที่ใช้ ODBC หรือ Open Database Connectivity ซึ่งโปรแกรมจะช่วยให้โปรแกรมแอบพลิเคชั่นที่แตกต่างกัน สามารถใช้ฐานข้อมูลร่วมกันได้โดยโปรแกรมที่ทำหน้าที่เป็น ODBCจะถูกเรียกว่า Driver ซึ่งหากต้องการติดต่อกับฐานข้อมูลชนิดใด ก็มักจะตั้งชื่อให้ตรงกับชื่อของฐานข้อมูลดังภาพผนวกที่ 1 จะแสดงตัวอย่าง ODBC



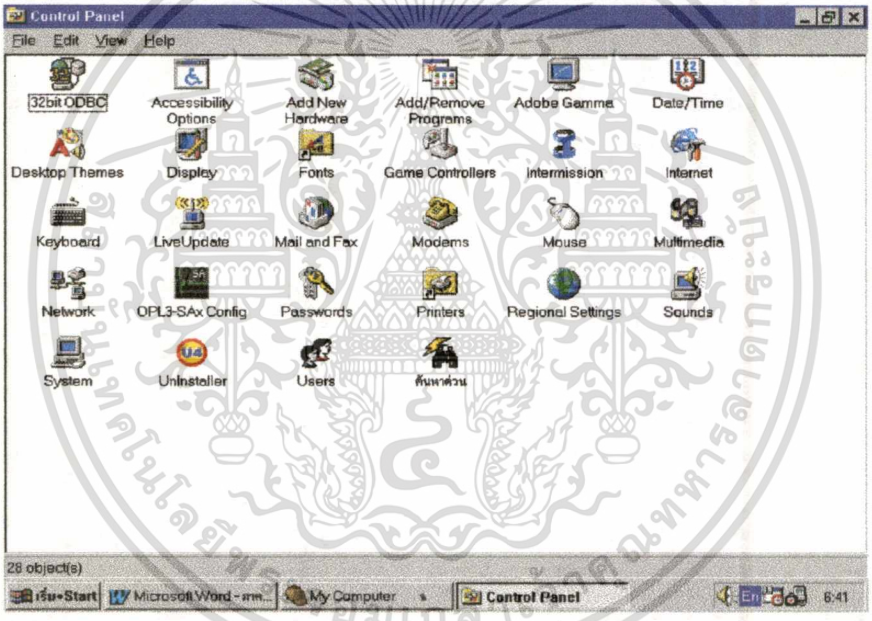
ภาพผนวกที่ 1 ตัวอย่าง ODBC

จากภาพผนวกที่ 1จะมีODBC Driver อาทิเช่น Microsoft Access Driver (*.MDB)ซึ่งจะเป็นตัวโปรแกรมที่ทำหน้าที่ในการเชื่อมโยงโปรแกรมแอบพลิเคชั่นต่างๆให้สามารถเรียกใช้งานฐานข้อมูลที่สร้างจาก Microsoft Access ได้โดยการเรียกผ่านทาง ODBC Driver

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ODBCและการสร้าง DSN

ก่อนที่โปรแกรมแอปพลิเคชั่นต่างๆจะสามารถใช้งานฐานข้อมูลที่สร้างโดย Microsoft Access ได้จะต้องสร้างส่วนที่ใช้ในการติดต่อฐานข้อมุลขึ้นมาก่อนซึ่งเรียกว่าเป็น Data Source Name หรือDSN เมื่อโปรแกรมแอปพลิเคชั่นใดต้องการใช้งานฐานข้อมูลดังกล่าว ก็เพียงแต่เรียกไปที่ DSNนั้นก็จะสามารถติดต่อกับฐานข้อมูลที่DSN นั้นติดต่ออยู่ได้การสร้างDSN สามารถทำได้โดยการเรียกโปรแกรม 32bit ODBC ที่อยู่ในcontrol panelขึ้นมาดังภาพผนวกที่ 2

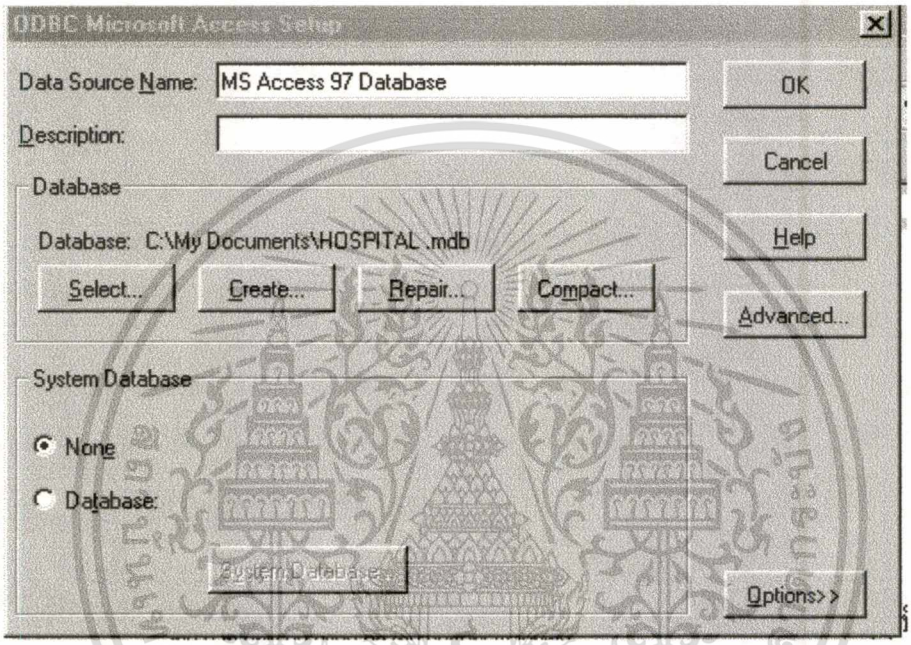


ภาพผนวกที่ 2 Control Panel

เมื่อเรียกโปรแกรม 32 bit ODBCขึ้นมา โปรแกรม 32 bit ODBC ประกอบด้วยส่วนประกอบหลายส่วนแต่ส่วนที่สำคัญซึ่งใช้ในการติดตั้ง DSN สำหรับติดต่อบานข้อมูลได้แก่ส่วนที่เรียกว่า system DSN ให้เรียกเข้าไปในส่วนนี้หากเคยสร้าง DSN ไว้ก่อนหน้านี้แล้วก็จะเห็นรายชื่อของ DSN ที่มีทั้งหมดที่เครื่องคอมพิวเตอร์นั้นมีอยู่ หากต้องการดูว่า DSN มีลักษณะประการใดให้คลิกปุ่ม configure

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DSN จะทำการติดต่อฐานข้อมูลชนิดที่มาจาก Microsoft Access หากโปรแกรมแอปพลิเคชั่นใดต้องการใช้ฐานข้อมูลนี้ก็จะต้องเรียกผ่าน ODBCแล้วโยงไปหา DSNก็จะทำให้สามารถเรียกใช้ฐานข้อมูลได้



ภาพผนวกที่ 3 แสดงคุณสมบัติของ DSN

จากภาพผนวกที่ 3 DSN ที่มีชื่อ MS Access 97และ DSN นี้จะติดต่อฐานข้อมูลชื่อ HOSPITAL.mdb ก็จะเรียกผ่านทาง ODBC แล้วโยงไปหา DSN ก็จะทำให้สามารถเรียกใช้ฐานข้อมูลดังกล่าวได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข
สถิติต่าง ๆ ที่ใช้ในการวิเคราะห์

1.การหาค่าเฉลี่ย (Mean) คำนวณหาค่าเฉลี่ยของอายุ ระยะเวลาที่ผู้ป่วยเข้ารับการรักษาในโรงพยาบาล ระยะเวลาที่นอนในโรงพยาบาลใช้การหาเฉลี่ย (จากโปรแกรม)

$$x = \frac{\sum x}{n}$$

เมื่อ x = ค่าเฉลี่ย

$\sum n$ = ผลรวมของค่าสังเกต

n = คำนวณค่าสังเกตทั้งหมด

2.ค่าเบี่ยงเบนมาตรฐาน (Strandard Deviation) การคำนวณหาส่วนเบี่ยงเบนมาตรฐานของอายุ ระยะเวลาที่ผู้ป่วยเข้ารับการรักษาในโรงพยาบาล

$$SD = \frac{\sqrt{\sum x^2}}{n}$$

เมื่อ SD = ส่วนเบี่ยงเบนมาตรฐาน

$\sum x^2$ = ผลรวมของค่าสังเกตทุกค่า แล้วนำมายกกำลังสอง

n = จำนวนค่าสังเกตทั้งหมด

3. ค่าความเป็นไปได้ (Entropy) ซึ่งถูกคำนวนจากมาจากสูตร

$$\sum_{i=1}^{2} \frac{\sum_{j=1}^{c} f_{ij}}{R} \sum_{j=1}^{c} P(c_{ij}) Ln(P(c_{ij}))$$

เมื่อ C = จำนวนของผลลัพธ์ที่วัดค่า

f = ความถี่ของผลลัพธ์ที่ได้จากแขนงที่แตกย่อย(j)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R = จำนวนทั้งหมดของการบันทึกในแต่ละแขนง

P(cij) = ค่าที่แสดงโดยใช้สูตรคือ

$$\frac{fij}{\sum_{k=1}^{c} fkj}$$

4. ค่าอัลกอลิทึม สำหรับผลลัพธ์ที่เป็น จำนวนเต็ม (Discrete)

ขบวนการสร้าง โครงสร้างต้นไม้ จาก Data set โครงต้นไม้ถูกสร้างโดยการแบ่งหลายครั้งเพื่อให้ได้ข้อมูลที่มีคุณสมบัติที่แตกต่างกันจนถึงจุดสุดท้ายการให้เหตุผลจะถูกใช้กับ โครงสร้างที่ขึ้นกับการวัดของระดับสามารถของคุณสมบัติแต่ละชนิด ข้อบ่อชี้ในการเลิกสาขา คือจะตัดสินใจเพื่อจุดสุดท้ายมาถึงนั้นคือมีการสร้างสาขา 2 ทางในทุกกิ่งที่แยกออกมาของต้นไม้ สำหรับต้นไม้ที่แบ่งเป็นประเภทผลลัพธ์ที่ต้องการการเลือกคุณสมบัติเพื่อแยกทุกๆขั้นตอนในการสร้างต้นไม้จะถูกทำผ่านรายละเอียดของข้อมูลของคุณสมบัติแต่ละประเภทใน Term ของกลุ่มผลลัพธ์แต่ละประเภทส่วนประกอบของข้อมูลนี้ถูกวัดโดยค่าความน่าจะเป็นกับข้อมูลมากสุดไปยังค่าความน่าจะเป็นที่น้อยที่สุดด้วยเหตุผลค่าความน่าจะเป็นที่น้อยที่สุดจะถูกเลือกทุก ๆกิ่ง สำหรับคุณสมบัติที่ต้องการค่าของกลุ่มจะถูกแยกระหว่าง 2 กิ่งที่มีค่าความน่าจะเป็นน้อยสำหรับคุณสมบัติเชิงปริมาณ (Numeric Attribute) จะแยกโดยค่าความน่าจะเป็นต่ำสุด โดยคำนวณจาก

$$\sum_{i=1}^{2} \frac{\sum_{j=1}^{c} fij}{R} \sum_{j=1}^{c} P(cij) Ln(P(cij))$$

C คือ ค่าผลลัพธ์ f คือ ความถี่ของผลลัพธ์ j ในกิ่ง i

R คือ จำนวนทั้งหมดที่ทำการบันทึกใน 2 กิ่งสาขา และ P(Cij) หาได้จากสูตร fij

$$\frac{fij}{\sum_{k=1}^{c} fkj}$$

สำหรับ Regression tree การเลือกข้อมูลจะใช้ค่า NSD (Normalized Standard Deviation) โดยจะใช้ค่าที่ต่ำสุด ของทุก ๆ กิ่ง สำหรับ Discrete Attribute ค่าของกลุ่มถูกแยกระหว่าง 2 กิ่งด้วยค่า NSD

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค

แบบฟอร์มรายงานการเฝ้าระวังโรคแบบต่อเนื่อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบฟอร์มรายงานการเฝ้าระวังโรคแบบต่อเนื่อง

โรงพยาบาลเกษมราษฎร์ ประชาชื่น

Month/year ...../.....

❶ WARD ..... ROOM ..... วันที่ ..... ถึง .....

❷ WARD ..... ROOM ..... วันที่ ..... ถึง .....

NAME ..... HN. ..... AN. .....

AGE ..... SEX ..... ADMITTED DATE ...../..... .....

DISCHARGE DATE ...../...../..... DURATION OF ADMISSION..... DAY.

DIAGNOSIS ON ADMISSION .....

.....

FINAL DIAGNOSIS .....

1. ผู้ป่วยได้รับการเข้ารักษาในโรงพยาบาลครั้งสุดท้ายเมื่อวันที่ ..... โดยมีอาการแสดง .....

2. อาการแสดงของผู้ป่วย

☐ มีไข้แรกรับ .....

☐ มีไข้หลังรับไว้ในโรงพยาบาลฯ เมื่อวันที่ .....

☑ ไม่มีไข้

3. อาการแสดงอื่น ๆ ที่ปรากฏ

☐ ไอ เริ่มเมื่อวันที่ .....

☐ ถ่ายเหลว เมื่อวันที่ .....

☐ ปัสสาวะแสบขัด เริ่มเมื่อวันที่ .....

☐ แผลมีหนอง / บวม / แดง เมื่อวันที่ ..... บริเวณ .....

☐ อื่น ๆ ระบุ ..... วันที่ .....

4. การใช้เครื่องมือหรืออุปกรณ์พิเศษ ใส่ให้ผู้ป่วย

☐ การใส่ท่อช่วยหายใจ การดูดเสมหะ การใช้เครื่องช่วยหายใจ วันที่ ..... ถึง .....

☐ การใส่สายสวนเข้าผู้ป่วย ชนิด ..... วันที่ ..... ถึง .....

☐ อื่น ๆ ระบุ ..... วันที่ ..... ถึง .....

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5. การผ่าตัดเมื่อวันที่ ชนิดการผ่าตัด ชนิดของแผลผ่าตัด ชนิดการดมยา

....

....

6. ANTIBIOTIC ที่ผู้ป่วยได้รับ

1. ..... เริ่มเมื่อวันที่ ..... จำนวน .... วัน
2. ..... เริ่มเมื่อวันที่ ..... จำนวน ..... วัน
3. ..... เริ่มเมื่อวันที่ ..... จำนวน ..... วัน

7. ผลการตรวจอื่น ๆ (เช่น X-ray, Endoscope, Ultrasound)

ชนิดการตรวจ วันที่ตรวจ ผลการตรวจ

....

....

....

8. ผลการตรวจทางห้องปฏิบัติการที่ผิดปกติ

Specimen วันที่ตรวจ ผลการตรวจ

....

....

....

9. ผล ANTIBIOTIC SENSITIVITY TEST

.....

.....

.....

10. ผลสรุป :

☐ ไม่ติดเชื้อในโรงพยาบาล

☐ ติดเชื้อในโรงพยาบาล บริเวณ .....

11. DISCHARGE STATUS

☐ ALIVE

☐ DEAD .....

☐ Directly due to nosocomial infecition .....

☐ Partly contributed by nosocomial infection .....

☐ Not relaled to nosocomial infection .....

ลงชื่อผู้บันทึก .....

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อผู้เขียน	นาง ดวงรัตน์ รักศักดิ์ศรี
วัน เดือน ปีเกิด	27 สิงหาคม 2505
สถานที่เกิด	กรุงเทพมหานคร
วุฒิการศึกษา	ปริญญาตรี วิทยาศาสตร์บัณฑิต (พยาบาลและผดุงครรภ์)
ปีที่สำเร็จการศึกษา	ปีการศึกษา 2527
ประสบการณ์ทำงาน	
พ.ศ. 2527 – 2529	พยาบาลประจำการ โรงพยาบาล เปาโลเมโมเรียลแผนกผู้ป่วยหนัก
พ.ศ. 2529 – 2534	พยาบาลประจำการ โรงพยาบาล วิภาวดี แผนกผู้ป่วยหนัก
พ.ศ. 2534 – 2538	หัวหน้าแผนกผู้ป่วยหนัก โรงพยาบาล ปากเกร็ดเวชการ
พ.ศ. 2538 – 2539	ผู้อำนวยการบริหาร โรงพยาบาล รักษ์สกล
พ.ศ. 2539 – 2541	ดำเนินธุรกิจส่วนตัว
พ.ศ. 2542 – ปัจจุบัน	โรงพยาบาลเกษมราษฎร์ (ประชาชื่น) ตำแหน่ง - ผู้จัดการคุณภาพพยาบาล - พยาบาลควบคุมการติดเชื้อในโรงพยาบาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้