

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของงาน

จากข้อมูลของสำนักนโยบายและยุทธศาสตร์ กระทรวงสาธารณสุข (รายงานผู้ป่วย ในช่วงปี พ.ศ. 2528-2552, 2554) พบว่าอัตราป่วยต่อประชากรแสนคนด้วยโรคหัวใจ โรคมะเร็ง โรคเบาหวาน โรคความดันโลหิตสูงและหลอดเลือดสมองมีแนวโน้มที่สูงขึ้นอย่างต่อเนื่อง โดยมี ปริมาณที่เพิ่มขึ้นจากปี พ.ศ. 2548 จำนวน 4.7 ล้านคนเป็น 5.4 ล้านคนในปี พ.ศ. 2552 ซึ่งการ เจ็บป่วยดังกล่าวเป็นโรคเรื้อรังที่สามารถป้องกันได้ในเบื้องต้นโดยประชาชนจะต้องดูแลสุขภาพ ตนเองด้วยการมีพฤติกรรมการบริโภคที่ถูกต้อง การควบคุมน้ำหนัก การออกกำลังกายอย่าง สม่าเสมอ ควรหลีกเลี่ยงการดื่มเครื่องดื่มที่มีแอลกอฮอล์ การพักผ่อนและการควบคุมความเครียด การผ่อนคลายในชีวิตประจำวัน รวมทั้งการตรวจสุขภาพอย่างสม่ำเสมอด้วยตนเองและสถาน บริการสุขภาพ นอกจากนี้ประชาชนยังต้องมีส่วนร่วมในการกำหนดมาตรการทางสังคมในการอยู่ ร่วมกัน เช่น เขตปลอดบุหรี่ การจำหน่ายอาหารปลอดภัย การณรงค์ออกกำลังกาย การลดปัจจัย สาเหตุอย่างครอบคลุมทั้งเรื่องของพฤติกรรม การบริโภค ปัจจัยเสี่ยงจากสิ่งแวดล้อม การจราจร และวิถีการดำเนินชีวิตประจำวัน อันจะนำไปสู่การมีสังคมสุขภาพที่ดีร่วมกัน

กระทรวงสาธารณสุขจึงได้กำหนดนโยบายให้สำนักงานสาธารณสุขจังหวัดทุกจังหวัด เป็นแกนนำในการดำเนินกิจกรรมเพื่อให้ประชาชนตระหนักถึงความสำคัญของการมีพฤติกรรม สุขภาพที่เหมาะสม ซึ่งมีความจำเป็นต้องดำเนินการอย่างเร่งด่วนและต่อเนื่อง เพราะปัญหาที่เกิดขึ้น ไม่เพียงแต่ส่งผลกระทบต่อสุขภาพเท่านั้น ยังมีผลต่อเนื่องถึงภาวะเศรษฐกิจและการพัฒนาประเทศ ด้วย

สำนักงานหลักประกันสุขภาพแห่งชาติจึงได้ส่งเสริมสนับสนุนงบประมาณในการเฝ้า ระวังโรคเรื้อรังทุกจังหวัด โดยการใช้แบบสำรวจสภาวะสุขภาพของประชาชนอายุตั้งแต่ 15 ปี ขึ้น ไป ซึ่งใช้ในการคัดกรองสุขภาพเบื้องต้นเป็นประจำทุกปี ทั้งในเขตเทศบาลและนอกเขตเทศบาล ศูนย์แพทย์ชุมชน โรงพยาบาลบุรีรัมย์ ได้เล็งเห็นความสำคัญของนโยบายดังกล่าว จึงได้จัดทำ การสำรวจสภาวะสุขภาพของประชาชนอายุตั้งแต่ 15 ปีในเขตเทศบาล ซึ่งข้อมูลประกอบด้วย 1) ข้อมูล ทั่วไป 2) ข้อมูลครอบครัว 3) ข้อมูลการเจ็บป่วยและพฤติกรรมเมื่อเจ็บป่วย 4) ข้อมูลคะแนน

พฤติกรรมด้านสุขภาพ เพื่อนำข้อมูลที่ได้มาออกแบบกิจกรรมและงบประมาณให้สอดคล้องกับผลที่สำรวจได้ทุกปี ดังเช่น ผลการคัดกรองสุขภาพเบื้องต้นในชุมชนหนึ่งพบว่ามีคนที่อยู่ในกลุ่มเสี่ยงเป็นจำนวนมาก และเมื่อนำมาวิเคราะห์ทำให้ทราบว่าส่วนใหญ่เป็นคนอ้วนลงพุง จึงจัดให้มีการอบรมเพื่อปรับเปลี่ยนพฤติกรรมในกลุ่มเสี่ยงของชุมชนนั้นด้วย โครงการชุมชนลดหวาน มัน เค็ม ลดอ้วน ลดโรคใน ปี 2555 เป็นต้น

เทคนิคการทำเหมืองข้อมูล (Data Mining) เป็นเทคนิคหนึ่งที่ได้รับการนิยมนำมาใช้วิเคราะห์ข้อมูลที่มีปริมาณมาก (Han, J., et al., 2006) โดยสามารถนำมาใช้หากฎเกณฑ์ความสัมพันธ์ และรูปแบบที่แฝงอยู่ในฐานข้อมูล ได้เป็นความรู้หรือสารสนเทศที่น่าสนใจ โดยจะต้องเป็นความรู้ที่ไม่ธรรมดา ไม่สามารถค้นพบได้ด้วยวิธีการเรียกดูข้อมูลจากตาราง (Query) ต้องไม่เคยรู้มาก่อน และต้องสามารถนำไปใช้ประโยชน์ได้ เช่น การค้นพบความสัมพันธ์ของสินค้าที่ถูกซื้อจากร้านสะดวกซื้อแห่งหนึ่งในประเทศสหรัฐอเมริกาพบว่าลูกค้ามักซื้อผ้าอ้อมกับเบียร์ไปด้วยกัน (Han, J., et al., 2006) ทำให้สามารถนำความรู้ที่ได้มาประยุกต์ในการวางกลยุทธ์เพื่อเพิ่มยอดขายหรือใช้ในการจัดชั้นสินค้าได้

ข้อมูลที่สำรวจได้จากแบบสำรวจสภาวะสุขภาพของประชาชนจะมีปริมาณเพิ่มสูงขึ้นทุกปี และยังขาดการจัดการข้อมูลเพื่อนำไปใช้ประโยชน์อย่างคุ้มค่า งานวิจัยนี้จึงนำเสนอการวิเคราะห์ข้อมูลจากแบบสำรวจสภาวะสุขภาพประชาชนด้วยเทคนิคการทำเหมืองข้อมูลโดยใช้เทคนิค Association Rule และ Classification เพื่อนำไปใช้ในการหาความสัมพันธ์ในการเกิดโรคต่างๆ และใช้ในการค้นหากฎในการจำแนกสุขภาพเบื้องต้นของประชาชนว่าอยู่ในกลุ่มปกติ กลุ่มเสี่ยงหรือกลุ่มป่วย และนำกฎที่ได้ไปสร้างระบบคัดกรองสุขภาพเบื้องต้นที่เปิดให้ประชาชนทั่วไปที่สนใจสามารถเข้ามาตรวจสอบสุขภาพได้ด้วยตนเอง ซึ่งช่วยลดภาระงานของผู้เชี่ยวชาญที่ต้องใช้เวลาวิเคราะห์ข้อมูลที่ละเอียดรายให้สามารถทำงานได้สะดวกและรวดเร็วมากยิ่งขึ้น

1.2 วัตถุประสงค์ของการศึกษา

1. เพื่อศึกษาการเตรียมข้อมูลเพื่อใช้ในการทำเหมืองข้อมูล
2. เพื่อศึกษาเทคนิคการทำเหมืองข้อมูลแบบ Association Rule และ Classification
3. เพื่อศึกษาลักษณะความสัมพันธ์ของการเกิดโรคต่างๆ
4. เพื่อศึกษารูปแบบการคัดกรองผู้ป่วยจากข้อมูลแบบสำรวจสภาวะสุขภาพประชาชนแบบ

อัตโนมัติ

1.3 ขอบเขตของงานวิจัย

ข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลของประชาชนในเขตเทศบาลเมืองจังหวัดบุรีรัมย์ ปี พ.ศ. 2555 จำนวนทั้งหมด 1,071 เรคอร์ด โดยเทคนิคที่ศึกษาประกอบด้วย

1. Association Rule ใช้ในการหาความสัมพันธ์ของการเกิดโรคต่างๆ ได้แก่ โรคเบาหวาน โรคความดันโลหิตสูง โรคตับ โรคอัมพาต โรคหัวใจ และโรคไขมันในเลือดผิดปกติ

2. Classification ใช้ในการหารูปแบบการพยากรณ์การคัดกรองสุขภาพเบื้องต้น โดยแบ่งออกเป็น กลุ่มปกติ กลุ่มเสี่ยง และกลุ่มป่วย

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถกฎการจำแนกที่ได้ไปสร้างเป็นระบบคัดกรองสุขภาพเบื้องต้น เพื่อช่วยอำนวยความสะดวกในการทำงานของเจ้าหน้าที่ศูนย์แพทย์ชุมชน และสามารถให้ประชาชนทั่วไปที่สนใจสามารถเข้ามาสำรวจสุขภาพได้ด้วยตนเอง

2. สามารถนำความรู้ที่ได้ไปประยุกต์ใช้ในการหาทางป้องกันและลดจำนวนผู้ป่วยที่มีความเสี่ยงในการเกิดโรคที่ศึกษา

1.5 นิยามศัพท์เฉพาะ

1. Data Mining เป็นกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น (Han, J., et al., 2006)

2. Association Rule เป็นการค้นหากฎความสัมพันธ์ของข้อมูล โดยค้นหาความสัมพันธ์ของข้อมูลทั้งสองชุดหรือมากกว่าสองชุดขึ้นไป (Han, J., et al., 2006)

3. Classification เป็นการจัดแบ่งประเภทของข้อมูล โดยหาชุดต้นแบบหรือชุดของการทำงานที่อธิบายและแบ่งประเภทข้อมูล วัตถุประสงค์เพื่อให้สามารถใช้เป็นต้นแบบทำนายประเภทของวัตถุหรือข้อมูลที่ไม่มีการระบุประเภทหรือชนิดของข้อมูล ซึ่งต้นแบบสร้างจากการวิเคราะห์ชุดของข้อมูลฝึกสอน (Training Data) โดยอาจจะเป็นกลุ่มข้อมูลที่มีการระบุประเภทหรือกลุ่มเรียบร้อยแล้ว รูปแบบของต้นแบบแสดงได้หลายแบบเช่น Classification Rules, Decision Trees หรือ Neural Networks เป็นต้น (Han, J., et al., 2006)

1.6 เครื่องมือ (Tool) ที่ใช้ในการดำเนินการ

1. โปรแกรมสำเร็จรูป Weka เพื่อใช้ในการทำเหมืองข้อมูล
2. ระบบปฏิบัติการ: Windows

3. โปรแกรม Microsoft Excel ใช้ในการทำความสะอาดข้อมูล ลดขนาดของข้อมูลและเตรียมข้อมูลสำหรับนำเข้าทดสอบในโปรแกรม Weka

4. โปรแกรม WordPad ใช้ในแก้ไขเพิ่มข้อมูล .arff

5. โปรแกรม Adobe Dreamweaver CS5.5, PHP, Apache และฐานข้อมูล MySQL สำหรับพัฒนาระบบคัดกรองสุขภาพเบื้องต้น

1.7 ระยะเวลาในการดำเนินการวิจัย

ระยะเวลาในการดำเนินการวิจัย และแผนผังขั้นตอนในการดำเนินการวิจัย สามารถสรุปได้ตารางที่ 1.1

ตารางที่ 1.1 ระยะเวลาในการดำเนินการวิจัย

ลำดับ	ระยะเวลาดำเนินงาน (เดือน)	1	2	3	4	5	6	7	8	9	10
1	ศึกษาและรวบรวมข้อมูล	←→									
2	เตรียมข้อมูล		←→								
3	ศึกษา ทดสอบและเลือกเทคนิคที่เหมาะสม			←→							
4	วิเคราะห์และสรุปผล				←→						
5	พัฒนาระบบ					←→					
6	สรุปผลการวิจัยและข้อเสนอแนะ						←→				
7	เรียบเรียงรายงานค้นคว้าอิสระ						←→				