การเปรียบเทียบโมเดลการเรียนรู้ของเครื่องสำหรับคัดกรองผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้า

The comparison of learning models of machine for screening diabetic patients who has foot numbness

อรุณรักษ์ ตันพานิช[1*], ดุษณี ศุภวรรธนะกุล[2], พิเชฐ บัญญัติ[3] และจรูญ จันแทน[4]

Aroonrak Tunpanit[1*], Dusanee Supawantanakul[2], Phichet Banyati[3] and Charoon Chantan[4]

[1] นักศึกษาระดับปริญญาเอก, หลักสูตรปรัชญาดุษฎีบัณฑิต, สาขาวิชาการจัดการเทคโนโลยี, มหาวิทยาลัยราชภัฏพระนคร

[1] Ph.D. Student in Technology Management, Phranakhon Rajabhat University.

[2] ผู้ช่วยศาสตราจารย์ ดร., มหาวิทยาลัยราชภัฏพระนคร

[2] Assistant professor Dr., Technology Management, Phranakhon Rajabhat University.

[3] นายแพทย์ ดร., พท., รองอธิบดีกรมวิทยาศาสตร์การแพทย์

[3] M.D. Dr., Thai Traditional Medicine., Deputy director general of Department of Medical Sciences

[4] อาจารย์ ดร., นักวิเคราะห์ระบบงานอาวุโส, ธนาคารกสิกรกลุ่มเทคโนโลยีทางธุรกิจ

[4] Lecturer Dr., Senior Solution Analyst, Kasikorn Business Technology Group (KBTG)

[*]**Corresponding author, E-mail:** aroonrak1982@windowslive.com

บทคัดย่อ

ปัญหาผู้ป่วยโรคเบาหวานมีแนวโน้มเพิ่มขึ้นทั่วโลกโดยเฉพาะเท้าเบาหวาน ระยะแรกผู้ป่วยโรคเบาหวานจะรู้สึกชาปลายเท้าหากปล่อยปละละเลยจะนำไปสู่แรงกดทับ เกิดแผลเท้า เนื่องจากเลือดมีความหนืดแผลจึงหายช้าและติดเชื้อในที่สุด แต่สามารถป้องกันได้ด้วยกระบวนการคัดกรองภาวะชาปลายเท้า ซึ่งงานวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อสังเคราะห์ข้อมูลและเปรียบเทียบโมเดลกระบวนการเรียนรู้ของเครื่องด้วยโปรแกรม RapidMiner Studio เปรียบเทียบประสิทธิภาพอัลกอริทึมของกระบวนการเรียนรู้ของเครื่อง ณ โรงพยาบาลส่งเสริมสุขภาพประจำตำบลท่าจีน จังหวัดสงขลา โดยมีกลุ่มตัวอย่างจำนวน 2 กลุ่ม ได้แก่ กลุ่มผู้ป่วยโรคเบาหวานชนิดที่ 2 ที่มีภาวะชาปลายเท้า จำนวน 50 คน และคนปกติ จำนวน 50 คน ที่มีช่วงอายุตั้งแต่ 40-70 ปี มีน้ำหนักตัว 60-80 กิโลกรัม ข้อมูลส่วนบุคคลที่ใช้สังเคราะห์ ได้แก่ รหัส เพศ อายุ น้ำหนัก ส่วนสูง รอบเอว ความดันโลหิต ประวัติโรคเบาหวาน ค่าดัชนีมวลกาย ค่า Risk Scale และค่า Gray Scale โดยมีใบนัดเพื่อเก็บข้อมูลจำนวน 3 ครั้ง รวมจำนวนทั้งสิ้น 300 ข้อมูล ผลการศึกษาพบว่าปัจจัยส่วนบุคคลด้านอายุมากกว่า 55 ปี และมีประวัติเกี่ยวกับความดันโลหิตสูง มีโอกาสที่จะเป็นเบาหวานมากกว่ากลุ่มในช่วงอายุอื่น และผลการเปรียบเทียบโมเดลทั้ง 3 วิธี พบว่า Support Vector Machine (SVM) มีค่าความเชื่อมั่น (Accuracy) สูงสุด โดยคิดเป็นร้อยละ 93.3 รองลงมาโมเดล Deep Learning มีค่าความเชื่อมั่น (Accuracy) โดยคิดเป็นร้อยละ 72.7 ลำดับที่ 3 โมเดล Random Forest มีค่าความเชื่อมั่น (Accuracy) โดยคิดเป็นร้อยละ 65.3 ซึ่งดังนั้นโมเดลแบบ Support Vector Machine (SVM) ซึ่งในบทความนี้เป็นส่วนหนึ่งของงานวิจัยเรื่องการพัฒนาอุปกรณ์คัดกรองภาวะชาปลายเท้าของผู้ป่วยโรคเบาหวาน

คำสำคัญ: ผู้ป่วยโรคเบาหวาน, ภาวะชาปลายเท้า, กระบวนการเรียนรู้ของเครื่อง

Abstract

The problem of diabetes is increasing worldwide. Especially diabetic foot initially, diabetic patients will feel the toe of the tea, if neglected, will lead to pressure, causing foot ulcers due to the viscous blood, the wound will eventually heal slowly and become infected. But can be prevented by the screening process for toe numbness which a study titled “A Comparison of learning models of machine for screening diabetic patients who has toe numbness” had the objectives to synthesize data and compare the learning process of the machine with Program named RapidMiner Studio in order to compare Algorithm performance of the machine learning process at Tachin Subdistrict Health Promotion Hospital, Songkhla. The population in this study were divided in 2 groups as follows, type 2 diabetic patients who has toe numbness totaled 50 people and normal people totaled 50 people, age between 40–70 years old, weight between 60-80 Kilograms. Personal information that is used to synthesize ; Code, Gender, Age, Weight, Height, Waistline, Blood pressure, Diabetes history, Body mass index, Risk Scale and Gray Scale. During doing the research, there are 3 times of appointment to collect the information with 300 information of population. The results of the study were as follows; Personnel information in terms of age above 55 years old and High Blood pressure patients have high risk to have diabetes more than other ages. The results of comparison of 3 models showed that Support Vector Machine Model (SVM) has the most accuracy by 93.3 percent. Secondly, Deep Learning Model has the reliability with 65.3 percent. Thus, Support Vector Machine Model (SVM) in this article is a part of a research titled Development of screening devices for patients with toe numbness Diabetes.

Keywords: Diabetes patient, Foot numbness, Machine learning

บทนำ

โรคเบาหวานเป็นโรคที่มีแนวโน้มผู้เสียชีวิตทั่วโลกเพิ่มขึ้นทุกปีและยังมีภาวะแทรกซ้อนต่อสุขภาพบั่นทอนคุณภาพชีวิตทำส่งผลให้มีค่าใช้จ่ายในการรักษาเพิ่มขึ้น ซึ่งผู้ป่วยโรคเบาหวานส่วนใหญ่เป็นโรคเบาหวานประเภทที่ 2 จากการสำรวจพบความชุกสถิติผู้ป่วยโรคเบาหวานทั่วโรค ปี ค.ศ. 2015 จำนวน 415 ล้านคน โดยมีช่วงอายุตั้งแต่ 20-79 ปี และคาดว่าในปี ค.ศ. 2040 จะมีผู้ป่วยโรคเบาหวานเพิ่มขึ้นเป็น 642 ล้านคน (International Diabetes Federation, 2015) สำหรับประเทศไทยโรคเบาหวาน 1 ใน 5 โรคที่ส่งผลกระทบต่อสาธารณสุขของประเทศไทย จากการสำรวจของกระทรวงสาธารณสุข พบว่า ประชากรไทยวัยผู้ใหญ่ป่วยเป็นโรคเบาหวานถึง 4.8 ล้านคน โดย พ.ศ. 2552 ความชุกเบาหวานจากร้อยละ 6.9 ในปี พ.ศ. 2558 ความชุกเพิ่มขึ้นเป็น ร้อยละ 8.9 คิดเป็นจำนวนเพิ่มขึ้น ปีละประมาณ 300,000 คน และผู้ป่วยโรคเบาหวานเสียชีวิตในกลุ่มอายุ 30-69 ปี โดยเป็นเพศหญิงมากกว่าเพศชาย โดยในปี พ.ศ. 2555 มีผู้เสียชีวิตจากโรคเบาหวานรวม 13.2 และในปี พ.ศ. 2558 เสียชีวิตเพิ่มเป็น 17.8 (สำนักโรคไม่ติดต่อกรมควบคุมโรค กระทรวงสาธารณสุข พ.ศ. 2560-2564, 2560) คาดการณ์ว่าความชุกของโรคเบาหวานจะ

ส่งผลต่อภาวะแทรกซ้อนต่ออวัยวะต่าง ของผู้ป่วยโรคเบาหวาน และคาดการว่าภายในปี พ.ศ.2583 จะเพิ่มสูงขึ้นถึง 5.3 ล้านคน ซึ่งโรคเบาหวานที่ถูกดูแลรักษาได้ไม่ดีอาจทำให้เกิดภาวะแทรกซ้อนต่ออวัยวะส่วนต่าง ๆ ของร่างกาย (ชัชลิต รัตรสาร, 2560)

ผู้ป่วยเบาหวานคือผู้ที่มีภาวะระดับน้ำตาลในเลือดสูงเกิน 125 มิลลิกรัมต่อเดซิลิตร (mg/dL) ติดต่อกันเป็นเวลา 5 ปีขึ้นไป (วีระศักดิ์ ศรินนภากร, 2550: 7) ส่งผลให้ระดับน้ำตาลในเลือดสูงและเกิดพยาธิสภาพของอวัยวะหลายส่วน เนื่องจากเลือดของผู้ป่วยเบาหวานมีความเข้มข้นทำให้ไม่สามารถไปเลี้ยงส่วนปลายของอวัยวะต่าง ๆ ได้ เช่น เลือดไปเลี้ยงส่วนปลายของหัวใจไม่เพียงพอจะทำให้เกิดหลอดเลือดหัวใจตีบ กล้ามเนื้อหัวใจตาย หรือเลือดไปเลี้ยงที่สมองไม่พอจะส่งผลให้สมองตีบตัน (อยุทธินี สิงหวินท์. 2557) และเส้นเลือดแตกทำให้เกิดการเป็นอัมพฤกษ์หรืออัมพาตถ้าหากเป็นที่ตาจะเกิดจอตาเสื่อมมีเลือดออกและตาบอดในที่สุด หากเกิดที่ไตจะทำให้ไตขาดเลือดเกิดภาวะไตเสื่อมและไตวาย และในกรณีที่เลือดไปเลี้ยงส่วนปลายของเท้าไม่เพียงพอจะทำให้เกิดผลกระทบของปลายประสาท เส้นประสาทถูกทำลาย การรับรู้ความรู้สึกต่าง ๆ ลดลง ทำให้มีบาดแผลที่เท้าเนื่องมาจากไม่สามารถรับรู้การสัมผัสอุณหภูมิของแหลมคมต่าง ๆ ผู้ป่วยเบาหวานจึงมีบาดแผลหรือแผลติดเชื้อลุกลามจนนำไปสู่การตัดเท้าในที่สุด (ดุษณี ศุภวรรธนะกุล, 2554: 2) ผู้วิจัยตระหนักถึงปัญหาผู้ป่วยโรคเบาหวานที่เพิ่มขึ้นและภาวะเสี่ยงที่จะสูญเสียเท้าจากโรคเบาหวาน ซึ่งนำข้อมูลสุขภาพของผู้ป่วยโรคเบาหวานมารวบรวมเป็นระบบข้อมูลสุขภาพ (HEALTH CARE) เพื่อการวิจัยและการวิเคราะห์เพื่อช่วยตัดสินใจทางการแพทย์ในการประเมินความเสี่ยงโรคเบาหวานด้วยค่า Risk score เพื่อให้สถาบันทางการแพทย์ดึงข้อมูลที่เป็นประโยชน์มาวินิจฉัยโรคด้วยระบบเครื่องคอมพิวเตอร์ก็มีการประมวลผลอย่างรวดเร็วและสามารถขุดหาข้อมูลที่มีความสำคัญด้วยกระบวนการเรียนรู้ของเครื่อง (Machine learning) ซึ่งเป็นเครื่องมือที่สามารถวิเคราะห์ประเมินเป็นค่าความถูกต้องมีความละเอียดอ่อนและเชื่อถือได้ในการค้นพบรูปแบบและความสัมพันธ์ของชุดข้อมูลสุขภาพด้วยซอฟต์แวร์ RapidMiner Studio

ในการประเมินความเสี่ยงโรคเบาหวานโดยใช้กระบวนการเรียนรู้ของเครื่องในการเรียนรู้ข้อมูลและทำการวิเคราะห์ข้อมูลและทำนายผลด้วยการเปรียบเทียบต่างโมเดล ต่างอัลกอริทึมเพื่อหารูปแบบการอธิบายข้อมูลที่มีอย่างมากมายเพื่อหาสาระสำคัญของชุดข้อมูลและช่วยพยากรณ์วิเคราะห์ข้อมูลความเสี่ยงโรคเบาหวานและสร้างแบบจำลองการทำนายโรคเบาหวาน ซึ่งในบทความนี้เป็นส่วนหนึ่งของงานวิจัยเรื่องการพัฒนาอุปกรณ์คัดกรองฝ่าเท้าผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้า

วัตถุประสงค์ของการวิจัย

1. เพื่อสังเคราะห์ข้อมูลระหว่างผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้าและข้อมูลคนปกติ
2. เพื่อเปรียบเทียบโมเดลกระบวนการเรียนรู้ของเครื่องด้วยโปรแกรม RapidMiner Studio

แนวคิด ทฤษฎี กรอบแนวคิด

1. โรคเบาหวานชนิดที่ 2

โรคเบาหวานชนิดที่ 2 (Diabetes Mellitus Type 2) หรือโรคเบาหวานชนิดไม่พึ่งอินซูลิน (Non - Insulin Dependent Diabetes Mellitus = NIDDM) เป็นโรคเบาหวานชนิดที่เป็นกันมาก พบอุบัติการณ์การเกิดโรคมากถึงร้อยละ 95-98 มักถ่ายทอดทางพันธุกรรมและเกิดกับคนอายุมากกว่า 40 ปีขึ้นไป อาจพบในเด็กหรือวัยหนุ่มสาวได้บ้าง โรคเบาหวานชนิดนี้ตับอ่อนยังสามารถสร้างอินซูลินได้แต่

ร่างกายไม่สามารถนำอินซูลินไปใช้ได้และอาจตรวจพบระดับอินซูลินปกติหรือสูงกว่าปกติเพราะมีภาวะดื้ออินซูลินสาเหตุอาจเกิดได้จากหลายปัจจัย ได้แก่ พันธุกรรม โรคอ้วน มีลูกดก การใช้ยาหรือพบร่วมกับโรคอื่น ๆ พยาธิสภาพของโรคเบาหวาน ชนิดที่ 2 เกิดจาก 2 ปัจจัย ได้แก่ 1) เซลล์ตับอ่อนสามารถหลั่งอินซูลินได้ในปริมาณน้อยไม่เผาพลาญไขมันนำมาใช้เป็นพลังงานทำให้มีภาวะระดับน้ำตาลในเลือดสูง ผู้ป่วยเบาหวานมีรูปร่างอ้วนร่างกายจึงไม่เกิดภาวะกรดคั่งแต่เกิดภาวะวิกฤตจากระดับน้ำตาลในเลือดที่สูง 2) เกิดภาวะดื้อต่ออินซูลิน คือ การที่ภาวะอินซูลินปกติแต่เซลล์ในเนื้อเยื่อไม่สามารถนำกลูโคสไปใช้ได้ นอกจากนี้ยังมีการผลิตน้ำตาลจากตับเพิ่มขึ้นการขาดอินซูลิน ทำให้กลูโคสจากอาหารไม่สามารถเก็บสะสมที่ตับในรูปไกลโคเจนได้ ระดับน้ำตาลในเลือดจึงสูงเมื่อระดับน้ำตาลในเลือดสูงเกินความสามารถของไตจะดูดซึมกลูโคสได้หมดทำให้ตรวจพบระดับน้ำตาลในปัสสาวะได้ (Powered by exteen blog. 2553: สื่ออิเล็กทรอนิกส์)

1.1 การประเมินความเสี่ยงในช่วงเวลาหรือเรียกอีกอย่างหนึ่งว่าการศึกษาชนิดตัดขวาง (Prevalence หรือ cross-sectional study) เป็นการประเมินความเสี่ยงรูปแบบนี้ใช้สำหรับการตรวจ คัดกรอง (Screening) เพื่อค้นหาผู้ป่วยเบาหวานที่ยังไม่มีอาการและหาทางป้องกันไม่ให้เกิดภาวะแทรกซ้อนในระยะเริ่มแรก (Keesukphan P. Chanprasertyothin S, Ongphiphadhanakul B, Puavilai G, 2007: 43) โดยใช้สมการในการประเมินความเสี่ยง (risk score) ดังสมการต่อไปนี้

$$\text{Risk Score} = (3 \times \text{AGE}) + (5 \times \text{BMI}) + (50 \times \text{Profile})$$

1.2 การประเมินความเสี่ยงเพื่อป้องกันโรค เป็นการประเมินความเสี่ยงเพื่อทำนายผู้ที่มีความเสี่ยงที่จะเกิดโรคเบาหวานในอนาคต (Incident case) เพื่อป้องกันให้บุคคลนั้นตระหนักถึงความเสี่ยงที่จะเกิดโรคเบาหวานในอนาคตได้สูงกว่าคนปกติ จึงสมควรดูและตนเองและรักษาสุขภาพเพื่อป้องกันปัจจัยเสี่ยงโรคเบาหวาน (สมาคมโรคเบาหวานแห่งประเทศไทย, 2554 : 19-20) หากผลคำนวณค่ามากกว่า 240_แสดงว่าบุคคลมีความเสี่ยงต่อโรคเบาหวานสูง

2. การเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง (machine learning) เป็นเทคนิคการเรียนรู้ส่วนมากเป็นการเรียนรู้เชิงอุปนัย (inductive learning) โดยการเรียนรู้เชิงอุปนัย คือ การเรียนรู้ที่หากฎเกณฑ์หรือความรู้ที่แฝงอยู่ในชุดตัวอย่างสอน (training example set) เพื่อเรียนรู้ให้ได้ความรู้ที่สอดคล้องกับชุดตัวอย่างสอน และอีกเทคนิคหนึ่งคือ การเรียนรู้เชิงวิเคราะห์ (analytical learning) เป็นการจัดรูปแบบของความรู้ใหม่เพื่อให้ใช้งานได้อย่างมีประสิทธิภาพมากขึ้น ทำงานได้รวดเร็วขึ้น (บุญเสริม กิจศิริกุล, 2548)

การพัฒนากระบวนการเรียนรู้ของเครื่อง (machine learning) เป็นการวิเคราะห์ข้อมูลในคลังข้อมูลเพื่อช่วยสนับสนุนการตัดสินใจในการบริหารจัดการองค์กร ประกอบด้วยระบบข้อมูลและโปรแกรมแอพพลิเคชั่นด้านการวิเคราะห์หลายระบบตัวแบบการทำนาย (Predictive Modeling) เป็นโมเดลการจำแนก (Classifier Model) ที่สามารถใช้ในการทำนายตัวแปรที่ไม่ทราบค่า (Class Label) ซึ่งกระบวนการสร้างโมเดลโดยการนำเอาข้อมูลมาแบ่งเป็น 2 ส่วน โดยนำส่วนแรกมาสอนแก่ระบบ (Training Data) เพื่อสร้างโมเดลการจำแนกและนำข้อมูลส่วนที่สองซึ่งเป็นข้อมูลทดสอบ (Testing Data) มาทำการทดสอบโมเดลที่ถูกสร้างขึ้นเพื่อวัดค่าความถูกต้องของโมเดลและเมื่อต้องการใช้งานก็นำข้อมูลที่ต้องการทำนายมาผ่านโมเดลการจำแนก

ตัวแบบการทำนาย (Predictive Modeling) เป็นโมเดลการจำแนก (Classifier Model) ที่สามารถใช้ในการทำนายตัวแปรที่ไม่ทราบค่า (Class Label) ซึ่งกระบวนการสร้างโมเดลโดยการนำเอาข้อมูลมาแบ่งเป็น 2 ส่วน โดยนำส่วนแรกมาสอนแก่ระบบ (Training Data) เพื่อสร้างโมเดลการจำแนกและนำข้อมูลส่วนที่สองซึ่งเป็นข้อมูลทดสอบ (Testing Data) มาทำการทดสอบโมเดลที่ถูกสร้างขึ้นเพื่อวัดค่าความถูกต้องของโมเดลและเมื่อต้องการใช้งานก็นำข้อมูลที่ต้องการทำนายมาผ่านโมเดลการจำแนก (สารัช ตันติกิตติ, 2559: 22)

ตัวแบบการทำนาย (Predictive Modeling) เป็นโมเดลการจำแนก (Classifier Model) ที่สามารถใช้ในการทำนายตัวแปรที่ไม่ทราบค่า (Class Label) ซึ่งกระบวนการสร้างโมเดลโดยการนำเอาข้อมูลมาแบ่งเป็น 2 ส่วน โดยนำส่วนแรกมาสอนแก่ระบบ (Training Data) เพื่อสร้างโมเดลการจำแนกและนำข้อมูลส่วนที่สองซึ่งเป็นข้อมูลทดสอบ (Testing Data) มาทำการทดสอบโมเดลที่ถูกสร้างขึ้นเพื่อวัดค่าความถูกต้องของโมเดลและเมื่อต้องการใช้งานก็นำข้อมูลที่ต้องการทำนายมาผ่านโมเดลการจำแนก

3. โปรแกรม RapidMiner Studio

Rapidminer Studio ซึ่งเป็น Open source software ทำให้สามารถประหยัดค่าใช้จ่าย กรณีต้องการใช้งานเพื่อทดสอบการวิเคราะห์ข้อมูล การทำความสะอาดข้อมูล (Data cleansing) และการรายงานข้อมูลในรูปแบบต่าง ๆ เช่น ตาราง กราฟ (อารีรัตน์ วงศ์สุวรรณ, 2557: ออนไลน์) และมีอัลกอริทึมช่วยในการจำแนกค่าออกเป็นตัวเลขซึ่งมีการยอมรับในด้านการศึกษาและใช้ในงานวิจัยอีกมากมาย โดยมีกระบวนการเรียนรู้ของเครื่องเป็นการสร้างฟังก์ชันจากข้อมูลสอน ข้อมูลสอนจะประกอบด้วยข้อมูลนำเข้าซึ่งมักจะเป็นแวกเตอร์และผลของการเรียนรู้ที่ต้องการ หลังจากการเรียนรู้จะได้ข้อมูลเป็นฟังก์ชันที่อาจจะให้ค่าต่อเนื่องหรือค่าทำนายประเภทของวัตถุ โดยหน้าที่ของการเรียนของเครื่อง คือการทำนายค่าของฟังก์ชันจากวัตถุเข้าที่ถูกต้องโดยใช้ตัวอย่างสอนจำนวนน้อย การเรียนรู้ของเครื่องมีข้อดีได้แก่ การเรียนรู้ของเครื่องถอดความสัมพันธ์และความเกี่ยวข้องที่สำคัญกลุ่มของข้อมูลใหญ่ จำนวนของข้อมูลเรียนรู้มากเกินความสามารถในการคำนวณโดยมนุษย์เครื่องสามารถที่จะเรียนรู้ความรู้นี้ทีละเล็กละน้อยและสามารถจับประเด็นได้ตรงตามที่ต้องการออกมาได้มากกว่ามนุษย์และข้อมูลมีการเปลี่ยนแปลงและเพิ่มเติมได้ตลอดเวลา เครื่องสามารถเรียนรู้ได้อย่างต่อเนื่องเมื่อมีข้อมูลเข้ามาใหม่

จากผลการสำรวจเว็บไซต์ KDnuggets ในปี 2014 พบว่า ผู้ร่วมตอบมีผู้สนใจ Rapidminer มากเป็นอันดับ 1 และมี Gartner ได้จัดให้ Rapidminer อยู่ในกลุ่ม Leaders สำหรับซอฟต์แวร์วิเคราะห์ข้อมูลปี 2016 ได้แก่ SAS, IBM, KNIME, Rapid Miner (นวคราญ คำวิชัย, 2559: 39)

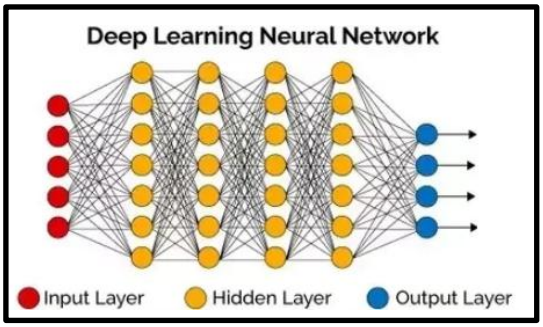
จากวิธีการเรียนรู้ของเครื่องเมื่อได้ค่าความแม่นยำที่สูงสุดแล้ว กระบวนการในการพัฒนาโปรแกรมในระบบปฏิบัติการวินโดวส์จึงเป็นขั้นตอนที่จะเชื่อมโยงระหว่างข้อมูลจากโมเดลผ่านกระบวนการเรียนรู้ของเครื่องไปสู่การใช้งานของผู้ใช้ ซึ่งการพัฒนาโปรแกรมมีส่วนประกอบในการเขียนผังงานเพื่อให้เห็นโครงสร้างโดยรวมของระบบและสามารถเขียนโปรแกรมได้ง่ายยิ่งขึ้น

4. โมเดลการเรียนรู้ของเครื่อง

4.1 Deep Learning

Deep learning คือ ซอฟแวร์คอมพิวเตอร์ที่เลียนแบบการทำงานของระบบโครงข่ายประสาท (neurons) ในสมองมนุษย์ ถือเป็นซับเซ็ตของ machine learning Algorithm ของ deep learning ถูกสร้างขึ้นจากการนำเอา neural network หลายๆ layer มาต่อกัน โดย layer แรกสุดจะทำหน้าที่ในการรับข้อมูล (Input layer) layer สุดท้ายจะทำหน้าที่ส่งผลลัพธ์การประมวลผลออกมา (Output layer) ส่วน

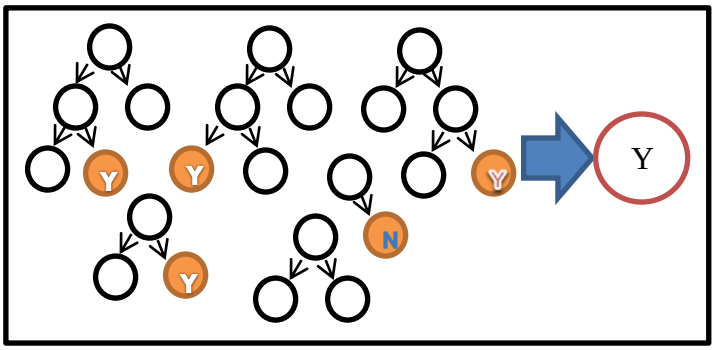
layer ระว่าง layer แรกสุดและ layer สุดท้าย จะถูกเรียกว่า Hidden layer คำว่า deep learning มีที่มาจากการที่ใช้ layer ของ neural network หลาย ๆ อันมาต่อกันโดย Hidden layer ของแต่ละ layer จะเปรียบเสมือนว่าประกอบด้วย เซลล์ประสาท (neural) จำนวนมาก ซึ่งมีหน้าที่ในการประมวลผล รับข้อมูลจาก layer ที่อยู่เหนือกว่า และส่งข้อมูลที่ประมวลผลเสร็จแล้วไปยัง layer ที่อยู่ล่างกว่า, ข้อดีของการส่งข้อมูลแบบนี้ก็คือ layer แต่ละ layer สามารถที่จะมี ค่าถ่วงน้ำหนัก (weight), ค่าความเอนเอียงของข้อมูล (bias) และวิธีการประมวลผลทางคณิตศาสตร์ (activation function) ที่เป็นอิสระต่อกันได้ ยิ่งเราป้อนข้อมูลให้กับ model มากเท่าไหร่, layer แต่ละ layer ก็จะสามารถสกัด feature ที่มีความซับซ้อนมากยิ่งขึ้น



ภาพประกอบที่ 1 โมเดล Neural Net อัลกอริทึม Deep Learning
ที่มา: G.E. Hinton and other 2012.

4.2 Random Forest

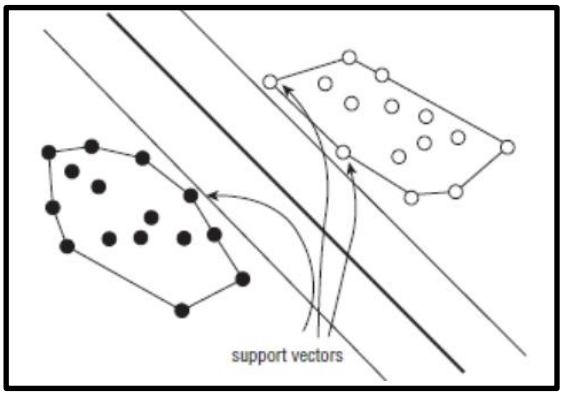
Random Forest แรนดอมฟอเรสต์ ประกอบด้วยแผนผังต้นไม้ช่วยในการประกอบการตัดสินใจ (Decision trees) จำนวนมาก โดยมีต้นไม้แต่ละต้นจะได้รับการจำแนกและต้นไม้จะทำการ “เลือกหมวด” คลาสที่มีประสิทธิภาพ ซึ่งลักษณะดีของแรนดอมฟอเรสต์คือ ดำเนินการได้อย่างมีประสิทธิภาพต่อฐานข้อมูลขนาดใหญ่ สามารถนำตัวแปรนำเข้าปริมาณหลายพันตัวได้โดยไม่ต้องลบตัวแปรออกไปบางส่วน และมีประสิทธิภาพในการประมาณค่าข้อมูลที่หายไปและความถูกต้องขณะที่ข้อมูลส่วนใหญ่หายไป



ภาพประกอบที่ 2 โมเดล Trees อัลกอริทึม Random Forest
ที่มา: Jiancho Han, Juan C. Rodriguze and Mohsen Beheshti.,2008: 26

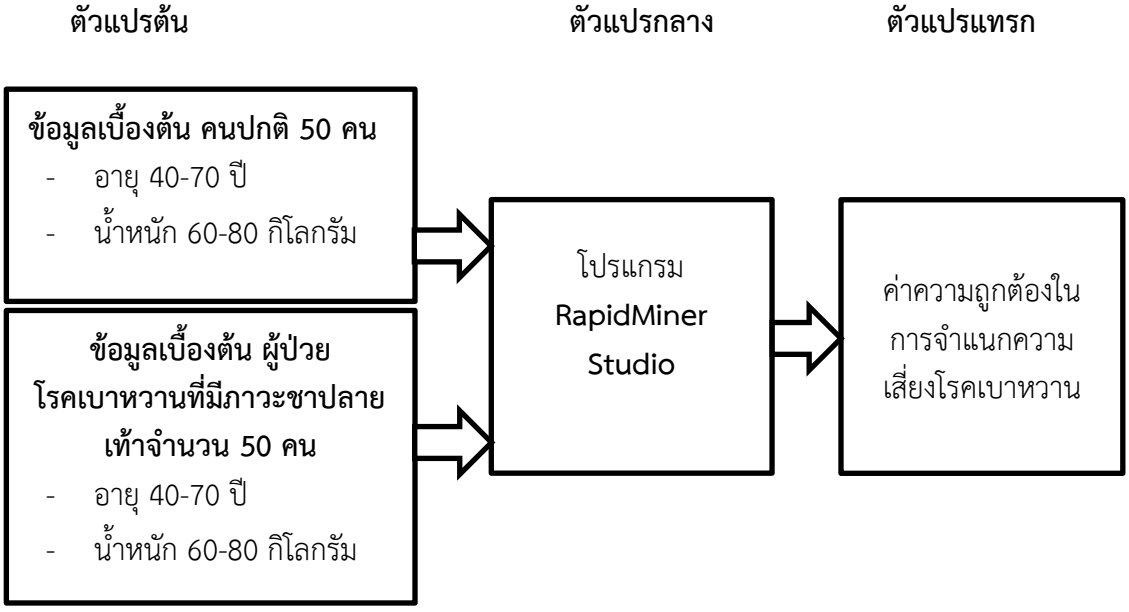
4.3 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines: SVM)

เป็นวิธีการที่นามาแยกกลุ่มของข้อมูลโดยอาศัยระนาบสาหรับแบ่งเขตแดนของกลุ่มข้อมูลออกจากกันเป็นสองฝั่งโดยที่มีเวกเตอร์ (Vector) แทนจำนวนเซตของคุณลักษณะที่ซึ่ง Hyperplane เป็นตัวแยกกลุ่มของเวกเตอร์ส่วน Margin เป็นระยะห่างจากเส้นตรง Hyperplane ถึงเส้นตรงที่ผ่านข้อมูลที่ใกล้ที่สุดและขนานกับ Hyperplane ของทั้งสองกลุ่ม โดยที่ SVM จะเลือก Hyperplane ที่มีค่า Margin สูงสุด (Maximum Margin) (Parker, J.R., 2010: 29)



ภาพประกอบที่ 3 Support Vector Machine ใน 2 มิติ
ที่มา: Parker, J.R., 2007

กรอบแนวคิด



ภาพประกอบที่ 4 กรอบแนวคิด

วิธีดำเนินการวิจัย

ในการพัฒนาแบบจำลองเพื่อช่วยทำนายผลความเสี่ยงโรคเบาหวาน มีขั้นตอนในการวิจัยดังต่อไปนี้

การเตรียมข้อมูลผู้วิจัยคัดเลือกกลุ่มตัวอย่างจำนวน 2 กลุ่ม คือ กลุ่มผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้าและกลุ่มคนปกติ ซึ่งกลุ่มตัวอย่างทั้งสองมีอายุระหว่าง 40 – 70 ปี และน้ำหนักตัวระหว่าง 60 – 80 กิโลกรัม โดยผู้วิจัยลงพื้นที่ ณ โรงพยาบาลส่งเสริมสุขภาพประจำตำบลท่าจีน จังหวัดสงขลา ซึ่งได้ทำหนังสือเพื่อขออนุญาตในการใช้อุปกรณ์ถ่ายภาพฝ่าเท้าและบันทึกข้อมูลผู้ป่วยโรคเบาหวานที่มีระยะเวลา 5 ปีขึ้นไป ซึ่งจะต้องมีภาวะชาปลายเท้าและไม่ถูกตัดอวัยวะใด ๆ จำนวน 50 คน และคนปกติจำนวน 50 คน จำนวน 3 ครั้ง รวมทั้งสิ้น 300 ข้อมูล ข้อมูลของบุคคลนำไปกำหนดเป็นดาต้าเซต (data set) โดยนำข้อมูลมากลั่นกรองตรวจสอบความซ้ำซ้อนและแก้ไขข้อมูลที่ขาดหายไป เช่น รหัสคนปกติครั้งที่ 1 H0N01-01 ครั้งที่ 2 H0N01-02 หรือค่าการเปลี่ยนแปลงน้ำหนักและค่า BMI เป็นต้น ซึ่งบันทึกข้อมูลส่วนบุคคลที่เปลี่ยนแปลงในการตรวจแต่ละครั้งและเพิ่มความแม่นยำในการเรียนรู้ของเครื่องให้มากขึ้น

ซึ่งการบันทึกข้อมูลจะเก็บไว้ในโปรแกรมไมโครซอฟเอ็กเซล (Mirosoft Excel) โดยเป็นไฟล์นามสกุล .xlsx ซึ่งข้อมูลประกอบด้วย รหัส เพศ อายุ น้ำหนัก ส่วนสูง รอบเอว ความดันโลหิต ประวัติโรคเบาหวาน ค่าดัชนีมวลกาย ค่า Risk Scale และค่า Gray Scale ดังตารางที่ 1

ตารางที่ 1 ชื่อตัวแปรสำหรับใช้ประมวลผลในโปรแกรมการเรียนรู้ของเครื่อง

ชื่อตัวแปร	คำอธิบาย	ชนิดข้อมูล
ID	รหัสกลุ่มตัวอย่าง	Numberic
SEX	เพศ	Numberic
AGE	อายุ	Numberic
Wight	น้ำหนัก	Numberic
High	ส่วนสูง	Numberic
Waistline	รอบเอว	Numberic
Pressure	ความดันโลหิต (Yes = 1/NO = 0)	Numberic
Profile	ประวัติโรคเบาหวาน(Yes = 1/NO = 0)	Numberic
BMI	ค่าดัชนีมวลกาย (0-255)	Numberic
Risk Score	ค่าความเสี่ยงโรคเบาหวาน	Numberic
Gray Scale	ค่า Gray Scale	Numberic

ตอนที่ 1 การสังเคราะห์ข้อมูลระหว่างผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้าและข้อมูลคนปกติ

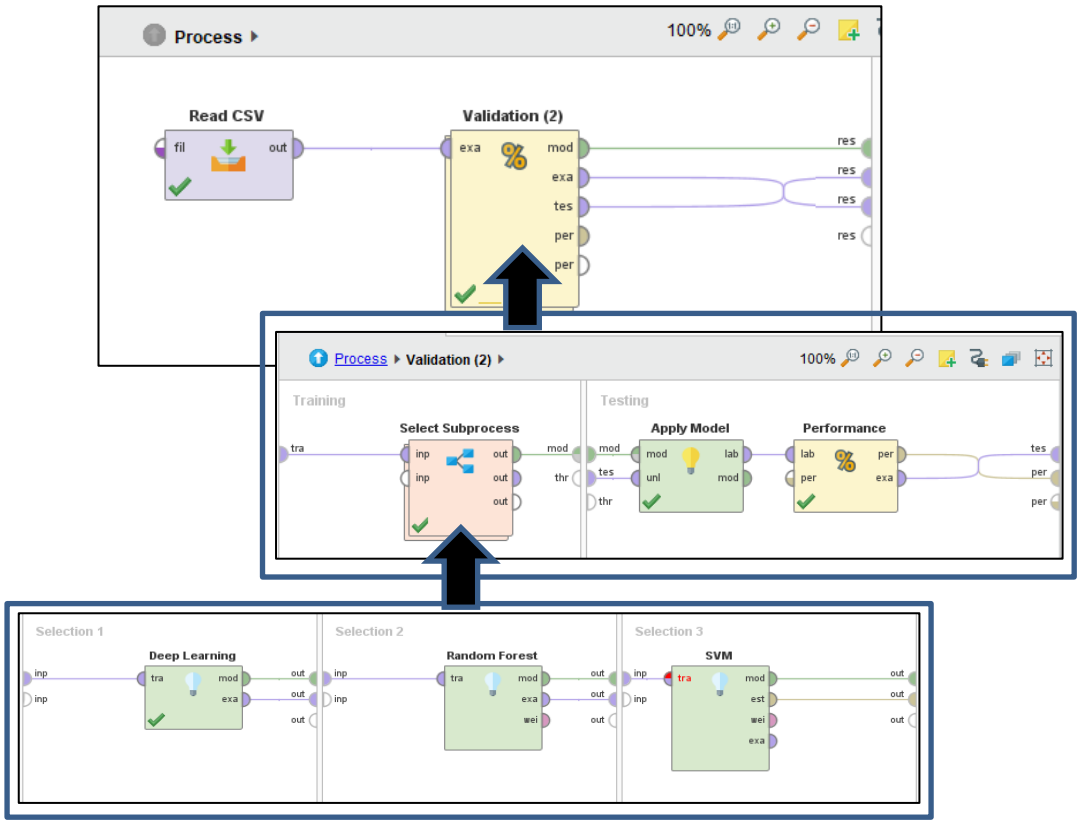
จากข้อมูลของคนปกติและผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้า ณ โรงพยาบาลส่งเสริมสุขภาพประจำตำบลท่าจีน จังหวัดสงขลา ผู้วิจัยได้จัดเก็บข้อมูลลงในโปรแกรมไมโครซอฟต์เอ็กเซล (Microsoft Excel) โดยได้กำหนดเป็นรหัสและจำนวนครั้งในการเก็บข้อมูล และได้คำนวณผลเพื่อหาค่า BMI ค่าประเมินความเสี่ยงในช่วงเวลานั้นและค่าประเมินความเสี่ยงเพื่อป้องกันโรค ดังตารางต่อไปนี้

ID	SEX	Age	Wight	Hight	Waistline	Pressure	Profile	BMI	Risk Score	Gray Scale
H0N48-01	ชาย	40	76	170	36	ไม่มี	ไม่มี	26.30	251.49	134
H0N49-01	ชาย	42	65	172	32	ไม่มี	ไม่มี	21.97	235.86	140.1
H0N50-01	ชาย	41	60	170	30	ไม่มี	ไม่มี	20.76	226.81	150.6
H1D01-01	หญิง	60	65	155	32	มี	มี	27.06	315.28	80.3
H1D02-01	หญิง	65	72	167	32	มี	มี	25.82	324.08	72.6
H1D03-01	ชาย	65	105	176	38	มี	มี	33.90	364.49	72
H1D04-01	หญิง	60	57	147	32	มี	ไม่มี	26.38	311.89	72.4

ภาพประกอบที่ 5 การบันทึกข้อมูลลงในโปรแกรมไมโครซอฟต์เอ็กเซล

ตอนที่ 2 การเปรียบเทียบโมเดลกระบวนการเรียนรู้ของเครื่องด้วยโปรแกรม RapidMiner Studio

ผู้วิจัยนำข้อมูลจากโปรแกรมไมโครซอฟเอ็กเซลมาเปลี่ยนบันทึกเป็นนามสกุล .csv เพื่อนำเข้าสู่โปรแกรม RapidMiner Studio ซึ่งเป็นโปรแกรมประเภทฟรีแวร์สำหรับการสร้างกระบวนการเรียนรู้ของเครื่อง โดยรับข้อมูลเข้าสู่โปรแกรมแบบ Read CSV และเลือกทำการเปรียบเทียบกับโมเดลจำนวน 3 โมเดล ได้แก่ Deep Learning, Random Forest และ Support Vector Machine จากข้อมูลดังกล่าวนำมาใช้สร้างแบบจำลองโดยหาค่าความเชื่อมั่น (Accuracy) ค่า Standard Deviation และค่า Runtime



ภาพประกอบที่ 6 การใช้โปรแกรม RapidMiner Studio เพื่อเปรียบเทียบโมเดลการเรียนรู้ของเครื่อง

ในการประมวลผลด้วยโปรแกรม RapidMiner Studio ใช้การประมวลผลจำนวน (Number of validation) 10 รอบ, อัตราส่วนตัวอย่าง (Simple ratio) = 1, เกณฑ์ (Criterion) = gain_ration, ค่าความลึกสูงสุด (Maximal deepth) = 20 ค่าความมั่นใจ (Confidence) = .25 โดยทำการเปรียบเทียบกับโมเดลจำนวน 3 วิธี ได้แก่ 1. Deep Learning 2.Random Forest และ 3.Support Vector Machine (SVM) โดยกำหนดสมมติฐาน (Assumption) ให้มีค่าความเชื่อมันมากกว่า 80% โดยมีการเรียนรู้โดยมีผู้สอน (Supervised learning)

ผลการวิจัย

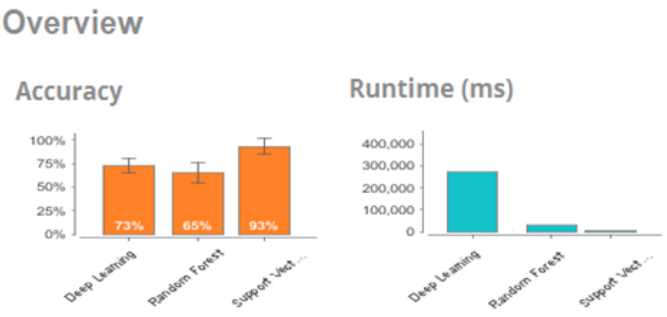
ตอนที่ 1 ผลการสังเคราะห์ข้อมูลระหว่างผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้าและข้อมูลคนปกติ

ผลการสังเคราะห์ในปัจจัยของข้อมูลเกี่ยวกับอายุที่สูงขึ้นและมีระดับความดันโลหิตสูง มีความเสี่ยงของในการเป็นโรคเบาหวานสูงกว่า ซึ่งผลการคำนวณค่า Risk score ทำให้มีความเสี่ยงโรคเบาหวานเพิ่มขึ้นโดยค่าผลรวมคำนวณที่มากกว่า 240 จะมีความเสี่ยงในระดับการคัดกรอง (Screening test) เบื้องต้น สอดคล้องกับมีความเสี่ยงโรคเบาหวานสอดคล้องกับ เกณฑ์การตัดสินคือคะแนนค่าความเสี่ยงโรคเบาหวาน (Risk score) ที่มากกว่า 240 (สำนักโรคไม่ติดต่อ กรมควบคุมโรค กระทรวงสาธารณสุข. 2560) แสดงว่าบุคคลนั้นมีความเสี่ยงต่อโรคเบาหวานสูง สมควรรับการตรวจหาเบาหวานโดยการตรวจเลือดและวัดระดับน้ำตาลในเลือดต่อไป

ส่วนการประเมินความเสี่ยงเพื่อป้องกันโรค โดยผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้าที่มีผลรวมตั้งแต่ 6-8 เป็นผู้มีความเสี่ยงโรคเบาหวานสูง โดยโอกาสเกิดเบาหวาน 1 ใน 7 ปี และผลรวมตั้งแต่ 9-11 เป็นผู้มีความเสี่ยงโรคเบาหวานสูง โดยโอกาสเกิดเบาหวาน 1 ใน 3 ปี

ตอนที่ 2 ผลการการเปรียบเทียบโมเดลกระบวนการเรียนรู้ของเครื่องด้วยโปรแกรม RapidMiner Studio

ผู้วิจัยได้ใช้โปรแกรมสำหรับการเรียนรู้ของเครื่องเพื่อเปรียบเทียบค่าความเชื่อมั่นด้วยโมเดลทั้ง 3 ประเภท ได้แก่ Deep Learning, Random Forest และ Support Vector Machine (SVM) โดยหาค่าความเชื่อมั่น (Accuracy) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) และระยะเวลาในการประมวลผลข้อมูล (Runtime)



Model	Accuracy	Standard Deviation	Runtime
Deep Learning	72.7%	+-7.4%	4 min 29 s
Random Forest	65.3%	+-10.9%	1 m 14 s
Support Vector Machine	93.3%	+-8.2%	46 s

ภาพประกอบที่ 7 ผลการเปรียบเทียบกระบวนการเรียนรู้ของเครื่องทั้ง 3 โมเดล

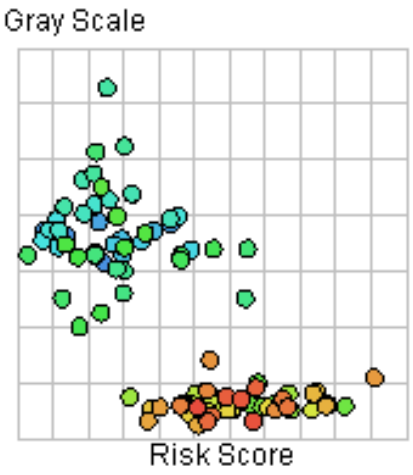
จากการเรียนรู้ของเครื่องด้วยโปรแกรม RapidMiner studio ผลการเปรียบเทียบทั้ง 3 โมเดลโดยเรียงตามลำดับของค่าความเชื่อมั่น พบว่าโมเดล Support Vector Machine (SVM) มีค่าความเชื่อมั่น (Accuracy) สูงสุด โดยคิดเป็นร้อยละ 93.3 ค่า S.D. คิดเป็นร้อยละ 8.2 และใช้ระยะเวลาในการประมวลผล 46 วินาที รองลงมาโมเดล Deep Learning มีค่าความเชื่อมั่น (Accuracy) โดยคิดเป็นร้อยละ 72.7 ค่า S.D. คิดเป็นร้อยละ 7.4 และใช้ระยะเวลาในการประมวลผล 4 นาที 29 วินาที และลำดับที่ 3 โมเดล Random Forest มีค่าความเชื่อมั่น (Accuracy) โดยคิดเป็นร้อยละ 65.3 ค่า S.D. คิดเป็นร้อยละ 10.9 และใช้ระยะเวลาในการประมวลผล 1 นาที 14 วินาที

Open in Turbo Prep Auto Model Filter (300 / 300 examples): all

Row No.	ID	Risk Score	prediction(Risk Score)	confidence(...	confidence(...	confidence(...
1	H1D21-03	345.71	332.79	0.025	0.001	0.006
2	H1D22-03	355.07	332.00	0.003	0.000	0.001
3	H1D23-03	285.42	285.42	0.001	0.003	0.001
4	H1D24-03	271.87	276.50	0.001	0.000	0.002
5	H1D25-03	300.21	325.72	0.001	0.000	0.000
6	H1D26-03	321.05	321.05	0.004	0.004	0.000
7	H1D27-03	298.09	289.34	0.011	0.000	0.008

ภาพประกอบที่ 8 ผลการทำนายค่า Risk score ด้วยโมเดล Support Vector Machine (SVM)

เมื่อได้แบบจำลองโมเดลทั้ง 3 โมเดลแล้ว ผู้วิจัยได้นำแบบจำลองดังกล่ามาทำนายผลค่า Risk score ด้วยโมเดล Support Vector Machine (SVM) ซึ่งมีค่าความเชื่อมั่นสูงสุดที่ร้อยละ 93.3 พบว่าค่าการทำนายสอดคล้องกับค่าการคำนวณประเมินความเสี่ยง (risk score) และผลการจำแนกได้ทำนายแบ่งกลุ่ม 2 กลุ่ม คือกลุ่มผู้มีความเสี่ยงโรคเบาหวานในระยะเริ่มต้นและกลุ่มคนปกติ ดังภาพที่ 5



ภาพประกอบที่ 9 การจำแนกกลุ่มตัวอย่างของคนปกติและผู้ป่วยโรคเบาหวานด้วย โมเดล SVM

ในการทดลองเพื่อหาผลการทดลองที่ดีที่สุดและจากการทบทวนงานวิจัยพบว่างานวิจัยส่วนใหญ่ใช้โมเดล Trees ในการพยากรณ์ข้อมูลความเสี่ยงโรคเบาหวานด้วยโปรแกรม RapidMiner โดยงานวิจัยของ Jiancho Han and other (2015) ได้ทำการศึกษาเรื่อง Diabetes Data Analysis and Prediction Model Discovery Using RapidMiner ทำการวิเคราะห์ข้อมูลเบื้องต้นของผู้ป่วยโรคเบาหวานในประเทศอินเดีย จำนวน 781 คน โดยกำหนดชนิดข้อมูล ได้แก่ ความเข้มข้นของกลูโคสในพลาสมาวัดระดับกลูโคสในช่องปาก (2 ชั่วโมง) การทดสอบระดับน้ำตาลในเลือด ความดันโลหิต ความหนาของผิวพับ เซรั่มอินซูลิน ค่า BMI ประวัติสายเลือดเบาหวาน อายุและระยะเวลาการเป็นโรคเบาหวานภายใน 5 ปี ผลการวิจัยพบว่าโปรแกรม RapidMiner และภาษา JAVA สำหรับเป็นส่วนติดต่อกับผู้ใช้ (GUI) การเลือกใช้โมเดล Trees โดยอัลกอริทึม ID3 ทำนายค่าความถูกต้องเท่ากับ 80% สอดคล้องกับ Sajida Perveen and other (2016) วิจัยเรื่อง Performance Analysis of Data Mining Classification Techniques to Predict Diabetes โดยการทำเหมืองข้อมูล (Data Mining) ใช้โปรแกรม WEKA ชุดข้อมูลที่ใช้ในการศึกษามาจากฐานข้อมูล CPCSSN (http://cpcssn.ca/) ฐานข้อมูล CPCSSN จำนวน 667,907 ชุดข้อมูล บันทึกเป็นระยะเวลาตั้งแต่ปี 2003 ถึง 2013 ในประชากรชาวแคนาดาที่ตรวจสอบความเสี่ยงโรคเบาหวานใน 8 ปี รวมถึงอายุเพศเลือดซิสโตลิก (diastolic) ความดันโลหิต, ไลโปโปรตีนความหนาแน่นสูง (HDL) ไตรกลีเซอไรด์ (TRG), ดัชนีมวลกาย (BMI) และจำนวนกลูโคส (FBG) โดยใช้อัลกอริทึมเปรียบเทียบจำนวน 3 อัลกอริทึม ได้แก่ AdaBoost, Bagging และ J48 ผลการวิจัยพบว่า โมเดล Trees ซึ่งใช้อัลกอริทึม J48 ได้ค่าความถูกต้องสูงสุดเพื่อจำแนกผู้ป่วยที่เป็นโรคเบาหวานโดยใช้ปัจจัยเสี่ยงโรคเบาหวาน การจำแนกประเภทนี้ทำในผู้ใหญ่สามกลุ่มที่แตกต่างกัน

สรุปและอภิปรายผล

จากข้อสรุปผู้วิจัยพบว่าในการใช้ข้อมูลผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้าในระยะเริ่มต้น เพื่อคำนวณค่าความเสี่ยงผู้ป่วยโรคเบาหวานในระยะแรกเริ่ม เพื่อป้องกันการเกิดแผลบริเวณเท้าและให้ใส่ใจสุขภาพเพื่อลดจำนวนผู้ป่วยโรคเบาหวานที่จะเกิดขึ้นในอนาคต ซึ่งปัจจัยส่วนบุคคลที่ส่งผลให้มีโอกาสเกิดความเสี่ยงในการเป็นโรคเบาหวานเพิ่มขึ้นคือ กลุ่มคนอายุมากกว่า 55 ปี และมีประวัติเกี่ยวกับความดันโลหิตสูง สำหรับการใช้โปรแกรม RapidMiner studio สำหรับการเรียนรู้ของเครื่องนั้น เพื่อเปรียบเทียบค่าความเชื่อมั่นในการคัดกรองความเสี่ยงโรคเบาหวาน ซึ่งโมเดลแบบ Support Vector Machine (SVM) มีค่าความเชื่อมั่นสูงสุดและใช้ระยะเวลาในการประมวลผลเร็วกว่าวิธีอื่น โดยการวิจัยในครั้งนี้กลุ่มตัวอย่างอยู่ในโรงพยาบาลส่งเสริมสุขภาพประจำตำบลท่าจีน จังหวัดสงขลา หากเพิ่มจำนวนกลุ่มตัวอย่างให้มากขึ้น และการเก็บข้อมูลเพิ่มจำนวนครั้งสำหรับการเก็บข้อมูลซ้ำ ๆ จะทำให้การเรียนรู้ของเครื่องมีความแม่นยำยิ่งขึ้น ซึ่งผลการวิจัยในครั้งนี้สามารถนำวิธีดังกล่าวเป็นแนวทางการพัฒนาโปรแกรมสำหรับคัดกรองผู้ป่วยโรคเบาหวานต่อไป ในการวิจัยครั้งนี้เป็นส่วนหนึ่งของงานวิจัยเรื่องการพัฒนาอุปกรณ์คัดกรองฝ่าเท้าผู้ป่วยโรคเบาหวานที่มีภาวะชาปลายเท้า

เอกสารอ้างอิง

ชัชลิต รัตรสาร. (2560). **สถานการณ์ปัจจุบันและความร่วมมือเพื่อปฏิรูปการดูแลรักษาโรคเบาหวานในประเทศไทย**. The Blueprint for Change Programe. NOVO Nordisk: กรุงเทพฯ.

ดุษณี ศุภวรรธนะกุลah. (2554). **การพัฒนาภูมิปัญญากะลามะพร้าวนวดเท้าชุดดุลพิธานเพื่อลดภาวะชาปลายเท้าในผู้ป่วยเบาหวาน**. หลักสูตรปรัชญาดุษฎีบัณฑิต. สาขาวิชาการจัดการเทคโนโลยี, มหาวิทยาลัยราชภัฏพระนคร.

นงคราญ คำวิชัย. (2559). **Practical Data Mining with RapidMiner Studio 7**. สาขาวิชาเทคโนโลยีสารสนเทศ, คณะวิทยาศาสตร์: มหาวิทยาลัยแม่โจ้.

บุญเสริม กิจศิริกุล. (2548). **ปัญญาประดิษฐ์.** เอกสารคำสอน**,** ภาควิชาวิศวกรรมคอมพิวเตอร์: จุฬาลงกรณ์มหาวิทยาลัย.

วีระศักดิ์ ศรินนภากร. (2550). **ภาวะน้ำตาลในเลือดสูง**. งานต่อมไร้ท่อ, กลุ่มงานอายุรศาสตร์: โรงพยาบาลราชวิถี.

สารัช ตันติกิตติ. (2559). **การประมวลผลภาพเพื่อบ่งชี้โรคไข้เลือดออกเดงกีจากเม็ดเลือดขาว.** สาขาเทคโนโลยีสารสนเทศ: มหาวิทยาลัยสยาม.

สมาคมโรคเบาหวานแห่งประเทศไทย ในพระราชูปถัมภ์สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี. (2554). **แนวทางเวชปฏิบัติสำหรับโรคเบาหวาน พ.ศ. 2555**. ศรีเมืองการพิมพ์จำกัด, กรุงเทพมหานคร.

สมาคมโรคเบาหวานแห่งประเทศไทย. (2554). **แนวทางเวชปฏิบัติสำหรับโรคเบาหวาน 2554**. สำนักงานหลักประกันสุขภาพแห่งชาติ, กรมการแพทย์: กระทรวงสาธารณสุข.

สำนักโรคไม่ติดต่อ กรมควบคุมโรค กระทรวงสาธารณสุข. (2560). **แผนยุทธศาสตร์การป้องกันและควบคุมโรคไม่ติดต่อระดับชาติ 5 ปี (พ.ศ. 2560- 2564)**. บริษัท อิโมชั่น อาร์ต จำกัด: กรุงเทพมหานคร.

อยุทธินี สิงหวินท์. (2557). นวัตกรรมใหม่ในการรักษาแผลในผู้ป่วยเบาหวาน. สืบค้นเมื่อ 17 ม.ค. 2562 จาก URL: http://www. phyathai. com/phyathai/service_center_ heartp2stemcell01.php

อารีรัตน์ วงศ์สุวรรณ. (2557). **การวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining โดยซอฟแวร์ Rapidminer Studio 6**. ออนไลน์ เข้าถึงได้จาก http://www.ubu.ac.th/blog/reerat-104. วันที่สืบค้น 2560 , สิงหาคม 18.

International Diabetes Federation.2015. IDF Diabetes Atlas Seventh Edition 2015. ออนไลน์ เข้าถึงได้จาก: www.diabetesatlas.org.

Jiancho Han, Juan C. Rodriguze and Mohsen Beheshti. (2015). **Diabetes Data Analysis and Prediction Model Discovery Using Rapidminer**. International Conference, Future Generation Communication and Networking.

Keesukphan P. Chanprasertyothin S, Ongphiphadhanakul B, Puavilai G. (2007). **The development and validation of diabetes risk score for high-risk Thai adults**, J med Assoc Thai 90: 149-54.

Parker, J. R. (2007). Algorithms for image processing and computer vision. **John Wiley & Sons.**

Sajida Perveen , Muhammad Shahbaz , Aziz Guergachi and Karim Keshavjee. (2016). **Performance Analysis of Data Mining Classification Techniques to Predict Diabetes**. Conference paper, Science Direct: Procedia Computer Science.

Powered by exteen blog. (2553). **พยาธิสภาพของโรคเบาหวาน.** [ออนไลน์]. สืบค้นได้จาก: http://phonphomuang.exteen.com/20100316/entry-3.วันที่สืบค้น 2558, สิงหาคม 28.

G.E. Hinton, N.Srivastava, A.Krizhevsky, I. Sutskever, and R.R. Salakhutdinov. 2012. Improving neural networks by preventing co-adaptation of feature detectors. ArXiv e-prints,July 2012.