

การหาประสิทธิภาพแบบจำลองการทำนายระดับความอ้วน

จากพฤติกรรมการกิน และสภาพร่างกาย

[ธัญชนก กิ่งปรุ]¹ [สุชาร์ตน์ กองฉลาด]² [ศรัณยพร ฉิมกุล]³ [น้ำทิพย์ บวรอารักษ์สกุล]⁴ [ณัฏพล ชูผล]⁵

และ[พิชญสินี กิจวัฒนาถาวร]⁶

[Thanchanok Kingpru]¹, [Sucharat Kongchalart]², [Saranyaporn Chimkun]³ [Namthip Bovornaraksakun]⁴

[Natchapol Choopol]⁵ and [Pitchayasini Kitwatthanathawon]⁶

สำนักวิทยาศาสตร์และศิลปดิจิทัล มหาวิทยาลัยเทคโนโลยีสุรนารี

[B6501297@g.sut.ac.th]¹, [SucharatKongchalart@gmail.com]² [saranyaporn4616@gmail.com]³

[B6530990@g.sut.ac.th]⁴ [sahapol1608@gmail.com]⁵ and [pichak@sut.ac.th]⁶

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ (1) สร้างแบบจำลองการทำนายทำนายระดับความอ้วนจากพฤติกรรมการกิน และสภาพร่างกายโดยใช้อัลกอริทึมเหมือนข้อมูล ได้แก่ โครงข่ายประสาทเทียม, ฟอเรสต์แบบสุ่ม, วิเคราะห์การถดถอยเชิงเส้น, นาอูฟเบย์, และต้นไม้ตัดสินใจ (2) เพื่อจะกลุ่มระดับของโรคอ้วน และ (3) เปรียบเทียบประสิทธิภาพของแบบจำลองด้วยวิธี 10-Fold Cross Validation โดยเครื่องมือที่ใช้ในการวิจัยคือโปรแกรม WEKA และ RapidMiner Studio และชุดข้อมูลเป็นข้อมูลสำหรับการประมาณระดับโรคอ้วนในบุคคลจากประเทศเม็กซิโก เปรู และโคลัมเบีย โดยพิจารณาจากพฤติกรรมการกินและสภาพร่างกายของพวกเขา ข้อมูลประกอบด้วยจำนวน 2111 ชุดข้อมูล 17 คุณลักษณะ ผลการวิจัยพบว่า...

คำสำคัญ: [แบบจำลอง] [เหมือนข้อมูล] [การทำนาย] [ระดับความอ้วน] [พฤติกรรมการกิน] [สภาพร่างกาย]

บทนำ

ประชากรที่ป่วยเป็นโรคอ้วนเพิ่มสูงขึ้นอย่างรวดเร็วทั่วโลก ซึ่งโรคอ้วน คือ ภาวะที่ร่างกายมีการสะสมไขมันมากเกินไปจนเกินกว่าปกติหรือมากเกินไปที่ร่างกายจะเผาผลาญ จึงสะสมพลังงานที่เหลือเอาไว้ในรูปของไขมันตามอวัยวะต่างๆ อาจมีความเสี่ยงต่อการเกิดปัญหาสุขภาพ และเป็นสาเหตุของการเกิดโรคเรื้อรังต่างๆ ตามมา สาเหตุที่ทำให้เกิดโรคอ้วน แบ่งออกเป็น ปัจจัยภายใน และปัจจัยภายนอก ซึ่งส่วนใหญ่แล้วผู้ที่เป็โรคอ้วน มักมีสาเหตุจากปัจจัยภายนอก เพราะมีพฤติกรรมมารับประทานที่ตามใจตนเอง จนทำให้รับประทานเกินความต้องการของร่างกาย จากผลการศึกษาของสหพันธ์โรคอ้วน World Obesity Federation ซึ่งเป็นองค์กรในสังกัดองค์การอนามัยโลกหรือ WHO ระบุว่า ปัจจุบันประชากรโลกที่มีภาวะน้ำหนักเกินหรือเป็นโรคอ้วนมีจำนวนประมาณ 2,600 ล้านคนหรือ 38% ของจำนวนประชากรโลกทั้งหมด 8,000 ล้านคน องค์การอนามัยโลก หรือ WHO พบว่า 1 ใน 3 ของประชากรที่เป็นโรคอ้วน หรือประมาณกว่า 600 ล้านคน มีอาการป่วยจากสาเหตุของโรคอ้วน เช่น เบาหวาน ความดันโลหิตสูง หลอดเลือดหัวใจ หลอดเลือดสมอง และรวมถึงมะเร็งบางชนิด ทางสหพันธ์โรคอ้วนได้คาดว่า ในปี 2035 ตัวเลขจะเพิ่มขึ้นเป็นมากกว่า 4,000 ล้านคน หรือ 51% ของจำนวนประชากรโลกทั้งหมด และสัดส่วนประชากรที่มีภาวะอ้วนรุนแรงจะเพิ่มจากจาก 1 คนต่อประชากร 7 คนในปัจจุบัน เป็น 1 คนต่อประชากร 4 คนในปี 2035 ประชากรโลกเกินครึ่งจะมีภาวะน้ำหนักเกินมาตรฐาน และโรคอ้วน หากรัฐบาลของประเทศต่าง ๆ ไม่แก้ไขปัญหานี้อย่างเร่งด่วน ผศ.พญ.ศานิต วิชานศวกุล อาจารย์ประจำหน่วยโภชนศาสตร์ ภาควิชาอายุรศาสตร์โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ กล่าวว่า อ้วนเป็นโรคที่ป้องกันได้ และควรได้รับการรักษาก่อนที่จะมีภาวะแทรกซ้อนจากโรคต่างๆ ไม่เฉพาะคนไข้ ครอบครัว ระบบสาธารณสุขต้องเข้ามาช่วยเหลือรักษาตั้งแต่เนิ่น จากปัญหาข้างต้น ผู้วิจัยจึงให้ความสำคัญในการนำทฤษฎีการทำเหมืองข้อมูล (Data Mining) ซึ่งเป็นกระบวนการวิเคราะห์ข้อมูลเพื่อค้นหารูปแบบ และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นๆ และในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท เช่น การพยากรณ์ผู้ป่วยเพื่อพยากรณ์การณการอุบัติของโรคต่างๆ มาวิเคราะห์เพื่อพยากรณ์หรือคาดการณ์โอกาสการเกิดโรคอ้วนโดยสร้างแบบจำลองสำหรับการคาดการณ์ระดับความอ้วนจากพฤติกรรมกรกิน และสภาพร่างกาย และเปรียบเทียบประสิทธิภาพของเทคนิคการทำเหมืองข้อมูล งานวิจัยนี้ได้ใช้เทคนิคการทำเหมืองข้อมูลที่หลากหลายเพื่อพยากรณ์ และเปรียบเทียบประสิทธิภาพของเทคนิคการทำเหมืองข้อมูล ซึ่งเทคนิคที่จะนำมาใช้ 5 เทคนิค ได้แก่ เทคนิคการเรียนรู้เชิงลึก (Deep Learning) คือวิธีการเรียนรู้แบบอัตโนมัติด้วยการ เลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาท (Neural Network) มาซ้อนกัน หลายชั้น (Layer) และทำการเรียนรู้ข้อมูลตัวอย่าง ซึ่งข้อมูล ดังกล่าวจะถูกนำไปใช้ในการตรวจจับรูปแบบ (Pattern) หรือจัด หมวดหมู่ข้อมูล (Classify the Data), เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เป็นเทคนิคหนึ่งของการทำเหมืองข้อมูลสำหรับการจำแนกข้อมูล (Classification Rules) โดยเป็นการนำข้อมูลมาสร้างแบบจำลองพยากรณ์เพื่อการทำนายหรือการจำแนกออกเป็นประเภทต่าง ๆ โดยมีโครงสร้างในลักษณะที่เป็นต้นไม้ (Sinsomboonthong, 2015) โครงสร้าง

ของต้นไม้ตัดสินใจประกอบด้วยโหนดราก (Root Node) กิ่ง (Branch) และโหนดใบ (Leaf Node), เทคนิคต้นไม้แบบสุ่ม (Random Forest) มีหลักการคือการเทรนโมเดลที่เหมือนกันหลายๆ ครั้ง (หลาย Instance) บนข้อมูลชุดเดียวกัน โดยแต่ละครั้งของการเทรนจะเลือกส่วนของข้อมูลที่เทรนไม่เหมือนกัน แล้วเอาการตัดสินใจของโมเดลเหล่านั้นมาโหวตกันว่า Class ไหนถูกเลือกมากที่สุด, เทคนิคนาอิวเบย์ (Naïve Bayes) หรือการเรียนรู้แบบเบย์ส์ เป็นวิธีการเรียนรู้ที่อาศัยความน่าจะเป็นตามทฤษฎีของเบย์ส์ (Bayes' Theorem) เป็นขั้นตอนวิธีในการจำแนกข้อมูลโดยการเรียนรู้ปัญหาที่เกิดขึ้นเพื่อนำมาสร้างเงื่อนไขการจำแนกข้อมูลใหม่ (Suwanco et al., 2017) เหมาะกับกรณีของเซตตัวอย่างที่มีจำนวนมาก และลักษณะของตัวอย่างไม่ขึ้นต่อกัน, การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression) คือการนำเอาข้อมูลหรือตัวแปรมาหาความสัมพันธ์กันโดยความสัมพันธ์ของข้อมูลจะออกมาในรูปแบบของการเรียงกันเป็นเส้นตรงหรือใกล้เคียง ผู้วิจัยได้นำเทคนิคที่กล่าวไปข้างต้นนี้ในการสร้างแบบจำลองเพื่อคาดการณ์ระดับความอ้วนจากพฤติกรรมกรกิน และสภาพร่างกาย

ทฤษฎีที่เกี่ยวข้อง

การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression) เกิดจากการรวมกันของคำ 2 คำได้แก่ Linear ที่แปลว่าเส้นตรง และ Regression ที่แปลว่าการถดถอย ซึ่งถ้าจะให้เข้าใจง่ายก็คือการนำเอาข้อมูลหรือตัวแปรมาหาความสัมพันธ์กันโดยความสัมพันธ์ของข้อมูลจะออกมาในรูปแบบของการเรียงกันเป็นเส้นตรงหรือใกล้เคียง ความสัมพันธ์ที่ได้กล่าวมานั้นคือการหาค่าสหสัมพันธ์หรือ Correlation นั่นเองโดยค่า Correlation หรือ “r” ยิ่งค่ามีความเข้าใกล้ 1 หรือ -1 กำลังจะบ่งบอกว่าค่าความสัมพันธ์ของข้อมูลนั้นยังมีความสัมพันธ์ในรูปแบบเส้นตรง (Linear) แต่ถ้าค่า “r” เข้าใกล้ 0 จะบ่งบอกว่าความสัมพันธ์ของข้อมูลเหล่านั้นไม่ได้อยู่ในรูปแบบเส้นตรง (Non-linear) Linear Regression สามารถแบ่งออกได้เป็น 2 ประเภทโดยจะแบ่งตามจำนวนตัวแปรอิสระที่จะมาเป็นตัวกำหนดค่าตัวแปรตามหรือค่า Y ได้แก่ Simple Linear Regression เป็นสมการเชิงเส้นที่มีตัวแปรอิสระหรือตัวแปร X มากำหนดตัวแปรตามหรือค่า Y แค่เพียงตัวแปรเดียวโดยสมการจะมีรูปแบบเหมือนกับที่ยกตัวอย่างไปข้างต้น เนื่องจากมีตัวแปรอิสระแค่เพียงตัวเดียวที่ส่งผลกระทบต่อค่า Y ทำให้สมการที่ได้ไม่มีความซับซ้อนในการแปรความหมาย และ Multiple Linear Regression เป็นสมการเชิงเส้นที่มีตัวแปรอิสระหรือตัวแปร X มากกว่า 1 ตัวมาเป็นตัวกำหนดตัวแปรตามหรือค่า Y ความหมายคือมีหลายปัจจัยที่ส่งผลกระทบต่อค่าที่เราให้ความสนใจทำให้การแปลความหมายของสมการมีความซับซ้อนมากยิ่งขึ้น

เทคนิคต้นไม้แบบสุ่ม (Random Forest) เป็นอัลกอริทึมประเภทหนึ่งของอัลกอริทึมต้นไม้ตัดสินใจที่มีลักษณะแบบ Unpruned หรือ Regression Trees ซึ่งถูกสร้างขึ้นโดยการสุ่มเลือกตัวอย่างข้อมูล. หลักการของ Random Forest คือการสร้าง model จาก Decision Tree หลายๆ model (ตั้งแต่ 10 model ถึงมากกว่า 1000 model) โดยแต่ละ model จะได้รับ data set ที่ไม่เหมือนกันซึ่งเป็น subset ของ data set ทั้งหมด. เมื่อทำการ prediction, แต่ละ Decision Tree ทำการ prediction แต่ละตัวและคำนวณผล prediction ด้วยการ

vote output ที่ถูกเลือกโดย Decision Tree มากที่สุด หรือหาค่าเฉลี่ยจาก output ของแต่ละ Decision Tree (ในกรณี regression)

เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เป็นอัลกอริทึมที่ใช้วิธีการแตกแขนงจากโหนดราก (Root Node) เป็นโหนดภายใน (Branch Node) แตกออกไปตามเงื่อนไขหรือข้อมูล จนไปสู่โหนดใบ (Leaf Node) เป็นแบบจำลองที่มีการเชื่อมโยงระหว่างสิ่งที่สนใจกับผลสรุปที่อาจเกิดขึ้นจากค่าของเหตุการณ์ (Jones, 2008) โหนดภายในของต้นไม้ตัดสินใจจะประกอบเป็นคุณลักษณะของข้อมูล ซึ่งเมื่อสอดคล้องกับข้อมูลใด ๆ ก็จะใช้คุณลักษณะนั้นเป็นตัวตัดสินใจว่าข้อมูลจะไปทิศทางใด โหนดภายในจะแตกกิ่งเป็นจำนวนเท่ากับจำนวนค่าของคุณลักษณะในโหนดภายใน และสุดท้ายคือ โหนดใบ เป็นกลุ่มผลลัพธ์ในการจำแนกประเภทข้อมูล ผลลัพธ์ที่ได้สามารถแปลงเป็นกฎ (Rule) ได้ การสร้างจะเริ่มต้นที่โหนดรากเป็นอันดับแรกก่อนจะดำเนินการพิจารณา โหนดใบและกิ่งก้านที่แตกแขนงต่อไป โดยต้องคำนวณหาข้อมูลที่เหมาะสมที่จะเป็น โหนดราก ซึ่งพิจารณาจากค่า Information Gain ที่มากที่สุด ที่ได้จากการคำนวณค่า Entropy เพื่อให้การจำแนกและแยกแยะข้อมูลให้อยู่ในกลุ่มเดียวกันมากที่สุด หลังจากที่ได้โหนดรากแล้วก็จะสร้าง Decision Tree ในลำดับต่อไป จนกระทั่งถึงและได้ Decision Tree ที่สมบูรณ์

เทคนิคการเรียนรู้เชิงลึก (Deep Learning) เป็นเทคนิคในกลุ่มโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) ที่มีโครงสร้างขนาดใหญ่ประกอบด้วยนิวรอนและชั้นซ่อนจำนวนมาก เป็นอัลกอริทึมที่ถูกสร้างขึ้นเพื่อการเรียนรู้ของเครื่องจักร แต่ละระดับ Hidden Layer ของการเรียนรู้เชิงลึกมีมากกว่า ANN ซึ่งแต่ละเลเยอร์จะเปรียบเสมือนประกอบด้วยเซลล์ประสาท (Neural) จำนวนมากที่มีหน้าที่ในการประมวลผล โดยเลเยอร์แรกสุดท้ายจะทำหน้าที่ในการรับข้อมูล (Input Layer) และส่งข้อมูลที่ประมวลผลเสร็จแล้วไปยังเลเยอร์สุดท้าย (Output Layer) การส่งข้อมูลแบบนี้มีข้อดีคือแต่ละเลเยอร์อาจทำให้มีค่าถ่วงน้ำหนัก (Weight) ค่าความเอนเอียงของข้อมูล (Bias) และวิธีการประมวลผลทางคณิตศาสตร์ (Activation Function) เป็นอิสระต่อกันถ้าป้อนข้อมูลเข้าไปให้กับโมเดลมากเท่าไร แต่ละเลเยอร์ก็จะสามารถสกัดคุณลักษณะที่มีความซับซ้อนมากขึ้นทำให้ระบบสามารถตัดสินใจได้ใกล้เคียงกับมนุษย์มากยิ่งขึ้น (สมศักดิ์ ศรีสุวรรณ และ สมัย ศรีสวญ, 2563)

เทคนิคนาอิวเบย์ (Naïve Bayes) เป็นเทคนิคที่ใช้ทฤษฎีความน่าจะเป็นตามกฎของเบย์ (Bayes' Theorem) (รุ่งโรจน์ บุญมา และ นิเวศ จิระวิชิตชัย, 2562) เพื่อหาสมมติฐานใดน่าจะถูกต้องที่สุด โดยใช้ความรู้ก่อนหน้า (Prior Knowledge) ได้แก่ ความน่าจะเป็นก่อนหน้าสำหรับสมมติฐานหนึ่งๆ ร่วมกับข้อมูล เช่น ความน่าจะเป็นที่สังเกตได้สำหรับสมมติฐานหนึ่งๆ เพื่อหาสมมติฐานที่ดีที่สุด การเรียนรู้แบบเบย์อาศัยหลักการของการคำนวณความน่าจะเป็นของแต่ละสมมติฐาน ในที่นี้คือคลาสเป้าหมายหรือผลลัพธ์การทำนายโดยการเรียนรู้แบบเบย์เป็นการเรียนรู้เพิ่มเติม เนื่องจากตัวอย่างใหม่ที่ได้มาถูกนำมาปรับเปลี่ยนการแจกแจงซึ่งมีผลต่อการเพิ่มหรือลดความน่าจะเป็นทำให้มีการเรียนรู้ที่เปลี่ยนไป วิธีการนี้ตัวแบบจะถูกปรับเปลี่ยนไปตามตัวอย่างใหม่ที่ได้โดยผนวก

กับความรู้เดิมที่มี ซึ่งการทำนายค่าคลาสเป้าหมายของตัวอย่างใช้ความน่าจะเป็นมากที่สุดของทุกสมมติฐานจาก ทฤษฎีของเบย์ เราสามารถคำนวณความน่าจะเป็นของสมมติฐานต่างๆ

งานวิจัยที่เกี่ยวข้อง

สุภาพร บรรดาศักดิ์, เบญญาภา ศรีสว่าง, และสุภาวดี ทองคำ (สุภาพร, เบญญาภา และสุภาวดี, 2559) ศึกษาถึงปัจจัยต่าง ๆ ที่ส่งผลต่อการเป็นโรคข้อเข่าเสื่อม โดยจากการศึกษาพบว่าอัลกอริทึม Naive byes มีประสิทธิภาพค่าความถูกต้อง เท่ากับ 92.1466 %, 99.7382 % อัลกอริทึม Sequential Minimal Optimization (SMO) มีประสิทธิภาพค่าความถูกต้องเท่ากับ 87.4346 %, 99.7382 % อัลกอริทึม Decision Tree (J48) มีประสิทธิภาพค่าความ ถูกต้องเท่ากับ 74.3455 %, 99.4764 % อัลกอริทึม Neural network มี ประสิทธิภาพค่าความถูกต้องเท่ากับ 89,0052 %, 98,9529 % จะเห็นได้ ว่าอัลกอริทึม Naive byes มี ประสิทธิภาพค่าความถูกต้องสูงที่สุด และมีการพยากรณ์ที่เที่ยงตรงสามารถนำไปใช้วิเคราะห์ความเสี่ยงคัดกรองผู้ ที่เสี่ยงต่อการเป็นโรคข้อเข่าเสื่อมได้เป็นอย่างดี

นพรัตน์ นนทศิริ, ราตรี มนัสศิลา, และกริช สมกันธา (นพรัตน์, ราตรี และกริช, 2564) ได้สร้าง แบบจำลองการจำแนกข้อมูลเพื่อวินิจฉัยความเสี่ยงการเป็นโรคเบาหวานจาก ผลการเปรียบเทียบพบว่า วิธีต้นไม้ ตัดสินใจให้ค่าประสิทธิภาพสูงสุด โดยมีค่าความถูกต้อง 93.73% วิธีนาอิวเบย์ค่าความถูกต้อง 88.92% วิธีความ ใกล้เคียงกันที่สุด และวิธีซัพพอร์ตเวกเตอร์แมชชีนค่าความถูกต้อง 86.97% และ 86.13% ตามลำดับ จะพบว่า วิธี ต้นไม้ตัดสินใจมีประสิทธิภาพในการสร้าง แบบจำลองมากที่สุดเมื่อเทียบกับวิธีที่ใช้เปรียบเทียบร่วมกัน เนื่องจาก เป็นวิธีที่ไม่มีการแจกแจงหรือไม่ใช้พารามิเตอร์ซึ่งไม่ได้ ขึ้นอยู่กับสมมติฐานการแจกแจงความน่าจะเป็น อีกทั้ง สามารถจัดการกับข้อมูลที่มีมิติสูงได้อย่างแม่นยำ เหมาะสมที่จะนำแบบ จำลองไปพัฒนาระบบจำแนกข้อมูลเพื่อ วินิจฉัยความเสี่ยงการเป็นโรคเบาหวาน ทางารแพทย์ในการวินิจฉัยความเสี่ยงการเป็นโรคเบาหวานต่อไป เพื่อ เป็นแนวทางในการสนับสนุนการตัดสินใจ

นงเยาว์ ในอรุณ (นงเยาว์, 2564) ได้สร้างแบบจำลองการทำนายความเสี่ยงโรคหัวใจและหลอดเลือดโดย ใช้อัลกอริทึมเหมือนข้อมูล พบว่าแบบจำลองที่มีประสิทธิภาพการทำนายดีที่สุดคือ แบบจำลองโครงข่ายประสาท เทียมพร้อมการเลือกคุณสมบัติ มีค่าความถูกต้อง 99.29% และต่ำสุดคือ แบบจำลองต้นไม้ตัดสินใจ มีค่าความ ถูกต้อง 70.39%

สุรวัชร ศรีเปารยะ, สายชล สีนสมบูรณ์ทอง (สุรวัชร, สายชล, 2560) ได้เปรียบเทียบประสิทธิภาพของ วิธีการจำแนกกลุ่ม โดยเลือกใช้วิธีความใกล้เคียงกันมากที่สุด เพื่อวัดประสิทธิภาพการจำแนกกลุ่มโดยใช้ข้อมูล ผู้ป่วยโรคไตเรื้อรังของโรงพยาบาลอโพลโล ประเทศอินเดีย จากการเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่ม ผู้ป่วยโรคไตเรื้อรัง โดยเปรียบเทียบจากค่าความถูกต้องและค่าความ คลาดเคลื่อนกำลังสองเฉลี่ย วิธีการจำแนก

กลุ่มที่มีประสิทธิภาพการจำแนกที่ดีที่สุดคือ วิธีต้นไม้ตัดสินใจ ซึ่งให้ค่าความ ถูกต้อง คือ 100 % และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0059

เบญจภัก จงหมื่นไวย (เบญจภัก, 2558) ได้เปรียบเทียบปัจจัยข้อมูลผู้สูงอายุ (เพศ อายุ และโรคประจำตัว) ซึ่งโรคประจำตัว ผู้สูงอายุสามารถจัดกลุ่ม พบว่า แบบจำลอง Decision Tree ร้อยละค่าความถูกต้องของการทำนายเท่ากับ 8.56% และ 91.43% เป็นค่าจริง แบบจำลอง Naïve Bayes ซึ่งพบว่า ร้อยละค่าความถูกต้องของการทำนายเท่ากับ 10.27 % และ 89.72 % เป็นค่าจริง ดังนั้น การวัดประสิทธิภาพการทางานค่าทำนายของผู้สูงอายุพบผู้สูงอายุที่ไม่มีโรคประจำตัวมากกว่าผู้สูงอายุที่มีโรคประจำตัวถึงร้อยละ 70.20 % และตัวแบบที่เหมาะสมสำหรับผู้สูงอายุในกลุ่มนี้คือ Decision Tree

กฤตกนก ศรีพิมพ์สอ และกิตติพล วิแสง (กฤตกนก และกิตติพล, 2566) พัฒนาแบบจำลองสำหรับการพยากรณ์ผู้ป่วยโรคเบาหวานโดยใช้เทคนิคการทำเหมืองข้อมูล ผลการวิจัยพบว่าเทคนิคแรนดอมฟอเรส (Random Forest) ให้ค่าความถูกต้องในการทำนายผลเป็นโรค เบาหวานมากที่สุดที่อยู่ที่ 99.75% มีค่าความแม่นยำ (Precision) ที่ 98.50% ค่าความครบถ้วน (Recall) ได้ 98.50% ค่าวัดประสิทธิภาพโดยรวม (F-Measure) ได้ 98.50% และค่าเส้นกราฟ (ROC) ได้ 95.20% สามารถนำผลที่ได้จากงานวิจัยนี้ไปประยุกต์ใช้ในการประกอบการรักษาผู้ป่วยโรคเบาหวานต่อไปในอนาคต

ณัฐพล แสนคำ, ทิพวัลย์ แสนคำ, และธนากร ปุรารัมย์ (ณัฐพล, ทิพวัลย์ และธนากร, 2560) ได้พัฒนาระบบสนับสนุนทางการแพทย์สำหรับคัดกรองผู้ป่วยโรคไตเรื้อรังโดยใช้เทคนิคเหมืองข้อมูลเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับทำนายโรคไตเรื้อรัง สรุปได้ว่า เทคนิค Random Forest ที่ใช้ชุดข้อมูลที่มีการเพิ่ม (Oversampling Data) มีประสิทธิภาพค่าความถูกต้อง (Accuracy) สูงที่สุดจากทุก ตัวแบบที่ค่า 97.29% ค่าความแม่นยำตรง (Precision) ที่ 95,76% และค่าวัดประสิทธิภาพโดยรวม (F-Measure) เท่ากับ 97,44% และนำเทคนิคเหมืองข้อมูลนี้มาพัฒนาเป็นตัวแบบในการทำนายโรคไต ผลการประเมินประสิทธิภาพในการทำนายโรคไตของระบบพบว่าสามารถทำนายโรคไตของข้อมูลใหม่ ได้ถูกต้อง 95.71% ทั้งนี้ เทคนิคต่าง ๆ และตัวแบบที่ได้พัฒนาขึ้นจะสามารถนำไปต่อยอด เพื่อพัฒนาระบบสนับสนุนทางการแพทย์ที่มีประสิทธิภาพในอนาคต

	เทคนิคการทำเหมืองข้อมูลที่ใช้ในวิจัย											ข้อมูลที่ใช้ในงานวิจัย			
	เทคนิค ฟอ เรสต์ แบบ สุ่ม	เทค นิค นา อึฟ เบย์	เทคนิค วิเคราะห์การ ถดถอย เชิงเส้น	เทคนิค โลจิสติก	เทคนิค ต้นไม้ ตัดสินใจ	เทคนิค โครงข่าย ประสาท เทียม	เทคนิค ซัพ พอร์ต เวกเทอ ร์แม ชชีน	เทคนิค เพอร์ เซ็ปตรอ น แบบ หลายชั้น	Sequential Minimal Optimization (SMO)	เค- เนียร์ เรสเน เบอร์	วิธีฐาน กฎ	ข้อมูล ส่วนตัว	ข้อมูล เกี่ยวข้องกับโรค	พฤติกรรม การใช้ ชีวิต	พฤติกรรม การกิน
สุภาพร บรรดาศักดิ์ และคณะ (2559)		X			X	X			X			X		X	
นพรัตน์ นนท์ศิริ และคณะ (2564)		X			X		X					X	X	X	
นงเยาว์ ใน อรุณ (2564)	X	X			X	X				X					
เบญจภาคี จงหนื่นไวย (2558)		X			X							X	X		
กฤตกนก ศรีพิมพ์สอ และกิตติ พล วิแสง (2566)	X	X		X	X			X							
ณัฐพล แสนคำ และคณะ (2560)	X				X										
สุรวัชร ศรี เปารยะ และสายชล สินสมบูรณ์ ทอง (2560)		X		X	X	X	X				X	X	X		
งานวิจัยนี้	X	X	X		X	X						X	X	X	X

วัสดุ อุปกรณ์ และวิธีดำเนินงาน

.....

.....

.....

.....

ผลการดำเนินงาน

.....

.....

.....

.....

ตารางที่ [ที่]: [พิมพ์ชื่อตาราง]

[หัวตาราง]	[หัวตาราง]	[หัวตาราง]
[ข้อมูลตาราง]	[ข้อมูลตาราง]	[ข้อมูลตาราง]
[ข้อมูลตาราง]	[ข้อมูลตาราง]	[ข้อมูลตาราง]
[ข้อมูลตาราง]	[ข้อมูลตาราง]	[ข้อมูลตาราง]
[ข้อมูลตาราง]	[ข้อมูลตาราง]	[ข้อมูลตาราง]

ที่มา:ชื่อผู้แต่ง (ปีที่พิมพ์)(ถ้ามี)

.....

.....

.....

.....



รูปที่ [ที่]: [พิมพ์ชื่อภาพ]

ที่มา: ชื่อผู้แต่ง (ปีที่พิมพ์)(ถ้ามี)

.....

.....

.....

.....

สรุปและอภิปรายผล

.....

.....

.....

.....

ข้อเสนอแนะ/งานวิจัยในอนาคต

เอกสารอ้างอิง

- ณัฐพล แส่นคำ และคณะ. (2560). ระบบสนับสนุนทางการแพทย์สำหรับคัดกรองผู้ป่วยโรคไตเรื้อรังโดยใช้เทคนิคเหมืองข้อมูล. วารสารวิชาการโรงเรียนนายร้อยพระจุลจอมเกล้า, ปีที่ 15, 161-170
- กฤตกนก ศรีพิมพ์สอ และคณะ. (2566). การพยากรณ์โรคเบาหวานด้วยเทคนิคเหมืองข้อมูล. วารสารวิชาการ “การจัดการเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม”, ปีที่ 10 ฉบับที่ 1, 52-63
- เบญจกัท จงหมื่นไวย. (2558). การเปรียบเทียบปัจจัยโรคประจำตัวผู้สูงอายุโดยใช้อัลกอริทึมการจัดกลุ่ม J48 และ Naïve Bayes: กรณีศึกษาสาธารณสุขสุโขทัยกลางนครราชสีมา. โครงการวิทยานิพนธ์คอมพิวเตอร์และเทคโนโลยีสารสนเทศ, ปีที่ 1 ฉบับที่ 1, 43-51
- สุรวัชร ศรีเปารยะ และคณะ. (2560). การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่งในประเทศอินเดีย. วิทยาศาสตร์และเทคโนโลยี, ปีที่ 25 ฉบับที่ 5, 840-853
- นงเยาว์ ในอรุณ. (2564). การเปรียบเทียบประสิทธิภาพของแบบจำลองการทำนายความเสี่ยงโรคหัวใจและหลอดเลือดโดยใช้อัลกอริทึมเหมืองข้อมูล. ไม่ปรากฏชื่อวารสาร, 138-147
- นพรัตน์ นนท์ศิริ และคณะ. (2564). การจำแนกข้อมูลเพื่อวินิจฉัยความเสี่ยงการเป็นโรคเบาหวานโดยใช้เทคนิคเหมืองข้อมูล. วิชาการพระจอมเกล้าพระนครเหนือ, ปีที่ 33 ฉบับที่ 2, 538-547
- มหาวิทยาลัยวลัยลักษณ์. (2559). งานประชุมวิชาการระดับชาติทาง ด้านเทคโนโลยีสารสนเทศ ครั้งที่ 8. กระบี่

bedee-expert. (2566).โรคอ้วน สัญญาณอันตรายจุดเริ่มต้นของโรคร้ายอื่น ๆ.สืบค้นเมื่อวันที่ 9 มกราคม 2567, BeDeebyBDMS: <https://www.bedee.com/articles/gen-med/obesity>

ไทยรัฐ ออนไลน์. (2563).เมื่อ "โรคอ้วน" คุกคามคนทั้งโลก!สืบค้นเมื่อวันที่ 9 มกราคม 2567, ไทยรัฐ: <https://www.thairath.co.th/lifestyle/woman/health/1971096>

สุรีย์ ศีลาวัณษ์. (2565)."โรคอ้วน" กระทบเศรษฐกิจ 13.2% ของงบประมาณสาธารณสุขทั่วโลก.สืบค้นเมื่อวันที่ 9 มกราคม 2567, กรุงเทพธุรกิจ: <https://www.bangkokbiznews.com/social/991651>

Reuters. (2566).ภายในปี 2035 คนครึ่งโลกจะมีภาวะน้ำหนักเกิน-โรคอ้วน.สืบค้นเมื่อวันที่ 9 มกราคม 2567, PPTV Online: <https://www.pptvhd36.com/health/news/2924>

สำนักสื่อสารความเสี่ยงฯ กรมควบคุมโรค. (2566).กรมควบคุมโรค ผลักดันคนไทยใส่ใจสุขภาพ ปรับเปลี่ยน มุมมองลด “โรคอ้วน”.สืบค้นเมื่อวันที่ 9 มกราคม 2567, สำนักสื่อสารความเสี่ยงและพัฒนาพฤติกรรมสุขภาพ: <https://ddc.moph.go.th/brc/news.php?news=32470&deptcode=brc>

วงการแพทย์. (2563).ตายเพราะอ้วนมากกว่าที่คิด.สืบค้นเมื่อวันที่ 9 มกราคม 2567, วงการแพทย์พลัสมีเดีย: <https://www.wongkarnpat.com/viewya.php?id=405>