



ระบบสารสนเทศเพื่อการจัดการข้อมูลการประกันสุขภาพ ถ้วนหน้าผ่านเครือข่ายอินเทอร์เน็ต

กาญจนา วิริยะพันธ์* และพยุง มีสัง**

บทคัดย่อ

การจัดการข้อมูลการประกันสุขภาพถ้วนหน้าเกิดจากความซับซ้อนของสิทธิการประกันสุขภาพ งานวิจัยนี้จึงเป็นการตรวจสอบและลงทะเบียนสิทธิการประกันสุขภาพผ่านเครือข่ายอินเทอร์เน็ต และวิเคราะห์ข้อมูลที่ได้ด้วยเทคนิคทางดาต้าไมนิ่ง โดยสร้างระบบเพื่อช่วยในการจัดการฐานข้อมูล คลังข้อมูล และการวิเคราะห์ข้อมูล ผลการดำเนินงานพบว่า จากการทดสอบประสิทธิภาพของระบบโดยผู้เชี่ยวชาญและผู้ใช้งานทั่วไปด้วยการใช้แบบประเมินพบว่าค่าเฉลี่ยที่ได้ขึ้นอยู่กับ 4.65 (ค่าเบี่ยงเบนเท่ากับ 0.476) และจากการทดสอบประสิทธิภาพโมเดลเหมืองข้อมูลที่สร้างขึ้นเพื่อใช้แบ่งกลุ่มข้อมูล แสดงให้เห็นว่า การทดสอบด้วยข้อมูลชุดเดียวกับข้อมูลชุดเรียนรู้ โมเดลการค้นหากฎความสัมพันธ์ โมเดลการแบ่งกลุ่มแบบต้นไม้ตัดสินใจ โมเดลการแบ่งกลุ่มเบย์เซียนประเภทหาข้อผิดพลาด โมเดลทั้งสามมีประสิทธิภาพดีที่สุด มีค่าเฉลี่ยเปอร์เซ็นต์ความถูกต้องเท่ากับ 100 เปอร์เซ็นต์ โมเดลโครงข่ายประสาทเทียมมีค่าเฉลี่ยเท่ากับ 99.99 เปอร์เซ็นต์รองลงมา และโมเดลการจัดการกลุ่มข้อมูลมีประสิทธิภาพต่ำสุดมีค่าเฉลี่ยเท่ากับ 99.45 เปอร์เซ็นต์ ดังนั้นสามารถสรุปได้ว่า ระบบที่ได้พัฒนาขึ้นนี้มีประสิทธิภาพดีมาก และได้ข้อมูลเพื่อช่วยในการตัดสินใจของผู้บริหารในการประกันสุขภาพถ้วนหน้า

คำสำคัญ : ระบบสารสนเทศ การจัดการข้อมูล คลังข้อมูล การประกันสุขภาพ ดาต้าไมนิ่ง

1. บทนำ

จากข้อมูลในอดีตที่ผ่านมา ระบบการให้บริการทางด้านสุขภาพในประเทศไทยมีการให้บริการกระจายทั่วทั้งประเทศเป็นไปตามนโยบายการสร้างหลักประกันสุขภาพถ้วนหน้าของรัฐบาลที่มีให้กับประชาชนทุกคน โดยมีแนวคิดที่เน้นด้าน

การรักษาสุขภาพ ดังนั้นจึงถือเป็นภาระหน้าที่การทำงานที่ทางรัฐบาลจะต้องเร่งดำเนินการปฏิรูประบบสาธารณสุขเพื่อสร้างหลักประกันให้กับประชาชนทุกคนให้สามารถเข้าถึงบริการตามความจำเป็น และได้รับการบริการอย่างมีมาตรฐาน โดยถือเป็น “สิทธิ” ขั้นพื้นฐานของประชาชน ทั้งนี้เพื่อให้เป็นไปตามเจตนารมณ์ของรัฐธรรมนูญแห่งราชอาณาจักรไทย พุทธศักราช 2540 ตามมาตรา 52 ที่ว่า “บุคคลย่อมมีสิทธิเสมอกันในการรับบริการสาธารณสุขที่ได้มาตรฐาน และผู้ยากไร้มีสิทธิได้รับการรักษาพยาบาลจากสถานบริการสาธารณสุขของรัฐ โดยไม่เสียค่าใช้จ่าย ทั้งนี้ตามที่กฎหมายบัญญัติ” [1] จากการดำเนินงานที่ผ่านมาพบว่า อุปสรรคที่สำคัญในการให้บริการแก่ประชาชนในโครงการประกันสุขภาพถ้วนหน้า เกิดปัญหาเนื่องจากระบบการให้บริการเดิมเป็นแบบเครื่องเดียว (Stand alone) ไม่สามารถเชื่อมโยงข้อมูลถึงกันได้ โดยการจัดการกองทุนประกันสุขภาพต้องมีการประสานงานด้านฐานข้อมูลกับกองทุนอื่นหลายกองทุน ซึ่งจะก่อให้เกิดความซ้ำซ้อน และความไม่สมบูรณ์ของฐานข้อมูลทะเบียนผู้มีสิทธิ และมีการแบ่งเขตการบริการเป็นพื้นที่ต่าง ๆ ทำให้ประชาชนในบางส่วนไม่ได้รับความสะดวกในการรับบริการของพื้นที่นั้น ๆ ระบบในการส่งข้อมูลระหว่างพื้นที่ทำได้ยาก และมีข้อผิดพลาดเกิดขึ้น นอกจากนั้นยังมีปัญหาเกี่ยวกับโปรแกรมประยุกต์ที่ใช้ในการดำเนินงานลงทะเบียน ดังนั้นจึงมีแนวคิดที่จะพัฒนาระบบสารสนเทศเพื่อการจัดการข้อมูลการประกันสุขภาพแบบเชื่อมโยงข้อมูลกันผ่านทางเว็บไซต์ (Web Site) บนเครือข่ายอินเทอร์เน็ต พร้อมทั้งพัฒนาระบบการลงทะเบียน การค้นหาข้อมูล การปรับปรุงข้อมูลของการประกันสุขภาพถ้วนหน้า เพื่อแก้ไขปัญหาดังกล่าวที่เกิดขึ้น อีกทั้งหน่วยงานที่ให้บริการสามารถดูแล และปรับปรุงฐานข้อมูลได้ตลอดเวลาทำให้สามารถตอบสนองความต้องการของประชาชนอย่างแท้จริง

* นักศึกษา ระดับปริญญาโท คณะเทคโนโลยีสารสนเทศ สจพ.

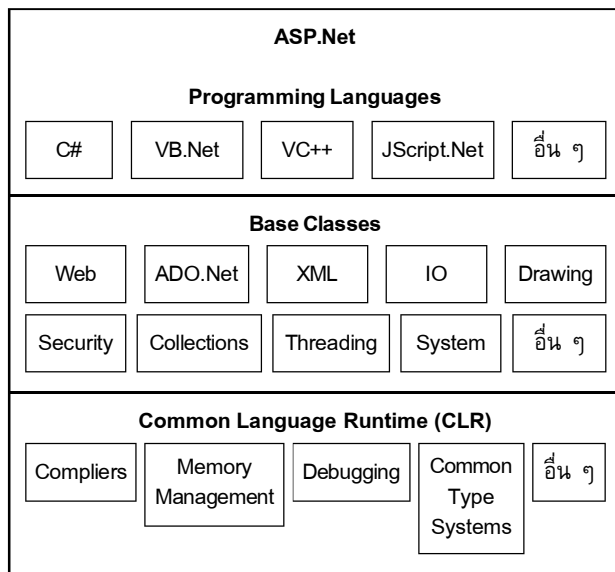
** อาจารย์ ระดับ 7 ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ สจพ. และ รองคณบดีฝ่ายบริหาร คณะเทคโนโลยีสารสนเทศ สจพ.



2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ไมโครซอฟท์ดอทเน็ต

.Net Framework คือ โครงสร้างการพัฒนาโปรแกรมคอมพิวเตอร์ที่ถูกออกแบบมาเพื่อช่วยอำนวยความสะดวกต่อการพัฒนาโปรแกรมในสมัยใหม่ที่ใช้งานในระบบเครือข่าย [2] ประกอบไปด้วยส่วนประกอบต่าง ๆ หลายส่วนด้วยกัน ส่วนประกอบเหล่านี้ถูกออกแบบมาเพื่อให้ทำงานได้เข้ากันได้ดียิ่งขึ้น โดยส่วนประกอบหลัก ๆ ของ .Net Framework แบ่งเป็นชั้น ๆ ได้ดัง diagram ต่อไปนี้



ภาพที่ 1 องค์ประกอบของ .Net Framework

2.2 แนวคิดเกี่ยวกับคลังข้อมูล

คลังข้อมูล (Data Warehouse) คือ การรวบรวมข้อมูลทั้งหมดจากฐานข้อมูลของส่วนปฏิบัติงานต่าง ๆ (Operational Database) หลาย ๆ แหล่ง จากนั้นทำการแปลงข้อมูล หรือสรุปข้อมูลเพื่อให้อยู่ในรูปแบบที่เหมาะสมในการเก็บและง่ายต่อการนำกลับมาใช้ แล้วนำข้อมูลที่ผ่านการแปลงหรือสรุปแล้วเข้าไปเก็บในคลังข้อมูลเพื่อใช้เป็นข้อมูลทางธุรกิจสำหรับนักวางแผน นักวิเคราะห์ ผู้จัดการแผนกหรือผู้บริหารในการตัดสินใจด้านธุรกิจอย่างถูกต้องสำหรับองค์กรต่อไป

คลังข้อมูลถือว่าเป็นเทคโนโลยีที่มีความสำคัญเป็นอย่างยิ่งในการวางแผนพื้นฐานกับองค์กร (Infrastructure) กล่าวคือทุกองค์กรจำเป็นต้องมีระบบฐานข้อมูลขององค์กร (Corporate Database Information) มีการจัดการข้อมูลที่ดี และนำข้อมูลเหล่านั้นเก็บไว้ในคลังข้อมูล (Data Warehouse) เพื่อให้เรียกใช้ข้อมูลได้อย่างมีประสิทธิภาพ

2.3 ดาต้าไมนิ่ง (Data Mining)

ดาต้าไมนิ่ง [3] เป็นการค้นหาสารสนเทศที่มีคุณค่าจากข้อมูลที่มีขนาดใหญ่ซึ่งเป็นฐานข้อมูลอะไรก็ได้ เป็นการค้นหาและวิเคราะห์โดยเอามาคำนวณหรือทำด้วยวิธีการต่าง ๆ และการกลั่นออกมาเป็นรูปแบบ หรือกฎที่สำคัญในธุรกิจ เพื่อนำไปช่วยในการตัดสินใจของผู้บริหาร ซึ่งในการทำสามารถทำเป็นแบบอัตโนมัติหรือแบบกึ่งอัตโนมัติได้

ดาต้าไมนิ่งถือเป็นขั้นตอนหนึ่งใน KDD Process (Knowledge Discovery in Database) [4] ซึ่ง KDD คือ การใช้เครื่องมือเข้าไปค้นหาสิ่งที่มีความหมาย หรือองค์ความรู้ในฐานข้อมูลที่มีขนาดใหญ่มหาศาลซึ่งซ่อนอยู่ภายในและเราไม่สามารถมองออกได้

2.3.1 เทคนิคต่าง ๆ ของดาต้าไมนิ่ง

2.3.1.1 การจัดกลุ่มข้อมูล (Clustering) จัดเป็นแนวคิดที่เกี่ยวกับการจัดกลุ่มข้อมูลตามลักษณะหรือคุณสมบัติที่คล้าย หรือใกล้เคียงกัน เพื่อให้ได้มาซึ่งผลของการคาดคะเนที่น่าจะเกิดขึ้นในอนาคต และได้มีการกำหนดข้อมูลที่ใช้เป็นข้อมูลชุดการเรียนรู้ (Training Data) ในการประมวลผลข้อมูลของเทคนิคการจัดกลุ่มนี้ คำนวณได้จากวิธี K-Mean ซึ่งเป็นการใช้วิธีการจัดกลุ่มโดยใช้ค่าจริง อยู่ในรูปที่เฉลี่ยได้ มีขั้นตอนดังนี้

- (1) กำหนดจำนวนคลัสเตอร์ทั้งหมดซึ่งดูจากความคล้ายของข้อมูลที่มีอยู่
- (2) นำข้อมูลแต่ละตัวมาหาระยะห่างของข้อมูลจากคลัสเตอร์ที่กำหนดไว้ จากสูตร

$$d(i, j) = \sqrt{(x_{i1} - c_{j1})^2 + \dots + (x_{ip} - c_{jp})^2} \quad (1)$$

เมื่อ d แทน Euclidean distance

c แทนจุดกลางของคลัสเตอร์

x แทนข้อมูล

i แทนลำดับของข้อมูล

j แทนลำดับของคลัสเตอร์

p แทนจำนวนมิติของข้อมูล

- (3) เลือกปรับคลัสเตอร์ ตัวที่มีระยะห่างน้อยที่สุด โดยปรับจากสูตร

$$C_{new} = \frac{X + C_{old(X)}}{2} \quad (2)$$

เมื่อ C แทนคลัสเตอร์ที่อยู่ใกล้ X มากที่สุด หรือมีค่า

ระยะ d ต่ำที่สุด

X แทนเวกเตอร์ข้อมูล

(4) ทำขั้นตอน (2) และ (3) จนครบทุกข้อมูล (1 รอบ)

(5) ทำวนจนครบรอบที่กำหนด

(6) นำข้อมูลสุดท้ายที่ได้ไปหาระยะห่างกับคลัสเตอร์ที่กำหนดไว้ (ที่ปรับค่าเรียบร้อยแล้ว) เลือกสรุปว่าข้อมูลนั้นอยู่ในคลัสเตอร์ไหน โดยดูจากค่าที่มีระยะห่างน้อยสุด

2.3.1.2 การค้นหากฎความสัมพันธ์ (Association Rule) หลักการทำงานวิธีนี้คือ การค้นหาความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่ที่มี เพื่อนำไปใช้ในการวิเคราะห์หรือทำนายปรากฏการณ์ ซึ่งสามารถแบ่งได้ 2 ขั้นตอน คือ การหาเซตไอเท็มที่มีค่าสนับสนุนมากกว่าค่าสนับสนุนน้อยที่สุดที่กำหนดให้ เรียกเซตนี้ว่า เซตไอเท็มปรากฏบ่อย (Frequent) และการนำเซตไอเท็มเหล่านี้มาสร้างเป็นกฎความสัมพันธ์ต่อไป

ในการประมวลผลของเทคนิคการค้นหาความสัมพันธ์นี้ได้ใช้อัลกอริทึม Apriori [5] ซึ่งเป็นอัลกอริทึมที่ใช้แพร่หลาย โดยทำการค้นหาในแนวทางกว้างก่อน และใช้การนับทรานแซคชัน ซึ่งจะสร้างและตรวจสอบเซตไอเท็มปรากฏบ่อยทีละชั้น เริ่มจากเซตไอเท็มที่มีจำนวนสมาชิกเท่ากับหนึ่ง ถ้าเซตไอเท็มใดมีค่าสนับสนุน (Support) น้อยกว่าค่าสนับสนุนที่กำหนดให้ก็จะตัดเซตไอเท็มนั้นออก ไม่นำไปสร้างไอเท็มในชั้นถัดไป การทำงานของอัลกอริทึมนี้จะวนอย่างนี้ไปเรื่อย ๆ จนกระทั่งไล่ไปทุกระดับชั้น หรือไม่เหลือเซตไอเท็มที่จะสร้างเซตไอเท็มในชั้นถัดไป และทำการหาค่าความเชื่อมั่น (Confidence) ของไอเท็มเซตที่สร้างขึ้น โดยมีสูตรการคำนวณดังนี้

$$Support(A \Rightarrow B) = \frac{\#tuple_containing_both_A_and_B}{total_ \#_of_tuples} \quad (3)$$

$$Confidence(A \Rightarrow B) = \frac{\#tuple_containing_both_A_and_B}{\#tuples_containing_A} \quad (4)$$

2.3.1.3 การแบ่งกลุ่มแบบต้นไม้ตัดสินใจ (Decision Tree) โดยรูปแบบของ Tree นั้นประกอบด้วยโหนดแรกสุดที่เรียกว่า Root node จาก Root node จะแตกออกไปเป็นโหนดลูก และที่โหนดลูกก็จะมีโหนดลูกของตัวเอง ซึ่งโหนดลูกแต่ละระดับอาจจะมีมากกว่า 2 โหนดก็ได้ ส่วนในโหนดระดับสุดท้ายเรียกว่า โหนดใบ (Leaf node)

การนำข้อมูลมาสร้างต้นไม้ตัดสินใจ มีขั้นตอนพื้นฐาน คือ

(1) หาลักษณะประจำที่สำคัญที่สุดมาแบ่งข้อมูล โดยลักษณะประจำนี้จะถูกนำมาสร้างเป็นโหนดภายใน โดยจะมี

เป้าหมายลักษณะประจำเป็นผลลัพธ์ซึ่งเป็นโหนดใบถูกกำหนดไว้ก่อน

(2) นำค่าที่เป็นไปได้ในลักษณะประจำที่ถูกเลือกมาแตกออกเป็นกลุ่มของตัวเอง

(3) แบ่งข้อมูลทั้งหมดตามกลุ่มที่แตกออกจากโหนดภายใน

(4) วนกลับไปทำแบบขั้นตอนแรก คือ หาลักษณะประจำที่สำคัญที่สุดจากข้อมูลที่เข้ามา เพื่อหาลักษณะประจำที่ใช้แบ่งข้อมูลต่อไป

สำหรับอัลกอริทึมที่ใช้สร้างต้นไม้ตัดสินใจมีหลายอัลกอริทึม ซึ่งในที่นี้ได้เลือกอัลกอริทึม การทำงานของ ID3, C4.5 โดยเลือกเทคนิค Information Gain มาใช้ในการจัดกลุ่มข้อมูลจากสูตร

$$I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n} \quad (5)$$

เมื่อ i แทนสถานะทั้งหมดของโหนดภายใน

p, n แทนค่าที่เป็นไปได้ของคลาสที่กำหนด

จากนั้นหา $Remainder(A)$ จากสูตร

$$Remainder(A) = \sum_{i=1}^v \left[\frac{p_i + n_i}{p+n} I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) \right] \quad (6)$$

เมื่อ A แทนลักษณะประจำทั้งหมด

v แทนค่าที่เป็นไปได้ทั้งหมดของลักษณะประจำนั้น

จากนั้นหา $Gain$ จากสูตร

$$Gain = I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) - Remainder(A) \quad (7)$$

เลือกตัวที่มีค่า $Gain$ มากที่สุดของลักษณะประจำทั้งหมดเพื่อนำมาเป็นตัวโหนดภายในเริ่มต้น จากนั้นจึงทำการตัดตัวที่เป็นโหนดภายในแล้วออก และเริ่มหาค่า $I, remainder(A)$ และค่า $Gain$ ใหม่อีกครั้ง ทำไปเรื่อย ๆ จนสามารถแบ่งคลาสได้ครบ

2.3.1.4 โครงข่ายประสาทเทียม (Neural Network) แนวคิดของโครงข่ายประสาทเทียมได้มาจากการทำงานของเซลล์สมองของมนุษย์ ซึ่งหน่วยที่ย่อยที่สุดของโครงข่ายประสาทเทียมนี้เรียกว่าเพอร์เซพตรอน (Perceptron) ซึ่งเทียบได้กับเซลล์สมองของมนุษย์หนึ่งนิวรอน



(Neuron) เพอร์เซพตรอนนี้จะทำหน้าที่รับอินพุตซึ่งเป็นเวกเตอร์ของจำนวนจริงเข้ามา พร้อมทั้งคำนวณค่าเหล่านี้โดยให้นำหนักของอินพุตแต่ละตัวแตกต่างกันและเอาต์พุตที่ได้จะถูกนำไปคำนวณค่าผิดพลาด (error) เพื่อนำมาปรับน้ำหนักของอินพุตต่อไป

แบ็กพรอพพาเกชันนิวรอลเน็ตเวิร์ก (Backpropagation neural network) เป็นเน็ตเวิร์กที่มีได้หลายนิวรอนและมีได้หลายชั้น (multilayer) และทำงานกับฟังก์ชันซิกมอยด์ (Sigmoid Function) ซึ่งเป็นฟังก์ชันที่สามารถแยกตัวได้อย่างได้แบบไม่เชิงเส้น ทำให้ทำงานได้ดีกว่าเพอร์เซพตรอนเดี่ยว ๆ

2.3.1.5 การแบ่งกลุ่มเบเชียนประเภทนาอิวเบเชียน (Naive Bayesian Classifier) [6] เป็นอัลกอริทึมที่สามารถเรียนรู้และเข้าใจได้ง่าย อีกทั้งการเรียนรู้และแบ่งกลุ่มสามารถทำได้อย่างรวดเร็ว โดยเป็นการคำนวณหาความน่าจะเป็นของแต่ละกลุ่มข้อมูล ซึ่งในที่นี้เรียกว่า คลาส (Class) เมื่อกำหนดลักษณะประจำ (Attribute) และค่าของลักษณะประจำแต่ละตัวมาให้ การทำนายจะคำนวณหาความน่าจะเป็นของทุก ๆ คลาสมาเปรียบเทียบกัน แล้วเลือกค่าความน่าจะเป็นที่สูงที่สุดของคลาสใด ๆ มาเป็นผลของการทำนายเพียงค่าเดียว โดยที่ถือว่าลักษณะประจำแต่ละตัวจะมีความเป็นอิสระต่อกัน

2.4 งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยมากมายที่ได้นำเสนอการลงทะเบียนข้อมูลต่าง ๆ ผ่านเครือข่ายอินเทอร์เน็ต สำหรับระบบการลงทะเบียนเรียนของนักศึกษาแห่งหนึ่งผ่าน Web [7] เป็นระบบที่ช่วยในการให้บริการการลงทะเบียนเรียนได้ด้วยตนเอง ผ่านทางเครือข่ายอินเทอร์เน็ต โดยไม่จำกัดเวลาและสถานที่ ทำให้เป็นการอำนวยความสะดวกแก่นักศึกษา และช่วยลดภาระงานของเจ้าหน้าที่ อีกระบบเป็นการจัดเก็บข้อมูลของผู้ป่วยและเสียชีวิตในโรงพยาบาล [8] เป็นวิธีการจัดการสารสนเทศนำคอมพิวเตอร์มาใช้งานเพื่ออำนวยความสะดวกแก่ผู้ใช้ในการดำเนินการสืบค้น แทนการค้นหาจากเอกสารที่ได้บันทึกไว้แต่เดิม ซึ่งในงานวิจัยต่างก็เป็นการจัดการข้อมูลแต่ไม่สามารถนำข้อมูลที่ได้นำไปใช้ในการตัดสินใจหรือทำนายเหตุการณ์ล่วงหน้า

อย่างไรก็ตามในงานวิจัยนี้จึงทำการศึกษาในอีกด้านหนึ่ง นั่นคือ ทำการวิเคราะห์และแบ่งกลุ่มข้อมูลเพื่อเป็นการนำข้อมูลที่ได้นำไปให้เกิดประโยชน์สูงสุด และสามารถเข้าถึงข้อมูลที่มีขนาดใหญ่มาได้อย่างรวดเร็วยิ่งขึ้น จึงได้มีการเลือกใช้เทคนิค

การจัดกลุ่มข้อมูล ต้นไม้ตัดสินใจ กฎการหาความสัมพันธ์ โครงข่ายประสาทเทียม การแบ่งกลุ่มข้อมูลแบบเบเชียน จึงได้นำเทคนิคดังกล่าวมารวมในการวิเคราะห์เพื่อจัดกลุ่มและแบ่งกลุ่มข้อมูล ในกรณีที่ข้อมูลเชิงปริมาณและทำการศึกษาเปรียบเทียบค่าความถูกต้องในการแบ่งกลุ่มข้อมูลของเทคนิคทั้งหมด ทั้งนี้เพื่อนำผลที่ได้มาวิเคราะห์หาวิธีที่เหมาะสมและนำไปใช้ในการจัดกลุ่มและแบ่งกลุ่มข้อมูลให้มีความถูกต้องและแม่นยำมากยิ่งขึ้นต่อไป

3. การออกแบบวิธีการดำเนินงาน

ในการทำงานของระบบ ผู้ใช้จะถูกแบ่งออกเป็น 2 กลุ่ม คือ ประชาชนทั่วไป และเจ้าหน้าที่ โดยความสามารถในการเข้าถึงระบบของผู้ใช้แต่ละกลุ่มจะมีความแตกต่างกันดังนี้คือ

(1) ประชาชนทั่วไป สามารถเข้ามาค้นหา ตรวจสอบสิทธิ์ และลงทะเบียนสิทธิ์ข้อมูลการประกันสุขภาพถ้วนหน้า การอ่านข่าวสารต่าง ๆ

(2) เจ้าหน้าที่ แบ่งได้เป็น 2 ประเภทคือ

เจ้าหน้าที่ทั่วไป สามารถล็อกอินเข้ามาจัดการฐานข้อมูล ผู้ที่มีสิทธิ์การประกันสุขภาพ การค้นหาข้อมูล การออกรายงาน การประกาศข่าวสารต่าง ๆ

ผู้บริหารระดับสูง สามารถล็อกอินเข้ามาค้นหาข้อมูลผู้ที่มีสิทธิ์การประกันสุขภาพ การออกรายงานทั่วไป รวมถึงการออกรายงานที่มีการแบ่งกลุ่มโดยใช้เทคนิคทางดาต้าไมนิง

จากการศึกษาและวิเคราะห์ข้อมูลทำให้ได้ทราบถึงขั้นตอนการทำงานของระบบ ข้อมูลที่จะนำเข้าสู่ระบบ บุคคลที่เกี่ยวข้องกับระบบ และแหล่งจัดเก็บข้อมูล จึงได้ใช้แผนภาพ UML (Unified Modeling Language) แสดงให้เห็นถึงขั้นตอนการทำงานของระบบ ข้อมูลที่เข้าและออกจากระบบ รวมทั้งข้อมูลที่ไหลเวียนภายในระบบจากขั้นตอนหนึ่งไปยังอีกขั้นตอนหนึ่ง และได้ใช้โปรแกรม Rational Rose 2000 เป็นเครื่องมือช่วยในการสร้างไดอะแกรมของ UML โดยได้อธิบายการออกแบบโครงสร้างออกเป็น 3 ส่วนคือ Use Case Diagram, Sequence Diagram และ Class Diagram

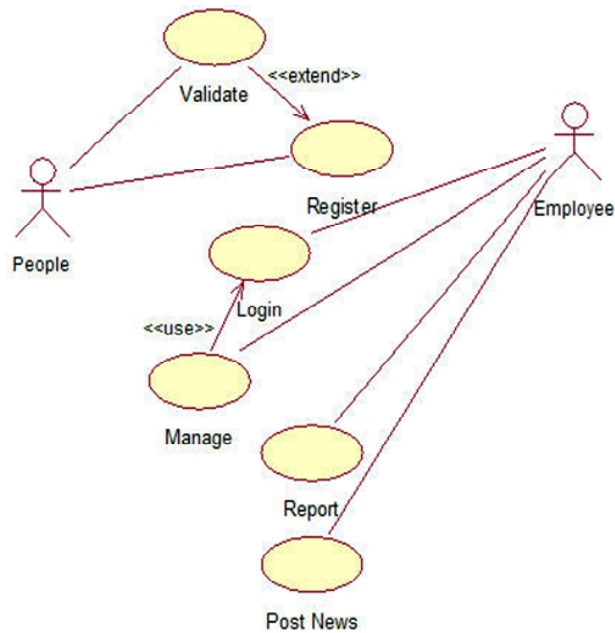
Use Case Diagram ของระบบประกอบด้วยผู้ใช้งาน (Actor) 2 กลุ่ม คือ ประชาชนทั่วไป และเจ้าหน้าที่

ผู้ใช้งานเข้ามาทำกิจกรรมกับระบบ หรือเรียกว่า Use Case โดยระบบสารสนเทศเพื่อการจัดการข้อมูลการประกันสุขภาพถ้วนหน้าผ่านเครือข่ายอินเทอร์เน็ตจะประกอบไปด้วย Use Case ทั้งหมด 6 Use Case ดังภาพที่ 2

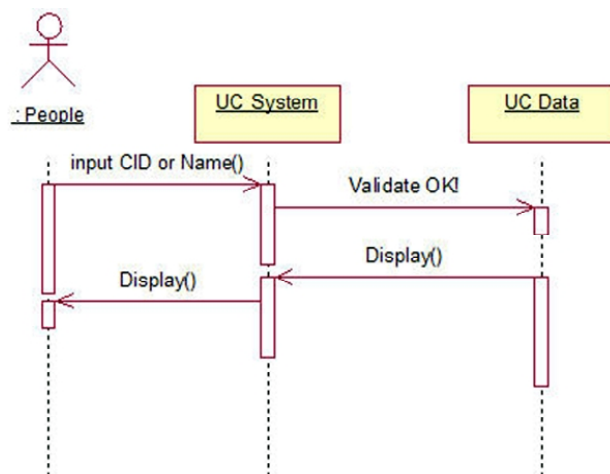


Sequence Diagram ซึ่งจะแบ่งได้ออกเป็น 6 Diagram ดังนี้

Validate ผู้ใช้ทำการเข้ามาตรวจสอบสิทธิ์ที่ระบบ UC ซึ่งจะสามารถทำการค้นหาโดยใช้เลขบัตรประจำตัวประชาชน หรือชื่อของประชาชนทั่วไป เมื่อใส่แล้วระบบจะทำการตรวจสอบสิทธิ์จากข้อมูล UC Data และทำการแสดงผลการตรวจสอบไปที่ระบบ UC



ภาพที่ 2 Use Case Diagram ของระบบ

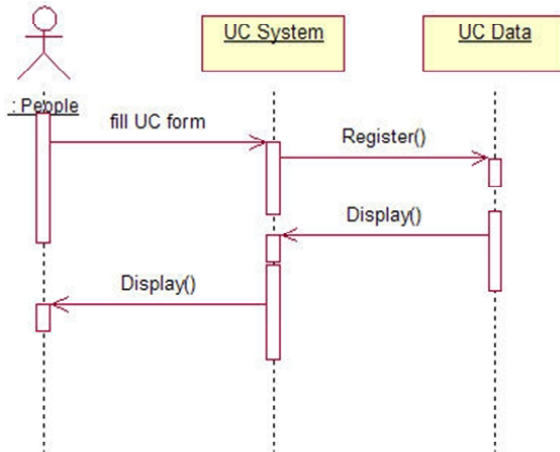


ภาพที่ 3 Sequence Diagram ของระบบ Validate

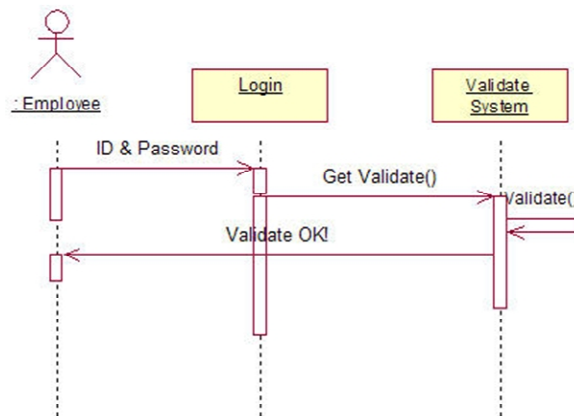
Register ผู้ใช้สามารถลงทะเบียนเพื่อขอสิทธิ์การประกันสุขภาพได้โดยการกรอกข้อมูลในแบบฟอร์มให้ครบถ้วนที่

ระบบ จากนั้นระบบจะทำการลงทะเบียนและบันทึกข้อมูลผู้ขอสิทธิ์ที่ UC Data และแสดงผลไปที่ระบบ UC

Login เจ้าหน้าที่ที่ต้องการจะเข้าใช้งานระบบ UC จะต้องใส่เลขประจำตัวพนักงาน และ Password ของตัวเอง โดยระบบใน Validate จะเอาข้อมูลที่ใส่ไปทำการตรวจสอบสิทธิ์ของเจ้าหน้าที่ผู้นั้น เมื่อระบบตรวจสอบถูกต้องจะสามารถให้เจ้าหน้าที่ผู้นั้นใช้งานระบบได้



ภาพที่ 4 Sequence Diagram ของระบบ Register



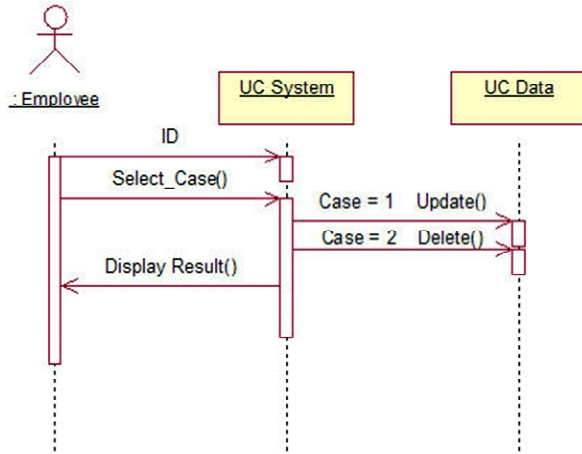
ภาพที่ 5 Sequence Diagram ของระบบ Login

UC Management เจ้าหน้าที่ใส่เลขบัตรประชาชนผู้มีสิทธิ์แล้วเลือกว่าต้องการทำงาน Update หรือ Delete ข้อมูลของผู้มีสิทธิ์นั้น เมื่อเจ้าหน้าที่ปรับปรุงข้อมูลเรียบร้อยแล้วระบบจะแสดงข้อมูลที่ได้แก้ไขปรับปรุงแล้ว

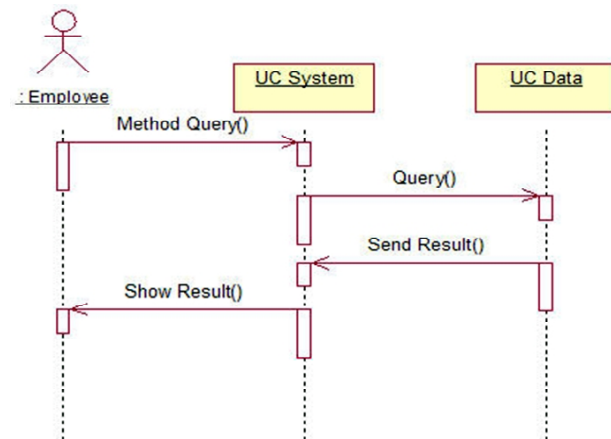
Report เจ้าหน้าที่เข้ามาเลือกประเภท Report ที่ต้องการดูระบบ UC จะทำการดึงข้อมูลจาก UC Data และส่งผลลัพธ์ไปที่ระบบ UC เพื่อแสดงผลออกมา



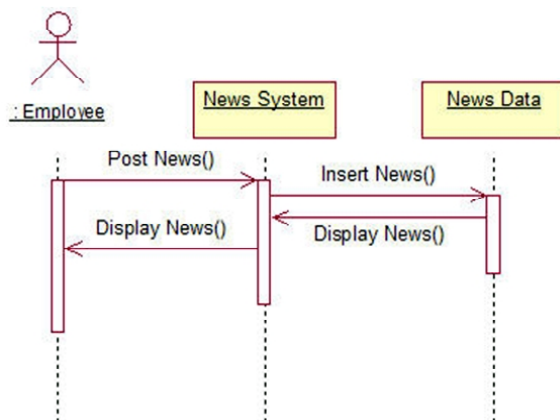
Post News เจ้าหน้าที่ประกาศข่าวสารต่าง ๆ เข้าไปที่ระบบ UC จากนั้นระบบจะเพิ่มข้อมูลพื้นฐานข้อมูล News Data และแสดงผลไปที่ระบบ UC



ภาพที่ 6 Sequence Diagram ของ ระบบ UC Management



ภาพที่ 7 Sequence Diagram ของระบบ Report



ภาพที่ 8 Sequence Diagram ของระบบ News

จากการที่ได้วิเคราะห์ระบบเรียบร้อยแล้ว ได้ทำการจำลองข้อมูลที่เกิดขึ้นทั้งหมดในระบบ ด้วยแผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (Class Diagram) ดังภาพที่ 9 และได้ทำการออกแบบฐานข้อมูล โดยประกอบด้วยตารางทั้งหมดจำนวน 9 ตารางที่มีความสัมพันธ์กัน

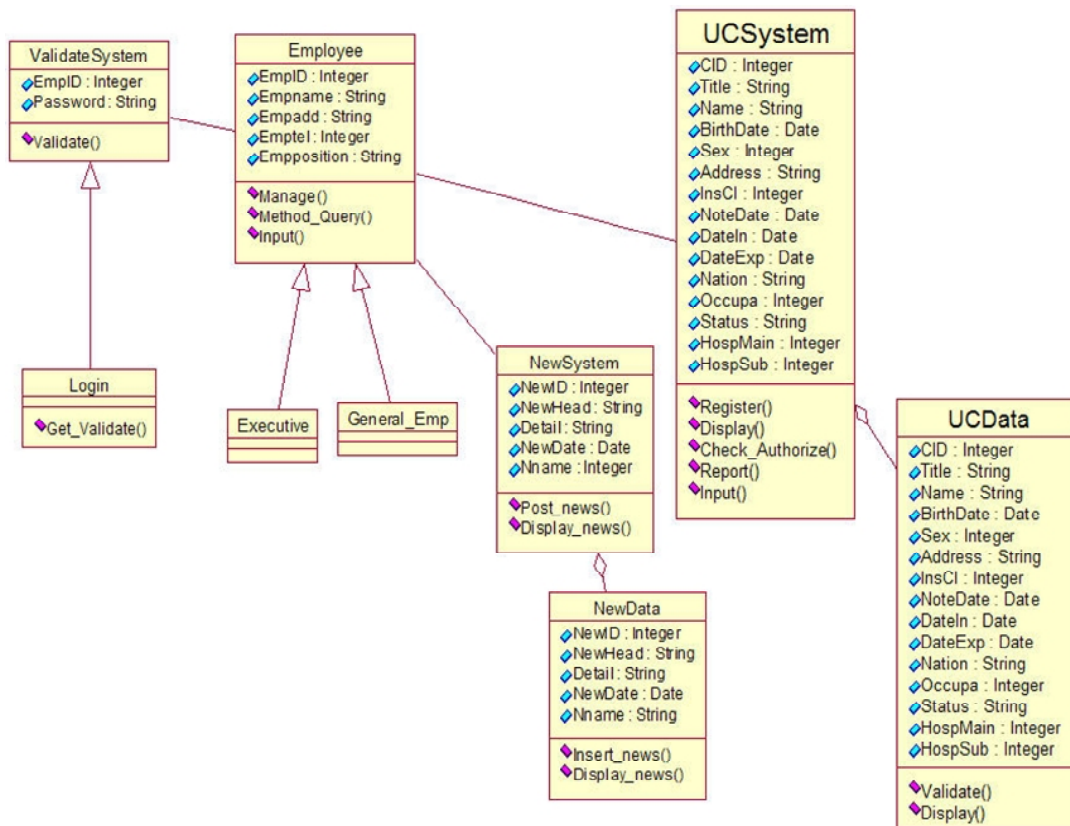
การพัฒนาาระบบ

การพัฒนาาระบบสารสนเทศสำหรับการวิเคราะห์ข้อมูลการประกันสุขภาพถ้วนหน้าผ่านเครือข่ายอินเทอร์เน็ต โดยใช้โปรแกรม Microsoft Visual Studio.Net พัฒนา ASP.Net ด้วยภาษา VB.net ร่วมกับโปรแกรมจัดการฐานข้อมูล Microsoft SQL Server 2005 ซึ่งโปรแกรมทั้งสองนี้มีความยืดหยุ่นในการทำงานสูงและสามารถรองรับการสนับสนุนการทำงานที่มีการติดต่อฐานข้อมูลอินเทอร์เน็ตได้ดี ซึ่ง ASP.Net ในปัจจุบันถือว่าเป็นโปรแกรมที่มีความรู้และสามารถในการทำงานได้หลากหลายรูปแบบ คือใช้ได้กับระบบปฏิบัติการหลากหลายและสามารถทำงานร่วมกันได้หลาย ๆ ภาษา คุณสมบัติเด่นของ ASP.Net เป็นการพัฒนาร่วมกันในหลายภาษา มีการออกแบบระบบเป็นเชิงวัตถุและยังเป็นการติดต่อกับโปรแกรมการจัดการระบบฐานข้อมูลที่มีอยู่มากมายในปัจจุบัน ทางด้านโปรแกรมจัดการฐานข้อมูล Microsoft SQL Server 2005 สามารถติดตั้งได้ง่ายและสามารถรองรับระบบปฏิบัติการการทำงานที่เป็นแบบ Windows จุดเด่นที่สำคัญของโปรแกรมการจัดการฐานข้อมูล Microsoft SQL Server 2005 คือ เป็นฐานข้อมูลตัวล่าสุดซึ่งช่วยให้งานของผู้บริหารระบบฐานข้อมูลทำได้ง่ายขึ้นและสามารถรองรับการทำงานกับระบบขนาดใหญ่ที่ต้องการความน่าเชื่อถือสูง ๆ ได้อย่างดี โดยรองรับกับแนวคิดของการบริหารงาน วิเคราะห์ และการจัดการข้อมูลขนาดใหญ่ได้อย่างมีประสิทธิภาพ ทำให้ตอบสนองกับความต้องการใช้งานในปัจจุบัน

การทดสอบระบบ

เมื่อพัฒนาระบบสารสนเทศเพื่อการจัดการข้อมูลการประกันสุขภาพถ้วนหน้าผ่านเครือข่ายอินเทอร์เน็ตเสร็จเรียบร้อยแล้ว เพื่อทดสอบความถูกต้องในการทำงานจึงได้ทำการตรวจสอบในเรื่องต่าง ๆ โดยแบ่งการทดสอบเป็น 2 ส่วน คือ

- (1) ตรวจสอบความถูกต้องของระบบงาน
- (2) ตรวจสอบความถูกต้องของโมเดลที่สร้างขึ้น โดยวิธีการทางด้าไมนิ่งและนำมาใช้ในการออกรายงาน โดยใช้



ภาพที่ 9 Class Diagram ของระบบ

ข้อมูลประชาชนทั่วไปที่ลงทะเบียนบัตรประกันสุขภาพถ้วนหน้าจำนวน 57,000 รายการ โดยการทดสอบจะแบ่งข้อมูลเป็นชุด ๆ ดังนี้

ข้อมูลประชาชนทั่วไปที่ลงทะเบียนบัตรประกันสุขภาพถ้วนหน้าแบ่งเป็นข้อมูลชุดเรียนรู้ จำนวน 57,000 รายการ และใช้ข้อมูลทดสอบชุดเดียวกับชุดเรียนรู้

ข้อมูลประชาชนทั่วไปที่ลงทะเบียนบัตรประกันสุขภาพถ้วนหน้าแบ่งเป็นข้อมูลชุดเรียนรู้ 500 เรคคอร์ด ข้อมูลชุดทดสอบ 500 เรคคอร์ด โดยทำการสุ่มเลือกเพื่อให้ได้ข้อมูลต่างกันจำนวน 10 ชุด

4. ผลการดำเนินงาน

ผลการออกแบบหน้าจอการใช้งานและรายงานต่าง ๆ แสดงดังภาพที่ 10 ถึง 15 ในแต่ละหน้าจอได้ออกแบบเพื่อป้อนข้อมูลและนำข้อมูลที่ได้มาประมวลผลและสรุปเป็นรายงานต่อไป

5. สรุปผล

จากการทำแบบประเมินของผู้เชี่ยวชาญและผู้ใช้งานทั่วไป

นั้น ผลการทดสอบทุกด้านมีค่าเฉลี่ยเท่ากับ 4.65 (ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.476) สามารถสรุปผลการประเมินประสิทธิภาพของระบบโดยรวมได้ว่า ระบบสารสนเทศเพื่อการจัดการข้อมูลการประกันสุขภาพถ้วนหน้าผ่านเครือข่ายอินเทอร์เน็ตนี้มีประสิทธิภาพอยู่ในระดับดีมาก สามารถนำไปใช้งานได้อย่างมีประสิทธิภาพและตอบสนองต่อความต้องการของผู้ใช้งาน

การเปรียบเทียบประสิทธิภาพของโมเดลที่สร้างขึ้นระหว่างเทคนิคการจัดกลุ่มข้อมูล (Clustering) การค้นหากฎความสัมพันธ์ (Association Rule) การแบ่งกลุ่มแบบต้นไม้ตัดสินใจ (Decision Tree) โครงข่ายประสาทเทียม (Neural Network) และการแบ่งกลุ่มเบย์เซียนประเภทนาอิวเบเซียน (Naive Bayesian Classifier) โดยกำหนดให้

NN คือ เทคนิคโครงข่ายประสาทเทียม

AR คือ เทคนิคการค้นหากฎความสัมพันธ์

DT คือ เทคนิคการแบ่งกลุ่มแบบต้นไม้ตัดสินใจ

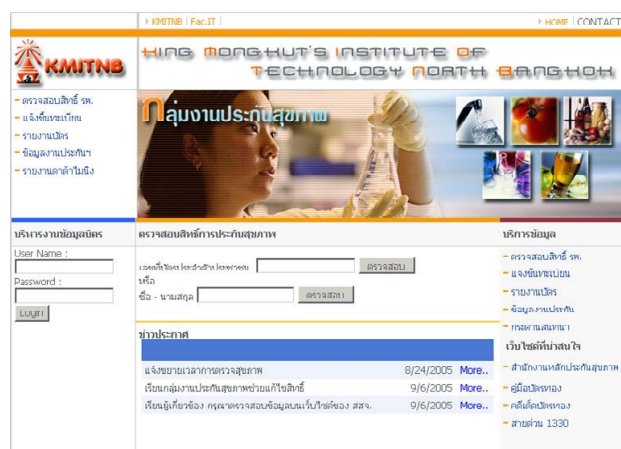
CT คือ เทคนิคการจัดกลุ่มข้อมูล

NB คือ เทคนิคการแบ่งกลุ่มเบย์เซียนประเภทนาอิวเบเซียน

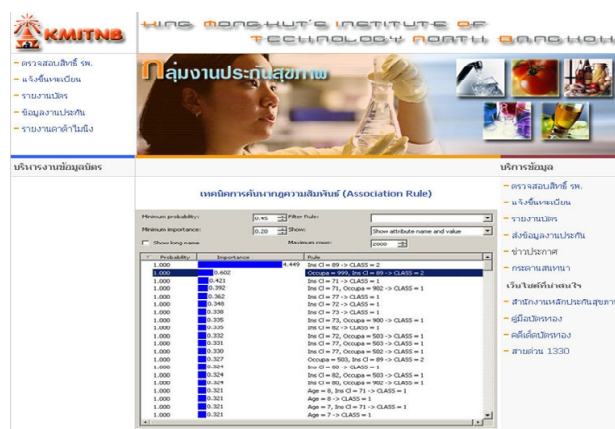


จากผลการหาประสิทธิภาพโมเดลดาต้าไมนิ่งด้วยข้อมูลชุดเรียนรู้ โมเดลที่สร้างโดยเทคนิค AR มีประสิทธิภาพที่ดีที่สุด เนื่องจากมีค่าเฉลี่ยเท่ากับ 100% ในขณะที่โมเดลที่สร้างโดยเทคนิค NN มีประสิทธิภาพรองลงมา คือ มีค่าเฉลี่ยเท่ากับ 99.71% และโมเดลที่สร้างโดยเทคนิค CT มีประสิทธิภาพต่ำสุด คือมีค่าเฉลี่ยเท่ากับ 94.45% การทดสอบด้วยข้อมูลชุดทดสอบ พบว่า โมเดลที่สร้างโดยเทคนิค AR มีประสิทธิภาพที่ดีที่สุด เนื่องจากมีค่าเฉลี่ยเท่ากับ 100% ในขณะที่โมเดลที่สร้างโดยเทคนิค NN มีประสิทธิภาพรองลงมา คือ มีค่าเฉลี่ยเท่ากับ 99.72% และโมเดลที่สร้างโดยเทคนิค CT มีประสิทธิภาพ

ต่ำสุด คือ มีค่าเฉลี่ยเท่ากับ 95.10% จากผลการเปรียบเทียบระหว่างโมเดลที่สร้างขึ้นโดยเทคนิคโครงข่ายประสาทเทียม (Neural Network) การค้นหากฎความสัมพันธ์ (Association Rule) การแบ่งกลุ่มแบบต้นไม้ตัดสินใจ (Decision Tree) การจัดกลุ่มข้อมูล (Clustering) และการแบ่งกลุ่มเบย์เซียนประเภทนาอีฟเบย์เซียน (Naive Bayesian Classifier) จะเห็นว่า เทคนิคการค้นหากฎความสัมพันธ์ สามารถสร้างโมเดลที่มีเปอร์เซ็นต์ความถูกต้องในการแยกข้อมูลที่ดีที่สุด ส่วนเทคนิคการจัดกลุ่มข้อมูลมีเปอร์เซ็นต์ความถูกต้องในการแยกข้อมูลน้อยที่สุด



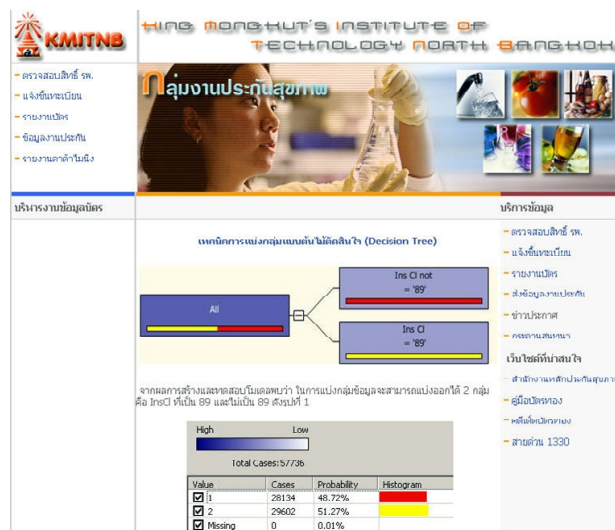
ภาพที่ 10 หน้าจอหลักของระบบ



ภาพที่ 12 รายงานดาต้าไมนิ่งโดยเทคนิคการค้นหากฎความสัมพันธ์



ภาพที่ 11 รายงานดาต้าไมนิ่งโดยเทคนิคการจัดกลุ่มข้อมูล



ภาพที่ 13 รายงานดาต้าไมนิ่งโดยเทคนิคการแบ่งกลุ่มแบบต้นไม้ตัดสินใจ



ภาพที่ 14 รายงานดาต้าไมนิ่งโดยเทคนิคโครงข่ายประสาทเทียม



ภาพที่ 15 รายงานดาต้าไมนิ่งโดยเทคนิคการแบ่งกลุ่มเบย์เซียนประเภทนาอ์ฟเบเซียน

เอกสารอ้างอิง

- [1] สำนักนโยบายและแผน กระทรวงสาธารณสุข. “เครื่องชี้วัดการประเมินผลการพัฒนาสุขภาพตามแผนพัฒนาสุขภาพแห่งชาติ ในช่วงแผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ ฉบับที่ 9 (พ.ศ.2545-2549).” กรุงเทพฯ : โรงพิมพ์สยามพาณิชย์, 2544.
- [2] ทวีชัย หงษ์สุมาลย์, สงวนชัย วุสธรณชีวะศิริ. “อินไซต์ ASP.NET ฉบับสมบูรณ์.” กรุงเทพฯ : บริษัท โปรวิชั่น จำกัด, 2546.
- [3] บุญเสริม กิจศิริกุล. “รายงานวิจัยเรื่องอัลกอริทึมการทำเหมืองข้อมูล.” คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2545.
- [4] U. Fayyad, G. P.-Shapiro, and P. Smyth. “**Knowledge Discovery and Data Mining: Towards a Unifying Framework.**” Proc. Of the Second Intl. Conf. On Knowledge Discovery and Data Mining (KDD-96), Aug. 1996.
- [5] R. Agrawal and R. Srikant. “**Fast algorithms for mining association rules in large databases.**” Research Report RJ 9839, IBM Almaden Research Center, San Jose, California, June 1994.
- [6] L. Nagle P., Iba W. and Thompson, K. **An analysis of bayesian classifiers.** In Proceedings of the tenth national conference on artificial intelligence. AAAI Press and MIT Press 1992 : 223-228.
- [7] ศิริพร เวชยโสภณ, ศิริรักษ์ ดันติกุลไพบูลย์, วิภาพร สันทอง. “ระบบการลงทะเบียนเรียนของนักศึกษาผ่าน Web.” คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, 2543.
- [8] วริน เจริญรัตน์. “การจัดเก็บข้อมูลของผู้ป่วยและเสียชีวิตในโรงพยาบาล.” คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2540.