

ระบบช่วยพยากรณ์พื้นที่เสี่ยงต่อโรคเบาหวาน และความดันโลหิตสูงในพื้นที่ จังหวัดพิษณุโลก

ใจรัก นุ่มพิษณ¹ และ เทวิน ณะวงษ์¹

¹ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร พิษณุโลก
Emails: Jairakn56@email.nu.ac.th, tawint@nu.ac.th

บทคัดย่อ

ในงานวิจัยนี้ได้นำเสนอการใช้เทคนิคดาต้าไมนิ่งและระบบสารสนเทศภูมิศาสตร์ในการพัฒนาระบบพยากรณ์พื้นที่เสี่ยงต่อโรคเบาหวาน และความดันโลหิตสูงในพื้นที่จังหวัดพิษณุโลก ซึ่งพัฒนาด้วยโปรแกรมเวก้า 3.6.13 โดยใช้เทคนิคต้นไม้ช่วยตัดสินใจภายใต้การเปรียบเทียบอัลกอริธึมระหว่าง J48 กับ J48graft เพื่อสร้างแบบจำลองที่เหมาะสม โดยใช้รูปแบบการทดสอบข้อมูลด้วยการแบ่งข้อมูลด้วยวิธีการตรวจสอบไขว้แบบ 5Fold Cross validation และ 10Fold Cross Validation และวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งแบบ Percentage Split 10% ,20% และ 66% ตามลำดับ ผลลัพธ์ที่ได้จากการสร้างแบบจำลองพยากรณ์ จะพบว่า แบบพยากรณ์พื้นที่เสี่ยงการเกิดโรคเบาหวานที่ถูกพัฒนาโดยอัลกอริธึม J48 และ J48graft ด้วยวิธี Percentage split20 มีค่าความถูกต้องมากที่สุดเท่ากับ 100% และด้วยวิธี Cross-Validation -5Fold และ-10Fold มีค่ามีค่าความถูกต้องมากที่สุดเท่ากับ 100% เช่นกัน และแบบจำลองพยากรณ์พื้นที่เสี่ยงการเกิดโรคความดันโลหิตสูงที่ถูกพัฒนาโดยอัลกอริธึมทั้ง J และ J48graft ด้วยวิธี Percentage split 66% มีค่าความถูกต้องมากที่สุดเท่ากับ 90% และด้วยวิธี Cross validation 10-Fold มีค่ามีค่าความถูกต้องมากที่สุดเท่ากับ 100%

ABSTRACT

This research proposes using techniques data mining modeling and geographic information system development forecasting system in risk areas of diabetes and hypertension at Phitsanulok province area. The development of model with weka3.6.13 program, which consisted of algorithm J48 and J48graft. The data were divided by the random data method which was 5-fold and 10-fold cross validation, and the percentage split 10%, 20% and 66%

respectively. The research result shown that, the algorithm j48 and j48graft by percentage split20% and 66%, the correctly level of highest diabetes classified data were 100 and 90 percent, and the 5-fold, 10-fold and 5-fold cross validation of diabetes and hypertension classified data were 100 percent equal.

คำสำคัญ— โรคเบาหวาน; โรคความดันโลหิตสูง; เหมืองข้อมูล; สารสนเทศภูมิศาสตร์

1. บทนำ

โรคไม่ติดต่อเรื้อรังเป็นสาเหตุหลักของการเสียชีวิตของคนไทย ได้แก่ โรคเบาหวานความดันโลหิตสูง หลอดเลือดสมอง หัวใจ และมะเร็งในปี 2552 คิดเป็นร้อยละ 73 ของคนไทยที่เสียชีวิต จากกลุ่มโรคไม่ติดต่อเรื้อรังทั้งสิ้น 314,340 คน สำหรับค่าใช้จ่ายในการรักษาพยาบาลเฉลี่ยของผู้ป่วยโรคไม่ติดต่อเรื้อรัง จากข้อมูลหลักประกันสุขภาพถ้วนหน้า ประกันสังคม และสวัสดิการ รักษาพยาบาลข้าราชการและครอบครัว พบว่าภาระทางเศรษฐกิจสำหรับประเทศไทยในปี 2551 มีค่าใช้จ่ายในการรักษาพยาบาลผู้ป่วยนอกและผู้ป่วยในด้วยโรคไม่ติดต่อเรื้อรังทั้งสิ้นเฉลี่ย 25,225 ล้านบาทต่อปี โดยเฉพาะโรคเบาหวาน 3.984 ล้านบาทต่อปี คิดเป็นร้อยละ 15.79 ของภาพรวมโรคไม่ติดต่อเรื้อรังหลัก ซึ่งสถานบริการจะต้องเสียค่ารักษาพยาบาลประมาณ 335,358 ล้านบาทต่อปี คิดเป็นร้อยละ 2.94 ของผลิตภัณฑ์มวลรวมภายในประเทศและผู้ป่วยโรคเบาหวานมีโอกาสเสี่ยงต่อการเกิดโรคแทรกซ้อนเรื้อรังที่ร้ายแรงต่างๆ ส่งผลต่อการดำรงชีวิต [5,10] การทำงานของผู้ป่วยและครอบครัวทำให้เสียค่าใช้จ่ายในการรักษาสูง เมื่อผู้ป่วยมีจำนวนมากขึ้นจะส่งผลกระทบต่อสถานะเศรษฐกิจของประชาชาติ ในปัจจุบันมีผู้ป่วยเป็นโรคเบาหวาน และความดันโลหิตสูง ในพื้นที่จังหวัดพิษณุโลกมีจำนวนมาก จากสถิติสำนักงานสาธารณสุขจังหวัดพิษณุโลก ปีงบประมาณ 2558 [18] มีจำนวนผู้ป่วยโรคเบาหวาน

และโรคความดันโลหิตสูงประมาณ 142,780 คน จากประชากร 851,357 คิดเป็นร้อยละ 16.77 และคาดว่าจะมีจำนวนมากขึ้นและเทคโนโลยีที่นำเสนอรายงานจำนวนผู้ป่วยโรคเบาหวาน และโรคความดันโลหิตสูง เป็นเพียงการนำเสนอในรูปแบบสถิติ ยังไม่สามารถพยากรณ์จำนวนผู้ป่วยได้

ดังนั้นผู้วิจัยจึงเล็งเห็นความสำคัญของปัญหาการเกิดโรคเบาหวาน และโรคความดันโลหิตสูง จึงนำระบบสารสนเทศภูมิศาสตร์ และเทคนิคดาต้าไมนิ่ง [2,14] มาประยุกต์ในการสร้างระบบที่ช่วยพยากรณ์การเกิดโรคเบาหวาน และโรคความดันโลหิตสูง เชิงพื้นที่ในจังหวัดพิษณุโลกเพื่อวิเคราะห์และกำหนดพื้นที่ที่มีความเสี่ยงของการเกิดโรคเบาหวาน และโรคความดันโลหิตสูง โดยนำเว็บแอปพลิเคชันมาใช้ในการสร้างตัวแบบพยากรณ์เพื่อช่วยให้เจ้าหน้าที่ และผู้ที่เกี่ยวข้องทางด้านสาธารณสุขทราบถึงแนวโน้มการเกิดโรคดังกล่าว เพื่อใช้ในการวางแผนนโยบายทางเชิงรุกทางด้านสาธารณสุขต่อไป

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 โรคเบาหวาน [15]

เกิดจากความผิดปกติของการทำงานของฮอร์โมนอินซูลิน (Insulin) ส่งผลให้อินซูลินซึ่งมีหน้าที่นำน้ำตาลภายในเลือด เข้าสู่เซลล์ต่างๆ ทำงานได้ไม่เต็มประสิทธิภาพและเกิดการคั่งของน้ำตาลในเส้นเลือดแดงส่งผลให้อวัยวะต่างๆ เสื่อม ซึ่งอาการนี้จะส่งผลให้เกิดโรคและอาการแทรกซ้อนต่ออวัยวะต่างๆ ได้

2.2 โรคความดันโลหิตสูง

ความดันโลหิตจะเป็นแรงผลักดัน ให้เลือดไปเลี้ยงอวัยวะต่างๆ ของร่างกาย ดังนั้นทุกคนควรจะเรียนรู้เกี่ยวกับความดันโลหิต และรักษาให้ความดันโลหิตอยู่ในเกณฑ์ปกติ เพราะความดันโลหิตสูงจะทำให้เกิดหลอดเลือดแข็งและตีบ เมื่อหัวใจบีบตัวหัวใจจะบีบเลือดไปยังหลอดเลือดแดง ทำให้เกิดความดันโลหิตซึ่งเกิดจากการบีบตัวของหัวใจ และแรงต้านทานของหลอดเลือด หัวใจคนเราเต้น 60-80 ครั้ง ความดันก็จะเพิ่มขึ้นขณะที่หัวใจบีบตัว และลดลงขณะที่หัวใจคลายตัว ความดันโลหิตของคนเราไม่เท่ากันตลอดเวลาขึ้นกับท่า ความเครียด การออกกำลังกาย การนอนหลับ ค่าปกติของคนเราคือ 120/80 มิลลิเมตรปรอท แต่ไม่ควรเกิน 140/90 หากสูงกว่านี้แสดงว่าคุณเป็นโรคความดันโลหิตสูง โรคความดันโลหิตสูงเป็นปัจจัยเสี่ยงทำให้เกิดโรคหัวใจ โรคไต โรคหลอดเลือดหัวใจตีบ โรคอัมพาต โรคหัวใจเป็นโรคที่มีอัตราตายสูง ดังนั้นการป้องกันความดันโลหิตสูงสามารถป้องกันอัตราการตายจากโรคหัวใจ และโรคอัมพาต โรคความดันโลหิตสูงเป็นภัยเงียบที่คุกคามชีวิตของทุกท่านเนื่องจากไม่มีอาการเตือน

ดังนั้น การจะทราบว่าเป็นความดันโลหิตสูงจำเป็นต้องวัดความดันโลหิต

2.3 ระบบสารสนเทศภูมิศาสตร์ (GIS) [3,12]

คือกระบวนการทำงานเกี่ยวกับข้อมูลในเชิงพื้นที่ด้วยระบบคอมพิวเตอร์ที่ใช้กำหนดข้อมูลและสารสนเทศที่มีความสัมพันธ์ในเชิงพื้นที่เช่น ที่อยู่ บ้านเลขที่สัมพันธ์กับตำแหน่งในแผนที่ ตำแหน่ง เส้นรุ้ง เส้นแวงข้อมูลและแผนที่ใน GIS เป็นระบบข้อมูลสารสนเทศที่อยู่ในรูปของตารางข้อมูลและฐานข้อมูลที่มีส่วนสัมพันธ์กับข้อมูลเชิงพื้นที่ (Spatial Data) จะสามารถนำมาวิเคราะห์ด้วย GIS และทำให้สื่อความหมายในเรื่องการเปลี่ยนแปลงที่สัมพันธ์กับเวลาได้เช่น การแพร่ขยายของโรคระบาด การเคลื่อนย้ายถิ่นฐาน การบุกรุกทำลาย การเปลี่ยนแปลงของการใช้พื้นที่ ฯลฯ ข้อมูลเหล่านี้ เมื่อปรากฏบนแผนที่ทำให้สามารถแปลและสื่อความหมาย ใช้งานได้ง่าย

2.4 โปรแกรม Quantum GIS [9]

Quantum GIS หรือ QGIS เป็นโปรแกรม Desktop GIS ประเภทหนึ่งที่มีประสิทธิภาพในการนำมาใช้จัดการข้อมูลปริภูมิ จัดอยู่ในกลุ่มซอฟต์แวร์รหัสเปิดที่ใช้งานง่าย ลักษณะการใช้งานเป็นแบบ Graphic User Interface ซึ่งสะดวกต่อการใช้งาน ไม่ว่าจะเป็นการเรียกใช้ข้อมูลภาพ ข้อมูลตาราง การแสดงผลตาราง การแสดงผลกราฟ ตลอดจนสามารถสืบค้นข้อมูลวิเคราะห์ข้อมูล และนำเสนอข้อมูลได้ในรูปแบบแผนที่ สามารถเรียกใช้ข้อมูลเวกเตอร์แรสเตอร์ในรูปแบบที่เป็นมาตรฐานแพร่หลายเช่น Shapefile และ GeoTIFF QGIS สามารถแก้ไข Shape File format ได้ซึ่งเป็นที่ต้องการมากในเวลานี้ 2.5 ArcGIS ArcGIS Online ของค่าย ESRI มีฟังก์ชันสนับสนุนการใช้งานแผนที่ออนไลน์อย่างครบถ้วน โดยหลัก ๆ แล้วไม่ต่างจาก Google Map แต่ในรายละเอียดก็มีความแตกต่างที่หลากหลาย การออกแบบ UI ดูสวย ทันสมัย แต่เรียบง่าย ไม่ซับซ้อน สร้างความคุ้นเคยได้เร็ว มีเครื่องมือช่วยเหลือที่ละเอียด รวมถึงคลิปวิดีโอแนะนำการใช้งานเกือบทุกขั้นตอน ArcGIS Online แบ่งเป็น 2 ส่วน คือ Personal Use และ Organizations ซึ่งจะมีความแตกต่างกันในด้านข้อจำกัดและสิทธิประโยชน์แน่นอนว่าแบบ Organizations ใช้งานได้กว้างขวางกว่า มีสิ่งอำนวยความสะดวกมากกว่า ผู้ใช้สามารถที่จะสร้าง Web Maps ให้แผนที่ในรูปแบบต่างๆ รวมถึง ใส่ชั้นข้อมูลเพื่อแสดงบนแผนที่ มีฟังก์ชันการค้นหาสถานที่ การใช้สัญลักษณ์ และอื่น ๆ อีกหลายอย่าง โดยใช้เครื่องมือที่อยู่แล้วบนเว็บไซต์ ArcGIS.com สามารถที่จะนำข้อมูลขึ้นแผนที่ได้หลากหลายรูปแบบ เช่น spreadsheets, shape files, KML

เป็นต้น ซึ่งผู้ใช้ไม่จำเป็นต้องติดตั้ง Servers หรือ ซอฟต์แวร์ใด ๆ ก็สามารถที่จะเผยแพร่ Web Maps ที่สร้างขึ้นผ่านเบราว์เซอร์ได้ โดยผู้อื่นสามารถเข้ามาดูได้ทั้งจาก Desktop และ mobile device จัดเก็บข้อมูลบน Cloud ข้อมูลต่าง ๆ ที่ผู้ใช้จัดทำขึ้น เพื่อแสดงบนแผนที่ ถูกจัดเก็บบน Cloud ของ ESRI จึงไม่มีความจำเป็นจะต้องตั้ง Servers ใด ๆ แบ่งปัน แผนที่ ข้อมูล และ แอปพลิเคชัน

2.5 Responsive Web [13, 16]

คือเว็บไซต์ที่สามารถรองรับการทำงานบนหน้าจออุปกรณ์[11] เครือข่ายได้ อุปกรณ์เช่น Desktop Internet, Mobile Internet ซึ่งอุปกรณ์เหล่านี้จะมีหน้าจอแตกต่างกันไป ตามขนาดความกว้างของเครื่อง ทำให้หน้าตาเว็บไซต์ที่ออกแบบให้ดูผ่านหน้าจอคอมพิวเตอร์อย่างเดียว มีปัญหาการทำงานเมื่อใช้งานผ่านอุปกรณ์ที่มีหน้าจอขนาดเล็กกว่า เช่น Mobile Internet Users อุปกรณ์มือถือ เพราะเนื่องจากปัจจุบันนี้ การใช้งานผ่าน Mobile Internet มีการเจริญเติบโตสูง และมีแนวโน้มที่จะใช้มากขึ้นกว่า Desktop Internet Users [17] ทำให้บุคคลทั่วไปสามารถเข้าชมเว็บไซต์ได้ตลอดเวลา

2.6 งานวิจัยที่เกี่ยวข้อง

การประยุกต์ใช้เทคนิคเหมืองข้อมูลพยากรณ์โรคเบาหวาน [1] การใช้สารสนเทศภูมิศาสตร์เพื่อเฝ้าระวังโรคระบาด[4] การวิเคราะห์พื้นที่เสี่ยงต่อโรคมะเร็งในจังหวัดพิษณุโลกโดยใช้ระบบสารสนเทศภูมิศาสตร์ [6] โดยนำระบบสารสนเทศภูมิศาสตร์มาใช้เป็นเครื่องมือวิเคราะห์พื้นที่เสี่ยงโรคมะเร็ง โดยการเก็บสถิติผู้ป่วยโรคมะเร็ง ย้อนหลัง 3 ปี (2548-2546) ยิงลาข้อมูลสำรวจจำนวนลูกน้ำ และสิ่งแวดล้อมมาวิเคราะห์แปลผล และนำเสนอ เพื่อมาพยากรณ์แนวโน้มการเกิดโรคมะเร็ง ส่งผลต่อการวางแผนนโยบายการป้องกันควบคุมโรคมะเร็ง การพยากรณ์ความต้องการน้ำประปาในพื้นที่เทศบาลนครพิษณุโลกโดยใช้แบบจำลองคณิตศาสตร์และระบบสารสนเทศภูมิศาสตร์[7] เพื่อให้ทราบถึงปริมาณความต้องการน้ำประปาของประชาชนในภาพรวมทั้งเขตเทศบาล และในระดับเขตการให้บริการบริการน้ำประปาย่อยที่จะเกิดในอนาคต และทราบถึงปัจจัยที่มีอิทธิพลต่อความต้องการน้ำประปาในเขตเทศบาลนครพิษณุโลก สามารถใช้เป็นแนวทางการแก้ปัญหากับเขตการให้บริการน้ำประปาที่พบปริมาณเกินเกณฑ์มาตรฐานที่กำหนดได้

3. วิธีการดำเนินการศึกษาวิจัย

3.2 ศึกษาและเก็บรวบรวมข้อมูล

การศึกษาและรวบรวมข้อมูลเพื่อใช้ในการดำเนินงานนั้น ผู้วิจัยได้ทำการศึกษาและรวบรวมข้อมูลที่เกี่ยวข้องและเหมาะสมกับระบบ โดยทำการเก็บรวบรวมข้อมูลแบบทุติยภูมิ โดยขอข้อมูลจาก สำนักงานสถิติจังหวัดพิษณุโลก และสำนักงานสาธารณสุขจังหวัดพิษณุโลก ข้อมูลจำนวนประชากร จำนวนรายได้ต่อปี จำนวนโรงพยาบาลส่งเสริมสุขภาพชุมชน จำนวนผู้ป่วยโรคเบาหวานและความดันโลหิตสูง 5 ปีย้อนหลัง ตั้งแต่ปี 54-58 ของแต่ละอำเภอในจังหวัดพิษณุโลก

3.3 การรวบรวม คัดเลือกและจัดเตรียมข้อมูล

3.3.1 การรวบรวมข้อมูล

ข้อมูลที่เก็บรวบรวมได้มาจากหน่วยงานของรัฐ และนำมาใช้ในการสร้างแบบจำลองและทดสอบแบบจำลอง โดยใช้กฎการจำแนกเทคนิคต้นไม้ตัดสินใจ ประกอบด้วยจำนวนตัวอย่าง 45 ชุดข้อมูล แบ่งเป็นตารางข้อมูลสำหรับแบบจำลองโรคเบาหวาน และข้อมูลแบบจำลองโรคความดันโลหิตสูงในรูปแบบไฟล์นามสกุล arff ดังรูปที่1 และ รูปที่2 ตามลำดับ

Relation diabetes	
@attribute District (BangFratthum,BangKakan,ChatTrakan,MueangPhitsanulok,NakhonThai,HonNagrang,PhromPhisan,WangThong,WatBot)	
@attribute HospitalHealth numeric	
@attribute AllPopulation numeric	
@attribute Male numeric	
@attribute Female numeric	
@attribute VulnerableGroups numeric	
@attribute Older numeric	
@attribute IncomeYears numeric	
@attribute HyperFetien numeric	
@attribute DiabetesFetien numeric	
@attribute Summary (medium,high,low)	
Data	
BangFratthum,13,48313,23571,24742,18816,6719,71733,4098,2147,medium	
BangFratthum,13,48390,23611,24779,18796,7615,76343,4124,2456,medium	
BangFratthum,13,48307,23612,24695,18570,8046,78418,2392,6816,high	
BangFratthum,13,48152,23494,24658,18570,8046,78418,2392,6816,high	
BangFratthum,13,48025,23439,24586,18531,8617,79559,2905,6990,high	
BangFratthum,13,48025,23439,24586,18531,8617,79559,2905,6990,high	
BangFratthum,20,94020,46235,47787,24820,10641,74579,4672,3548,low	
BangFratthum,20,94578,46489,48089,36376,12580,69078,6853,3765,low	
BangFratthum,20,94832,46429,48203,35932,13847,76149,3978,6574,low	
BangFratthum,20,94850,46484,48236,35932,13847,76149,3978,6574,low	
BangFratthum,20,95146,46699,48447,36372,14992,92100,11349,4626,medium	
ChatTrakan,10,40144,20335,19811,13016,3150,42763,2679,1264,low	
ChatTrakan,10,40482,20487,19978,13090,4228,51876,2679,1264,low	
ChatTrakan,10,40633,20529,20110,15275,4484,51652,3208,1487,low	
ChatTrakan,10,40801,20591,20210,15275,4484,51652,3208,1487,low	
ChatTrakan,10,41100,20749,20383,15275,4484,51652,3208,1487,low	
MueangPhitsanulok,24,280922,135421,145501,99309,29256,71052,17194,6469,low	
MueangPhitsanulok,24,280922,135421,145501,99309,29256,71052,17194,6469,low	
MueangPhitsanulok,24,281762,135667,146095,103609,40700,96198,23850,10758,low	
MueangPhitsanulok,24,283419,136184,147235,103609,40700,96198,23850,10758,low	
NakhonThai,20,85534,42886,42648,23035,8235,49354,5691,2118,low	
NakhonThai,20,86143,43161,43002,33224,11257,52023,4738,2476,low	
NakhonThai,20,86489,43464,43220,33415,13554,48540,3323,7259,low	
NakhonThai,20,87042,43603,43439,33415,12033,57589,3777,8271,medium	
NakhonThai,20,87352,43716,43636,34006,13212,73747,4374,9826,low	

รูปที่ 1. แสดงตัวอย่างข้อมูลสำหรับแบบจำลองโรคเบาหวาน

Relation Hypertension	
@attribute District (BangFratthum,BangKakan,ChatTrakan,MueangPhitsanulok,NakhonThai,HonNagrang,PhromPhisan,WangThong,WatBot)	
@attribute HospitalHealth numeric	
@attribute AllPopulation numeric	
@attribute Male numeric	
@attribute Female numeric	
@attribute VulnerableGroups numeric	
@attribute Older numeric	
@attribute IncomeYears numeric	
@attribute HyperFetien numeric	
@attribute Summary (medium,high,low)	
Data	
BangFratthum,13,48313,23571,24742,18816,6719,71733,4098,2147,4098,medium	
BangFratthum,13,48390,23611,24779,18796,7615,76343,4124,2456,4124,medium	
BangFratthum,13,48307,23612,24695,18570,8046,78418,2392,6816,4124,high	
BangFratthum,13,48152,23494,24658,18570,8046,78418,2392,6816,4124,high	
BangFratthum,13,48025,23439,24586,18531,8617,79559,2905,6990,4124,high	
BangFratthum,13,48025,23439,24586,18531,8617,79559,2905,6990,4124,high	
BangFratthum,20,94020,46235,47787,24820,10641,74579,4672,3548,4672,low	
BangFratthum,20,94578,46489,48089,36376,12580,69078,6853,3765,4672,low	
BangFratthum,20,94832,46429,48203,35932,13847,76149,3978,6574,4672,low	
BangFratthum,20,94850,46484,48236,35932,13847,76149,3978,6574,4672,low	
BangFratthum,20,95146,46699,48447,36372,14992,92100,11349,4626,4626,high	
ChatTrakan,10,40144,20335,19811,13016,3150,42763,2679,1264,2679,low	
ChatTrakan,10,40482,20487,19978,13090,4228,51876,2679,1264,2679,low	
ChatTrakan,10,40633,20529,20110,15275,4484,51652,3208,1487,3208,low	
ChatTrakan,10,40801,20591,20210,15275,4484,51652,3208,1487,3208,low	
ChatTrakan,10,41100,20749,20383,15275,4484,51652,3208,1487,3208,low	
MueangPhitsanulok,24,280922,135421,145501,99309,29256,71052,17194,6469,17194,low	
MueangPhitsanulok,24,280922,135421,145501,99309,29256,71052,17194,6469,17194,low	
MueangPhitsanulok,24,281762,135667,146095,103609,40700,96198,23850,10758,10758,low	
MueangPhitsanulok,24,283419,136184,147235,103609,40700,96198,23850,10758,10758,low	
NakhonThai,20,85534,42886,42648,23035,8235,49354,5691,2118,2118,low	
NakhonThai,20,86143,43161,43002,33224,11257,52023,4738,2476,4738,low	
NakhonThai,20,86489,43464,43220,33415,13554,48540,3323,7259,7259,low	
NakhonThai,20,87042,43603,43439,33415,12033,57589,3777,8271,8271,medium	
NakhonThai,20,87352,43716,43636,34006,13212,73747,4374,9826,9826,low	
HonNagrang,11,87716,23043,23875,20816,4209,54419,1494,3340,low	

รูปที่ 2. แสดงตัวอย่างข้อมูลสำหรับแบบจำลองโรคความดันโลหิตสูง

3.3.2 การคัดเลือกปัจจัยเพื่อการประมวลผล

จากข้อมูลที่ได้รับรวบรวมมาจำนวน 45 ชุดข้อมูล สามารถนำมาคัดเลือกปัจจัยหรือคุณลักษณะหรือแอททริบิวต์ (Attribute) ที่เกี่ยวข้องเพื่อนำมาใช้สำหรับสร้างแบบจำลองสำหรับพยากรณ์ของโรคเบาหวาน และความดันโลหิตสูงจำนวน 11 และ 10 คุณลักษณะ ตามลำดับ ดังตาราง 1 และ ตาราง 2 โดยแต่ละคุณลักษณะมีรายละเอียดนำเสนอเพียงบางส่วนดังนี้

ตาราง 1. แสดงคุณลักษณะเพียงบางส่วนของแบบจำลองโรคเบาหวาน

ชื่อข้อมูล	ชนิดข้อมูล	ความหมาย
District	string	ชื่ออำเภอ
Hospitals Health	int	จำนวนโรงพยาบาลส่งเสริมสุขภาพชุมชน
All Population	int	จำนวนประชากรทั้งหมด
Male	int	จำนวนประชากรเพศชาย
Female	int	จำนวนประชากรเพศหญิง
Vulnerable Groups	int	จำนวนผู้ที่อยู่ในกลุ่มเสี่ยง(อายุ 35-59 ปี)
Older	int	จำนวนผู้สูงอายุ(อายุ 60 ขึ้นไป)
Income Years	int	จำนวนรายได้ต่อปี
HyperPatien	int	จำนวนผู้ป่วยโรคความดันโลหิตสูง
DiabetesPatien	int	จำนวนผู้ป่วยโรคเบาหวาน
Summary	string	ค่าผลลัพธ์ที่ใช้การพยากรณ์

ตาราง 2. แสดงคุณลักษณะเพียงบางส่วนของแบบจำลองโรคความดันโลหิตสูง

ชื่อข้อมูล	ชนิดข้อมูล	ความหมาย
District	string	ชื่ออำเภอ
Hospitals Health	int	จำนวนโรงพยาบาลส่งเสริมสุขภาพชุมชน
All Population	int	จำนวนประชากรทั้งหมด
Male	int	จำนวนประชากรเพศชาย
Female	int	จำนวนประชากรเพศหญิง
Vulnerable Groups	int	จำนวนผู้ที่อยู่ในกลุ่มเสี่ยง(อายุ 35-59 ปี)
Older	int	จำนวนผู้สูงอายุ(อายุ 60 ขึ้นไป)
Income Years	int	จำนวนรายได้ต่อปี
HyperPatien	int	จำนวนผู้ป่วยโรคความดันโลหิตสูง
Summary	string	ค่าผลลัพธ์ที่ใช้การพยากรณ์

3.3.3 แปลงไฟล์ข้อมูลเพื่อการประมวลผล

เพื่อการประมวลผลนั้นจะทำการแปลงข้อมูลให้อยู่ในรูปแบบของไฟล์นามสกุล ARFF สำหรับนำไปใช้ในการประมวลผลด้วยโปรแกรมเวก้าเพื่อสร้างแบบจำลองการพยากรณ์ต่อไป โดยภายในแฟ้มข้อมูลจะมีส่วนประกอบต่างๆ ได้แก่ ส่วนชื่อของไฟล์ (Relation), แอททริบิวต์ (Attribute) และข้อมูล (Data) ดังรูป 2

3.4 การสร้างแบบจำลองพยากรณ์

การสร้างแบบจำลองพยากรณ์ โดยใช้กฎการจำแนกเทคนิคต้นไม้ตัดสินใจ โดยขั้นตอนวิธีการพัฒนาแบบจำลองของระบบช่วยพยากรณ์พื้นที่เสี่ยงต่อโรคเบาหวานและ ความดันโลหิตสูงใน

```

Relation diabetes

@attribute District {BangKachum,BangKakam,ChatTrakan,MuangPhitsanulok,NakhonThai,WoonBorabong,PhromPhirak,WangThung,WardBoi}
@attribute HospitalsHealth numeric
@attribute AllPopulation numeric
@attribute Male numeric
@attribute Female numeric
@attribute VulnerableGroups numeric
@attribute Older numeric
@attribute IncomeYears numeric
@attribute HyperPatien numeric
@attribute DiabetesPatien numeric
@attribute Summary {medium,high,low}

@data
BangKachum,13,48313,23071,24742,18816,4719,71733,4199,2147,medium
BangKachum,13,48388,23621,24779,18796,7615,74343,4124,2456,medium
BangKachum,13,48307,23612,24695,18570,8046,78618,6846,2592,high
BangKachum,13,48152,23494,24688,18570,8046,78618,6846,2592,high
BangKachum,13,48026,23498,24586,18531,8017,78559,6890,2595,high
BangKachum,20,34020,46235,47787,34820,10461,74879,4872,2546,low
BangKachum,20,34678,46489,48089,34874,12380,48078,6881,3746,low
BangKachum,20,34832,46629,48203,35832,13847,76149,6874,3978,medium
BangKachum,20,34800,46684,48236,35832,13847,76149,6868,4249,medium
BangKachum,20,35146,46689,48447,36372,14932,82100,11349,4614,medium
ChatTrakan,10,40146,20333,19811,33046,3106,40743,2479,1266,low
ChatTrakan,12,40432,20457,19975,15090,4228,51874,2478,1345,low
ChatTrakan,10,40633,20522,20110,15275,4484,51652,3008,1487,low
ChatTrakan,10,40801,20591,20210,15217,4594,45565,3364,1510,low
ChatTrakan,10,41102,20749,20355,15700,4940,71217,3597,1587,low
MuangPhitsanulok,24,20502,135421,145501,93509,20256,71052,17134,8463,low
MuangPhitsanulok,24,20902,135421,145501,104305,37655,74702,18057,8634,low
MuangPhitsanulok,24,20176,135687,146085,103609,40700,81460,20893,9680,low
MuangPhitsanulok,24,209419,136124,147120,103609,40700,96199,20850,10795,low
MuangPhitsanulok,24,206401,137465,148268,104704,44470,116202,21078,11204,low
NakhonThai,20,45534,42886,42444,20305,8235,49534,5681,2118,low
NakhonThai,20,46163,43161,43002,20324,11257,50023,4738,2476,low
NakhonThai,20,46684,43444,43220,20415,13854,45540,7259,3028,low
NakhonThai,20,47042,43603,43459,20415,12835,47589,6271,2777,medium
NakhonThai,20,47331,43714,43636,20404,13112,70747,5826,4574,high

```

รูปที่ 2. ตัวอย่างรูปแบบของไฟล์นามสกุล ARFF ของโรคเบาหวาน

พื้นที่จังหวัดพิษณุโลก โดยใช้กฎการจำแนกเทคนิคต้นไม้ตัดสินใจ นั้น ผู้วิจัยเลือกใช้โปรแกรมเวก้าในการสร้างแบบจำลองพยากรณ์ที่อาศัยเทคนิคต้นไม้ตัดสินใจด้วยอัลกอริทึม J48 และอัลกอริทึม J48graft ซึ่งอัลกอริทึมดังกล่าวเป็นอัลกอริทึมที่เป็นที่รู้จักและนิยมนำมาใช้อย่างแพร่หลาย

หลังจากนั้นข้อมูลที่ได้จากไฟล์นามสกุล arff จากรูปที่1 และรูปที่2 นั้น จะถูกนำมาสร้างแบบจำลอง เมื่อนำเข้าโปรแกรม Weka แสดงข้อมูลได้ดังรูปที่3 และรูปที่4 ดังต่อไปนี้



รูปที่ 3. ข้อมูลทั้งหมดที่เตรียมไว้สำหรับวิเคราะห์เพื่อสร้างแบบจำลองของโรคเบาหวาน



รูปที่ 4. ข้อมูลทั้งหมดที่เตรียมไว้สำหรับวิเคราะห์เพื่อสร้างแบบจำลองของโรคเบาหวาน

4. ผลการดำเนินงาน

ผลจากการดำเนินงานได้ทำการทดสอบแบบจำลองพยากรณ์ทุกแบบ มีความสามารถในการจำแนกกลุ่มได้ดี แบบจำลองการพยากรณ์พื้นที่เสี่ยงต่อโรคเบาหวานที่ถูกพัฒนาโดยอัลกอริทึม J48 และ J48graft Cross-Validation ชนิด 5 และ 10-Fold มีค่าความถูกต้องของแบบจำลองมากที่สุดซึ่งอยู่ที่ 100% แบบจำลองการพยากรณ์พื้นที่เสี่ยงต่อโรคความดันโลหิตสูงที่ถูกพัฒนาโดยอัลกอริทึม J48 และ J48graft ด้วยวิธี Cross-Validation 5 และ 10-Fold มีมีค่าความถูกต้องของแบบจำลองมากที่สุดซึ่งอยู่ที่ 100% (ดังตาราง 3 และตาราง 4 ตามลำดับ)

ตาราง 3. รูปแบบของแบบจำลองโรคเบาหวาน

รูปแบบของแบบจำลอง	J48 (%)	J48Graft (%)
Percentage split (10)	80.0	80.0
Percentage split (20)	100	100
Percentage split (66)	86.67	86.67
Cross-validation (Fold=5)	100	100
Cross-validation (Fold=10)	100	100

4.1 ตัวอย่างการทำงานของระบบ

การทำงานของระบบโปรแกรมจะแบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 เป็นส่วนการทำงานของพยากรณ์พื้นที่เสี่ยงต่อโรคเบาหวาน และสำหรับส่วนที่ 2 จะเป็นการทำงานการพยากรณ์พื้นที่เสี่ยงต่อโรคเบาหวานและโรคความดันโลหิตสูง (ดังรูปที่ 5 และรูปที่ 6 ตามลำดับ)

ตาราง 4. รูปแบบของแบบจำลองโรคความดันโลหิตสูง

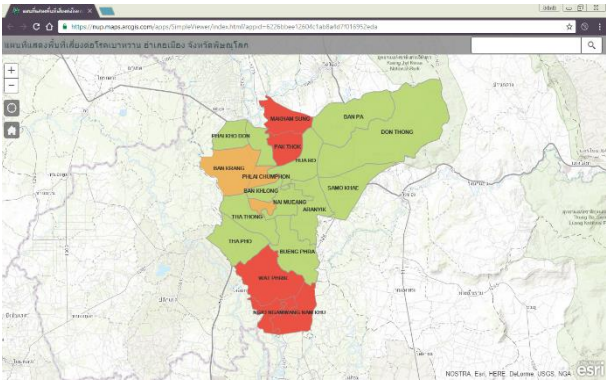
รูปแบบของแบบจำลอง	J48 (%)	J48Graft (%)
Percentage split (10)	80.0	80.0
Percentage split (20)	77.78	77.78
Percentage split (66)	90	90.0
Cross-validation (Fold=5)	77.78	77.78
Cross-validation (Fold=10)	100	100

รูปที่ 5. แสดงหน้าของระบบพยากรณ์พื้นที่เสี่ยงต่อโรคเบาหวาน

รูปที่ 6. แสดงหน้าของระบบพยากรณ์พื้นที่เสี่ยงต่อโรคความดันโลหิตสูง

โดยผลลัพธ์ที่แสดงออกมาเป็นข้อความและแผนที่แสดงพื้นที่ที่ผู้ใช้เลือก (ดังรูปที่ 7 และ รูปที่ 8 ตามลำดับ)

รูปที่ 7. ตัวอย่างหน้าจอแสดงผลเป็นข้อความ



รูปที่ 8. ตัวอย่างหน้าจอแสดงผลเป็นแผนที่

5. สรุป

จากผลการทดลองพบว่าเว็บแอปพลิเคชันระบบช่วยพยากรณ์พื้นที่เสี่ยงต่อโรคเบาหวานและความดันโลหิตสูง ด้วยเทคนิคต้นไม้ตัดสินใจที่พัฒนาขึ้นนั้นได้บรรลุตามวัตถุประสงค์ โดยสามารถนำเอาเทคนิคการทำเหมืองข้อมูล และระบบสารสนเทศภูมิศาสตร์มาใช้ประกอบในการทำเว็บแอปพลิเคชัน ซึ่งเป็นรูปแบบของโปรแกรมที่สามารถเข้าถึงได้ง่ายผ่านเว็บเบราว์เซอร์ของคอมพิวเตอร์และโทรศัพท์มือถือที่เชื่อมต่อกับอินเทอร์เน็ต โดยหน่วยงานทางสาธารณสุขสามารถใช้เว็บแอปพลิเคชันดังกล่าว เพื่อทำการกรอกข้อมูลแล้วดูผลการทำนายที่ได้ เพื่อเป็นแนวทางในการวางแผนนโยบายด้านสาธารณสุขต่อไป

เอกสารอ้างอิง

- [1] นงเยาว์ นออรุณ. การประยุกต์ใช้เทคนิคเหมืองข้อมูลในการพยากรณ์โรคเบาหวาน กรณีศึกษาโรงพยาบาลศูนย์ สวรรค์ประชารักษ์. มหาวิทยาลัยนเรศวร, พิษณุโลก, 2557.
- [2] บุญเสริม กิจศิริกุล. รายงานเรื่องอัลกอริทึมการทำเหมืองข้อมูล. คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2546.
- [3] บัณฑิตวิทยาลัยการจัดการและนวัตกรรม ม.เทคโนโลยีพระจอมเกล้าธนบุรี. HTML. สืบค้นเมื่อ 5 เมษายน 2559, เข้าถึงได้จาก <http://www.codingbasic.com/html>
- [4] ปรีดา ชลสุข และพจนวรัญ ไหลเจริญ. การประยุกต์ใช้ระบบสารสนเทศภูมิศาสตร์เพื่อเฝ้าระวังโรคระบาดในพื้นที่อำเภอแควศรีสำโรง จังหวัดสุโขทัย. มหาวิทยาลัยนเรศวร, 2548.
- [5] พวงทอง ไกรพิบูลย์. ความดันโลหิตสูง (Hypertension). สืบค้นเมื่อ 20 กุมภาพันธ์ 2559, เข้าถึงได้จาก <http://haamor.com/th/ความดันโลหิตสูง/>
- [6] วรวิทย์ บำรุงพงษ์, วัชรภรณ์ วิชุกร และเสาวนีย์ ติมูล. การวิเคราะห์หาพื้นที่เสี่ยงโรคมะเร็งในพื้นที่จังหวัดตาก. มหาวิทยาลัยนเรศวร, 2552.

[7] สุกลักษณ์ จันทรมบัติ. การพยากรณ์ความต้องการน้ำประปาในพื้นที่เทศบาลนครพิษณุโลกโดยใช้แบบจำลองคณิตศาสตร์และระบบสารสนเทศภูมิศาสตร์. มหาวิทยาลัยนเรศวร, 2547.

[8] สำนักงานสาธารณสุข จังหวัดพิษณุโลก. ตัวชี้วัดกระทรวงสาธารณสุข. สืบค้นเมื่อ 12 เมษายน 2559, เข้าถึงได้จาก http://plk.hdc.moph.go.th/hdc/main/?sele_kpiyearc=2015&level=3&sele_kpi=848b5045eab655b7be0069ecc445dd9

[9] หน่วยปฏิบัติการวิจัยระบบการจัดการแหล่งน้ำ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย. คู่มือการใช้งานโปรแกรม Quantum GIS. สืบค้นเมื่อ 16 มีนาคม 2559, เข้าถึงได้จาก http://projectwre.eng.Chula.ac.th/thai_waterplan/sites/default/km_wp/manual/quantum_gisv1.7.4.pdf

[10] Better Health. รู้ทัน เข้าใจ อยู่อย่างปลอดภัยกับโรคความดันและเบาหวาน. สืบค้นเมื่อ 20 พฤศจิกายน 2559, เข้าถึงได้จาก <https://www.bumrungrad.com/th/BetterHealth/2553/30thanniversary/living-with-hypertension-and-diabetes>

[11] Enjoyday.net. CSS คืออะไร. สืบค้นเมื่อ 11 เมษายน 2559, เข้าถึงได้จาก http://www.enjoyday.net/webTutorial/css/css_chapter01.html

[12] Gisthai.org. ระบบสารสนเทศภูมิศาสตร์. สืบค้นเมื่อ 20 มีนาคม 2559, เข้าถึงได้จาก <http://www.gisthai.org/about-gis/gis.html>

[13] RiderOOU. เว็บแอปพลิเคชัน (Web Application) คืออะไร ?. สืบค้นเมื่อ 20 มีนาคม 2559, เข้าถึงได้จาก <http://aicomputer.co.th/sArticle/002-what-is-Web-Application.aspx>

[14] Scorpio. Weka สำหรับทำ Data Mining. สืบค้นเมื่อ 11 เมษายน 2559, เข้าถึงได้จาก <http://goo.gl/wsY6Ew>

[15] Siamhealth.net. โรคเบาหวานและความดันโลหิตสูง. สืบค้นเมื่อ 1 มีนาคม 2559, เข้าถึงได้จาก http://www.siamhealth.net/public_html/Disease/endocrine/DM/htm#VxhVSHGqkqp

[16] Stoms Web. Responsive Web Design การแสดงผลเว็บไซต์ที่รองรับทุกหน้าจอการทำงาน. สืบค้นเมื่อ 20 พฤศจิกายน 2559, เข้าถึงได้จาก <http://web.stoms.co.th/2013/10/15/responsive-web-design/>

[17] Zarazi. JavaScript Overview. สืบค้นเมื่อ 11 เมษายน 2559, เข้าถึงได้จาก https://developer.mozilla.org/docs/Web/JavaScript/Guide/JavaScript_Overview