

Analysis of risk factors and severity of osteoarthritis from risk behaviors to osteoarthritis using data mining techniques

Benyapha Srisawang, Surasak Mungsing

Faculty of Information Technology, Sripatum University

Abstract

The purpose of this research was to study and analyze the data of those at risk of osteoarthritis and the severity of those suffering from osteoarthritis from behaviors at risk of developing osteoarthritis. Data were collected from a sample of 151 patients with osteoarthritis of the knee and a sample of 113 without osteoarthritis. The data used in the analysis were collected between March and June 2021 and were analyzed using four data mining techniques: Decision tree (J48), Naïve bayes, SMO and Neural network to find correlation between behavioral data and

knee pain levels in people with osteoarthritis. The results showed that the Naïve bayes algorithm was able to generate the most efficient models with an accuracy of 95.8115 %

Keywords: osteoarthritis, data mining techniques, data classification, severity prognosis

*Revised: 10 May 2022, Revise: 25 July 2022,
Accepted: 30 September 2022*

Correspondence: Surasak Mungkasing, Department of Computing, Faculty of Information Technology, Sripatum University, 2410/2 Phahonyothin Road, Chatuchat District, Bangkok 10900, E-mail: surasak.mu@spu.ac.th

การวิเคราะห์ปัจจัยเสี่ยงและระดับความรุนแรงของโรคข้อเข่าเสื่อมจากพฤติกรรมที่มีความเสี่ยงต่อการเกิดโรคข้อเข่าเสื่อมโดยใช้เทคนิคการทำเหมืองข้อมูล

เบญญาภา ศรีสว่าง, สุรศักดิ์ มั่งสิงห์

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม

บทคัดย่อ

ปัจจุบันมีผู้ป่วยที่ต้องการรับโลหิตเป็นจำนวนมาก ทั้งจากอุบัติเหตุหรือผู้ป่วยที่จำเป็นต้องรับโลหิตในการรักษา ประกอบกับสถานการณ์ระบาดของไวรัสโควิด 19 ทำให้จำนวนผู้บริจาคโลหิตลดลง ส่งผลให้โรงพยาบาลหลายแห่งมีโลหิตสำรองไม่เพียงพอต่อความต้องการของผู้ป่วย บทความนี้นำเสนอต้นแบบระบบค้นหาผู้ร่วมบริจาคโลหิต การทำงานของระบบแบ่งเป็น 2 ส่วน ส่วนแรกคือ โมบายแอปพลิเคชันสำหรับผู้ร่วมบริจาคโลหิต โดยผู้ร่วมบริจาคโลหิตจะลงทะเบียนผ่านแอปพลิเคชัน ซึ่งจะส่งข้อมูลตำแหน่งพิกัดปัจจุบันที่ผู้ร่วมบริจาคโลหิตโดยใช้อุปกรณ์ระบุตำแหน่ง (GPS) ไปยังฐานข้อมูลผ่านเว็บเซอร์วิส (Web services API) เพื่อที่สามารถให้สาขาบริการโลหิตแห่งชาติสามารถค้นหาผู้ร่วมบริจาคที่อยู่ใกล้กับสาขาบริการโลหิตแห่งชาติหรือศูนย์บริการโลหิตแห่งชาติได้ และระบบจะแสดงข้อความคำขอบริจาคโลหิตจากสถานที่ที่ต้องการโลหิตไปถึงผู้ร่วมบริจาค ส่วนที่ 2 คือ เว็บแอปพลิเคชันของผู้ดูแลระบบของสาขาบริการโลหิตแห่งชาติหรือศูนย์บริการโลหิตแห่งชาติ ซึ่งผู้ดูแลระบบของสาขาบริการโลหิตแห่งชาติจะลงทะเบียน

ข้อมูลเกี่ยวกับสาขาบริการโลหิตแห่งชาติเอาไว้ ผู้ดูแลระบบสามารถกรอกข้อมูลโลหิตที่ต้องการเพื่อแจ้งให้กับผู้ร่วมบริจาคโลหิตทางโมบายแอปพลิเคชันที่ใกล้ที่สุด โดยที่ระยะทางไม่เกิน 10 กิโลเมตรจากสาขาบริการโลหิตแห่งชาติ ซึ่งทางสาขาบริการโลหิตแห่งชาติสามารถแสดงรายการผู้ร่วมบริจาคโลหิตที่ตอบตกลงได้ได้ ทั้งนี้ระบบที่พัฒนาขึ้นเป็นเพียงต้นแบบเพื่อทดสอบแนวคิดทางเทคนิค (technical proof of concept) เท่านั้น หากนำไปใช้งานจริงต้องทำการวิเคราะห์และออกแบบตามความต้องการของศูนย์บริการโลหิตแห่งชาติ สภากาชาดไทย

คำสำคัญ: โรคข้อเข่าเสื่อม, เทคนิคการทำเหมืองข้อมูล, การจำแนกข้อมูล, การพยากรณ์ระดับความรุนแรง

วันที่รับต้นฉบับ : 10 พฤษภาคม 2565, วันที่แก้ไข: 25 กรกฎาคม 2565, วันที่ตอบรับ: 30 กันยายน 2565

บทนำ

ในปัจจุบันการใช้ชีวิตของคนในสังคมไม่ว่าจะเป็นการทำงานการใช้ชีวิตประจำวันล้วนมีความเสี่ยงต่อการเกิดโรคภัยไข้เจ็บมากมายโดยเฉพาะโรคข้อเข่าเสื่อมโรคนี้มักจะเกิดขึ้นกับวัยผู้สูงอายุซึ่งพฤติกรรมต่างๆในชีวิตประจำวันอาจทำให้เกิดโรคข้อเข่าเสื่อม อาทิเช่น การนั่งยองหรือพับเข่าเป็นเวลานาน การทำงานที่ยกของหนักและอยู่ในพื้นชัน หรือการเดินขึ้นลงจากบันไดบ่อยครั้งก็อาจทำให้เกิดภาวะความเสี่ยงที่จะเป็นโรคข้อเข่าเสื่อมหรือข้อเข่าอักเสบได้ด้วยเช่นกัน

ผู้วิจัยจึงได้ทำการศึกษาและหาวิธีที่สามารถวิเคราะห์ความเสี่ยงของโรคข้อเข่าเสื่อมที่มีประสิทธิภาพ เพื่อนำมาใช้วิเคราะห์ความเสี่ยงต่อการเป็นโรคข้อเข่าเสื่อม และระดับความรุนแรงของผู้ที่เป็นโรคข้อเข่าเสื่อม และศึกษาความสัมพันธ์จากข้อมูลพฤติกรรมต่างๆ

ที่ส่งผลต่อการเป็นโรคข้อเข่าเสื่อม เพื่อหลีกเลี่ยงต่อพฤติกรรมที่ทำให้เกิดความเสี่ยงของโรคข้อเข่าเสื่อม ข้อมูลที่ใช้ในการวิเคราะห์ได้จากการสอบถามอาการแรกเริ่มที่จะส่งผลต่อการเกิดความเสี่ยงโรคข้อเข่าเสื่อม, อายุ, เพศ, น้ำหนัก, ส่วนสูง, ปัจจัยเสี่ยงต่างๆ และพฤติกรรมที่ก่อให้เกิดความเสี่ยงโรคข้อเข่าเสื่อม ผู้วิจัยเลือกใช้เทคนิคการทำเหมืองข้อมูลในการวิเคราะห์ความเสี่ยงโรคข้อเข่าเสื่อม โดยจะเปรียบเทียบประสิทธิภาพความถูกต้องของอัลกอริทึม 4 อัลกอริทึมคือ Decision tree(J48), Naöve bayes, SMO และ Neural network เพื่อหาอัลกอริทึมที่สร้างแบบจำลองที่มีความน่าเชื่อถือจากชุดข้อมูลทดสอบ

วัตถุประสงค์

1. เพื่อศึกษาอาการและพฤติกรรมของการเกิดโรคข้อเข่าเสื่อม
2. เพื่อวิเคราะห์หาความสัมพันธ์ของข้อมูลพฤติกรรมต่างๆ ที่ส่งผลต่อการเป็นโรคข้อเข่าเสื่อม และความเสี่ยงของโรคข้อเข่าเสื่อมในระดับความรุนแรงต่างๆ

ผู้สนับสนุนประสานงาน: สุรศักดิ์ มั่งสิงห์, คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม 2410/2 ถนนพหลโยธิน เขตจตุจักร กรุงเทพฯ 10900 E-mail: surasak.mu@spu.ac.th

3. เพื่อจำแนกระดับความรุนแรงของผู้ป่วยเป็นโรคข้อเข่าเสื่อม จากข้อมูลพฤติกรรมต่างๆ ที่ส่งผลต่อการเป็นโรคข้อเข่าเสื่อม ด้วยเทคนิคการทำเหมืองข้อมูล

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

1. ข้อเข่าเสื่อม

ข้อเข่าเสื่อม หมายถึงโรคที่เกิดจากความเสื่อมของกระดูกอ่อนผิวข้อทั้งทางด้านรูปร่างโครงสร้างการทำงานของกระดูกข้อต่อและกระดูกบริเวณใกล้เคียงการเปลี่ยนแปลงที่เกิดขึ้นไม่สามารถกลับสู่สภาพเดิมและอาจมีความเสื่อมรุนแรงขึ้นตามลำดับจากการศึกษาเชิงวิเคราะห์ความสัมพันธ์มีวัตถุประสงค์เพื่อทำนายความเสี่ยงการเกิดโรคข้อเข่าเสื่อมของผู้ป่วยสรุปผลการวิจัยได้ว่าการวัดระดับความรุนแรงของข้อเข่าเสื่อมได้ทั้งหมด 4 ระยะ [7] คือ

ระยะที่ 0 เป็นระยะที่ไม่มีมีความรุนแรงต่อการเกิดโรคข้อเข่าเสื่อม
ระยะที่ 1 เป็นระยะที่มีความรุนแรงของโรคน้อย เริ่มจากมีอาการปวดเมื่อยเล็กน้อย หลังจากการทำงานหนัก เดินไหล และอาการจะค่อยๆ เพิ่มขึ้น แต่พอก้าวเดินไปนานๆ ลัก 10-15 ก้าว อาการก็จะเริ่มหายปวดหรือเรียกได้ว่าเป็นครั้งคราว หรือนานๆ ครั้ง

ระยะที่ 2 เป็นระยะที่มีความรุนแรงของโรคข้อเข่าเสื่อมปานกลาง อาการปวดจะมากขึ้นและบ่อยขึ้น จนทำให้ขึ้นลงบันไดลำบาก นั่งยองๆ หรือนั่งพับเพียบนานไม่ได้ เมื่อลุกต้องมีคนพยุง อาจกินเวลา 5-15 ปี

ระยะที่ 3 เป็นระยะที่มีความรุนแรงของโรคข้อเข่าเสื่อมมาก โดยมีอาการปวดเข่าอยู่ตลอดเวลา ไม่ว่าจะพับข้อ ใช้ข้อหรือเรียกได้ว่าทุกอย่างก้าวที่เคลื่อนไหว ต้องมีคนช่วยพยุงตลอดเวลาและไม่สามารถช่วยเหลือตัวเองได้

2. การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล [6] เป็นกระบวนการของการค้นพบองค์ความรู้ที่น่าสนใจจากจำนวนข้อมูลที่มีมากที่ถูกเก็บไว้ในฐานข้อมูล Data Warehouses หรือที่เก็บข้อมูลอื่นๆ ช่วยให้สามารถทำนายแนวโน้มของข้อมูลใหม่ที่จะเกิดขึ้นในอนาคตได้การเหมืองข้อมูลแบ่งได้ 4 ขั้นตอนหลักๆ ดังนี้

ขั้นตอนที่ 1 เตรียมข้อมูล (Data Preparation) ถ้าข้อมูลไม่อยู่ในรูปแบบที่ถูกต้องหรือเหมาะสมจะต้องมีการปรับข้อมูลให้อยู่ในรูปแบบเดียวกันและจัดข้อมูลที่ไม่ถูกต้องออกไป

ขั้นตอนที่ 2 ลดขนาดของข้อมูล (Data Reduction) : เนื่องจากถ้าข้อมูลมีปริมาณมากเกินไปการค้นหาโมเดลหรือแพทเทิร์นจากกลุ่มข้อมูลขนาดใหญ่ต้องใช้เวลาามากซึ่งถ้าลดจำนวนข้อมูลลงด้วยสัดส่วนที่ถูกต้องโมเดลที่ได้ยังคงเป็นเช่นเดิมจะช่วยเพิ่มประสิทธิภาพมากขึ้น

ขั้นตอนที่ 3 ค้นหาโมเดล (หรือความสัมพันธ์) จากข้อมูล (Data Modeling/Discovery) กระบวนการค้นหาโมเดลหรือความ

สัมพันธ์จะเริ่มจากข้อมูลเริ่มต้นจำนวนไม่มากนักจากนั้นผลที่ได้จากกระบวนการค้นหา (Learning Process/Method) ไปยืนยันกับข้อมูลทดสอบถ้าผลที่ได้ยังไม่น่าพอใจจะต้องปรับค่าพารามิเตอร์บางตัวของ Learning Method และเริ่มกระบวนการค้นหาใหม่กับข้อมูลจำนวนมากขึ้นจนกว่าผลที่ได้มีความถูกต้องอยู่ในระดับที่ยอมรับได้จึงจะจบกระบวนการค้นหา

ขั้นตอนที่ 4 ตรวจสอบและวิเคราะห์ผล (Solution Analyses) : โมเดลหรือความสัมพันธ์ที่หามาได้ในขั้นตอนที่ 3 จะต้องถูกมาทดสอบอัตราความผิดพลาดและวิเคราะห์ความซับซ้อนของรูปแบบโมเดลถ้าอัตราความผิดพลาดยังสูงเกินไปอาจจะต้องย้อนกลับไปขั้นตอนที่ 3 อีกครั้งเพื่อปรับปรุงโมเดลให้ถูกต้องยิ่งขึ้น

3. โปรแกรม Weka

โปรแกรม Weka เป็นซอฟต์แวร์ประเภทฟรีแวร์ซึ่งอยู่ภายใต้การควบคุมของ GPL License ถูกพัฒนาขึ้นบนพื้นฐานของภาษาจาวาโดยเน้นใช้งานทางด้านการเรียนรู้ด้วยเครื่อง (Machine Learning) และการทำเหมืองข้อมูล (Data Mining) มีโมดูลย่อยสำหรับใช้ในการจัดการข้อมูลและใช้ตัวแสดงผลผ่านทางหน้าจอในลักษณะ GUI (Graphic User Interface)

4. วิจัยที่เกี่ยวข้อง

รัตนาลิ กักตัมย์ [2] การวิจัยเชิงปฏิบัติการนี้มีวัตถุประสงค์เพื่อศึกษาปัญหาและพัฒนาการดูแลผู้สูงอายุที่มีอาการปวดข้อเข่าและข้อเข่าผิดในตอนเช้าจำนวน 30 คนในเขตพร.สบตบ้านบากโดยได้สัมภาษณ์เชิงลึกผู้ดูแลและประธานสม.ผู้ใหญ่บ้านและเจ้าหน้าที่สาธารณสุขผลการวิจัยพบว่าอาการปวดข้อเข่าของผู้สูงอายุเกิดจากการใช้งานผิดวิธีน้ำหนักเกินมาตรฐานและเจ้าหน้าที่ในหน่วยบริการขาดแนวทางในการดูแลดังนั้นจึงได้ประชุมระดมสมองหาแนวทางวางแผนการดูแลและพัฒนาคู่มือการอบรมสม. และแกนนำด้วยซึ่งดำเนินการโดยอบรมสม. แกนนำออกเยี่ยมบ้านให้ความรู้และมีคู่มือการบริหารข้อเข่าประเมินผลหลังปฏิบัติการ 1 เดือน พบว่าผู้สูงอายุสามารถบริหารข้อเข่าได้ด้วยตนเองและปรับวิถีชีวิตทำให้ข้อเข่าปวดลดลงทำให้เคลื่อนไหวข้อเข่าได้ดีขึ้น

กรกมลพนาพิทักษ์กุล [3] งานวิจัยนี้มีวัตถุประสงค์เพื่อช่วยในการพยากรณ์ความเสี่ยงที่จะเกิดโรคของแต่ละระดับช่วงอายุโดยใช้ Formal Concept Analysis สำหรับการจำแนกระดับช่วงอายุโดยมีขั้นตอนการทำงานดังนี้ 1) การเตรียมข้อมูลสถิติของโรค 10 อันดับที่เกิดขึ้นและระดับช่วงอายุ 2) การเลือกลักษณะเฉพาะโดยใช้ FCA และ 3) การจำแนกประเภทและการประเมินผลจากการทดลองพบว่าสามารถจำแนกได้เป็น 5 กลุ่มคือ 1) วัยทารกมีความเสี่ยงในการเกิดโรคปอดและโรคไข้เลือดออก 2) วัยเด็กมีความเสี่ยงในการเกิดโรคตาแดง 3) วัยรุ่นมีความเสี่ยงในการเกิดโรคช่องปากโรคไข้หวัดใหญ่และโรคไข้เลือดออก 4) วัยผู้ใหญ่มีความเสี่ยงในการเกิดโรคข้ออักเสบโรคเบาหวานโรคความดันโลหิตสูงและโรคข้อเสื่อมและ 5) วัยผู้สูงอายุมีความเสี่ยงในการเกิดโรคเบาหวาน

โรคความดันโลหิตสูงโรคเบาหวานและโรคไขมันในเลือดสูง ดังนั้นจากงานวิจัยนี้สามารถหาวิธีเพื่อป้องกันโรคที่เกิดขึ้นในแต่ละวัยต่อไปได้ วนาสนา กรุดไทย [4].งานวิจัยที่เสริมสร้างความเข้มแข็งในการมองโลก เป็นเสมือนฐานบุคลิกภาพของบุคคลในมิติของการส่งเสริมสุขภาพ และเป็นตัวกำหนดภาวะสุขภาพของบุคคลโดยมีความสอดคล้องกันว่า ความเข้มแข็งของการมองโลกมีความสัมพันธ์ทางบวกต่อสุขภาพ และคุณภาพชีวิตและมีความสัมพันธ์ทางลบกับความรู้สึกร้าย ผู้ที่มีความเข้มแข็งสูงจะมีวิธีเผชิญกับปัญหาเหมาะสมและมีการปรับตัว ที่ดีงานวิจัยนี้ศึกษาความสัมพันธ์ความรุนแรงของข้อเข่าเสื่อมและ คุณภาพชีวิตของผู้ป่วยข้อเข่าเสื่อมที่รักษาทางแพทย์แผนไทยว่ามี ลักษณะอย่างไรและมีการทำนายคุณภาพชีวิตพฤติกรรมของผู้ป่วย ข้อเข่าเสื่อมโดยใช้ระบบการตัดสินใจ data Mining จากกลุ่มตัวอย่าง 137 คนในโรงพยาบาลชุมชนจังหวัดเพชรบุรีโดยใช้เครื่องมือ แบบสอบถามและตรวจสอบความเชื่อมั่นของเครื่องมือด้วย สัมประสิทธิ์อัลฟาของครอนบาค ผลของงานวิจัยพบว่า เพศ อายุ ดัชนีมวลกาย จำนวนครั้งที่รักษา ความรุนแรงของข้อเข่าเสื่อมและ ความเข้าข้างในการมองโรคนี้สามารถทำนายคุณภาพชีวิตของ ผู้ป่วยข้อเข่าเสื่อมที่รักษาทางแพทย์แผนไทยอย่างมีนัยสำคัญ ทางสถิติ ($F=8.565, p<0.001$)

สุขภาพ อารีเอื้อ [5] โรคข้อเข่าเสื่อมพบมากในกลุ่มผู้สูงอายุ การปฏิบัติตัวที่ไม่ถูกต้องอาจเกิดอาการรุนแรงได้ภายใน 1-3 ปี การวิจัยเชิงสำรวจนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่สัมพันธ์และ ทำนายพฤติกรรมและจากข้อมูลอย่างง่ายเข้ากลุ่มทดลอง 30 ราย และกลุ่มควบคุม 30 ราย ซึ่งจับคู่ให้มีคุณลักษณะคล้ายคลึงกัน ในเรื่องเพศอายุการศึกษาพื้นฐานด้านสุขภาพเป็นต้นเครื่องมือที่ใช้ ประกอบด้วยแบบบันทึกข้อมูลทั่วไปของผู้สูงอายุข้อเข่าเสื่อมแบบ ประเมินความเจ็บปวดข้อเข่า NRS (numeric rating scale) แบบ วัดความพึงพอใจต่อตนเองด้านร่างกายและด้านจิตใจและแบบวัด ความพึงพอใจต่อโปรแกรมของผู้สูงอายุข้อเข่าเสื่อมผลการทดลอง พบว่าคะแนนเฉลี่ยความเจ็บปวดของข้อเข่าในผู้สูงอายุข้อเข่าเสื่อม หลัง ทดลองระหว่างกลุ่มทดลอง ($x = 5.9 \pm 1.9$) และหลังทดลอง ($x = 2.4 \pm 0.9$) แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($p=0.00$) คะแนนเฉลี่ยความเจ็บปวดของข้อเข่าในผู้สูงอายุข้อเข่าเสื่อมหลัง ทดลองระหว่างกลุ่มทดลอง ($x = 2.4 \pm 0.9$) และควบคุม ($x = 6.7 \pm 1.7$) แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ($p=0.00$) ความพึงพอใจต่อตนเองด้านร่างกายและด้านจิตใจก่อนและหลัง ทดลองไม่แตกต่างกันส่วนความพึงพอใจต่อโปรแกรมลดปวด อยู่ในระดับปานกลาง

วิธีการดำเนินการวิจัย

1. ศึกษาปัญหาและวิเคราะห์ข้อมูล

ศึกษาและรวบรวมสรุปงานวิจัย ที่เกี่ยวข้องกับการศึกษาการทำงาน ของเทคนิคเหมืองข้อมูลเกี่ยวกับข้อมูลทางการแพทย์ ได้ Decision tree(J48), Naïve bayes, SMO (Sequential Minimal Optimization) และ Neural network รวมทั้งวิธีการวิเคราะห์และ ประเมินผลศึกษาวิธีการใช้งานซอฟต์แวร์ที่ใช้ในการทำเหมืองข้อมูล และรวบรวมข้อมูลมาเก็บข้อมูลจากคนที่ เป็นโรคข้อเข่าเสื่อม และไม่เป็นโรคข้อเข่าเสื่อม

- เครื่องมือที่ใช้ในการวิเคราะห์วัดจากคนที่ทำแบบสอบถาม แบบเจาะกลุ่ม

- 1) แพทย์ พยาบาล
- 2) ผู้ป่วยที่เป็นโรคข้อเข่าเสื่อม
- 3) ผู้ป่วยที่ไม่เป็นโรคข้อเข่าเสื่อม
- 4) บุคคลทั่วไป

คุณภาพ ค่าที่คำนวณจากการหาค่า IOC (index of item objective congruence) โดยมีเงื่อนไข

- ให้คะแนน +1 ถ้าแน่ใจว่าข้อคำถามวัดได้ตรงตามวัตถุประสงค์
ให้คะแนน 0 ถ้าถ้าไม่แน่ใจว่าคำถามวัดได้ตรงตาม
ให้คะแนน -1 ถ้าเห็นว่าคำถามวัดได้ไม่ตรงตามวัตถุประสงค์
เกณฑ์

- (1) ข้อคำถามที่มีค่า IOC ตั้งแต่ 0.50-1.00 มีค่าความเที่ยงตรง ใช้ได้
- (2) ข้อคำถามที่มีค่า IOC ต่ำกว่า 0.50 ต้องปรับปรุง ยังใช้ไม่ได้

$$IOC = \frac{\sum R}{N}$$

เมื่อ IOC คือความสอดคล้องระหว่างวัตถุประสงค์กับแบบทดสอบ
R คือผลรวมของคะแนนจากผู้เชี่ยวชาญทั้งหมด
N คือจำนวนผู้เชี่ยวชาญ

ตารางที่ 1 การพิจารณาแบบสอบถามการวิเคราะห์ความเสี่ยง

จุดประสงค์เชิงพฤติกรรม	เนื้อหา	ระดับการพิจารณา		
		1	0	-1
เมื่อนำแบบสอบถามไปถามกลุ่มเป้าหมายสามารถบอกได้ว่ามีความเสี่ยงการเกิดโรคข้อเข่าเสื่อมได้ถูกต้องแม่นยำ	1. ข้อมูลทั่วไป (ตารางที่ 5) 2. ข้อมูลการเกิดโรค (ตารางที่ 4) 3. พฤติกรรมเสี่ยงการเกิดโรค (ตารางที่ 3)			

จากตารางที่ 1 มีการนำไปให้ผู้เชี่ยวชาญตามกลุ่มจำนวนกลุ่มละ 5 คนตรวจสอบความเที่ยงตรงและปรากฏผลดังตารางที่ 2

ตารางที่ 2 ผลการตรวจสอบความสอดคล้องของแบบสอบถาม

	คะแนนความคิดเห็นของผู้เชี่ยวชาญ					รวม	ค่า IOC
	คนที่ 1	คนที่ 2	คนที่ 3	คนที่ 4	คนที่ 5		
แบบ สอบถาม	1	1	0	1	1	4	0.80

จากตารางที่ 2 ผลการตรวจสอบความคิดเห็นปรากฏว่ามีทางด้านแพทย์เห็นด้วยกับความสอดคล้องของแบบสอบถาม

ความเสี่ยงข้อเข่าเสื่อมจำนวน 4 คนและไม่เห็นด้วยกับความสอดคล้องของข้อแบบสอบถามความเสี่ยงข้อเข่าเสื่อมจำนวน 1 คน ผลรวมของคะแนนพิจารณา ($\sum R$) จะได้เท่ากับ 4 ดังนั้นค่า IOC จึงหาได้จาก

$$IOC = \frac{4}{5} = 0.80$$

ค่า 0.80 แสดงถึงแบบสอบถามความเสี่ยงข้อเข่าเสื่อมข้อนี้มีความสอดคล้องกับวัตถุประสงค์ เนื่องจากมีค่า 0. ซึ่งใกล้เคียงค่า 1 ทั้งนี้ค่า IOC ที่ยอมรับไว้ว่า แบบสอบถามมีความเที่ยงตรง คือมีค่าตั้งแต่ 0.5 ขึ้นไป ถ้าหากมีค่าน้อยกว่า 0.5 ถือว่าข้อแบบสอบถามความเสี่ยงข้อเข่าเสื่อมข้อนั้นไม่มีความสอดคล้องกับวัตถุประสงค์เชิงพฤติกรรมจะต้องแบบสอบถามความเสี่ยงข้อเข่าเสื่อมข้อนั้นออกไปหรือทำการปรับปรุงแบบสอบถามความเสี่ยงข้อเข่าเสื่อมข้อนั้นใหม่

2. การเตรียมข้อมูล

การเก็บข้อมูลที่ใช้ในการวิเคราะห์ ระหว่างช่วงเดือน มีนาคม-มิถุนายน 2564 จำนวนข้อมูลทั้งหมด 264 คนเป็นข้อมูลจากกลุ่มตัวอย่างที่เป็นโรคข้อเข่าเสื่อมจำนวน 151 และกลุ่มตัวอย่างที่ไม่เป็นโรคข้อเข่าเสื่อมจำนวน 113 คน แบบสอบถามได้แบ่งกลุ่มผู้ตอบแบบสอบถามออกเป็น 2 และอีกกลุ่มเป็นกลุ่มบุคคลทั่วไปหรือคนที่ไม่เป็นโรคข้อเข่าเสื่อม

เมื่อเก็บข้อมูลจะนำมาวัดระดับความเจ็บปวดจากข้อมูลที่ได้ศึกษามาได้ 4 ระยะดังนี้

- ระยะที่ 1 ทำงานทุกอย่างได้ตามปกติ
- ระยะที่ 2 ทำงานหนักไม่ได้
- ระยะที่ 3 ทำกิจวัตรประจำวันได้
- ระยะที่ 4 เดินไม่ไหว

ตารางที่ 3 พฤติกรรมเสี่ยงการเกิดโรค

ลำดับ	ตัวแปร	ชื่อข้อมูล
1	Walk	เดิน
2	Walk up stair	เดินขึ้นบันได
4	Walk down stair	เดินลงบันได
5	Sleep	ขณะนอน
6	Stand up	ขณะลุกจากที่นอน
7	Stand	ยืน
8	Wake up	ตื่นนอน
9	Change stance	ขณะเปลี่ยนท่า
10	BentUp down	ขณะงอขา
13	Wearshoes	ขณะสวมรองเท้า
14	Takeoff shoes	ขณะถอดรองเท้า
15	Out of bed	ขณะออกเดิน
16	Sit	นั่งอยู่กับที่
17	Sit on toilet	ขณะลุกเข้าห้องน้ำ
18	Household work	ขณะทำงานบ้าน
19	Householdworkhard	ขณะทำงานบ้านหนัก

(ตัวอย่างแบบสอบถามและลักษณะคำตอบ แสดงในรูปที่ 1)

ตารางที่ 4 ข้อมูลการเกิดโรค

ลำดับ	ตัวแปร	ชื่อข้อมูล
1	FatFood	ทานอาหารไขมัน
2	SweetFood	ทานอาหารหวาน
4	SourFood	ทานอาหารเค็ม
5	Sound	เสียงในข้อเข่า
6	Workhard	ทำงานหนัก
7	Same position long time	ทำงานที่เดิมเป็นเวลานานๆ
8	Up and down	ขึ้นลงบันได
9	Exercise wrong	ออกกำลังกายบ่อย
10	Accident	เคยประสบอุบัติเหตุ
11	Knee pain	อาการปวดเข่า
12	Move difficult	ย้ายของหนัก
13	Difficulty walking	การเดินแบบยากลำบาก
14	Difficultychange sitting	ลุกนั่งลำบากเวลาเปลี่ยนท่านั่ง
15	Bent press	ใช้งานข้อเข่า
16	Impact weight Pregnancy	รับน้ำหนักจากการตั้งครรภ์

ตารางที่ 5 ข้อมูลทั่วไป

ลำดับ	ตัวแปร	ชื่อข้อมูล
1	Sex	เพศ
2	Age	อายุ
4	Weight	น้ำหนัก
5	Height	ส่วนสูง
6	BMI	ดัชนีมวลกาย
7	Job	อาชีพ
8	Exercise	การออกกำลังกาย
9	Smoke	การสูบบุหรี่

15.ประเมินการเจ็บปวดข้อเข่า

15.1 ขณะเดินบนพื้นราบ

☐ ไม่ปวด ☐ ปวดปานกลาง

☐ ปวดมาก ☐ ปวดมากที่สุด

☐ ปวดรุนแรงมาก

15.2 ขณะเดินขึ้นบันได

☐ ไม่ปวด ☐ ปวดปานกลาง

☐ ปวดมาก ☐ ปวดมากที่สุด

☐ ปวดรุนแรงมาก

15.3 ขณะเดินลงบันได

☐ ไม่ปวด ☐ ปวดปานกลาง

☐ ปวดมาก ☐ ปวดมากที่สุด

☐ ปวดรุนแรงมาก

15.4 ขณะนอนบนเตียงในต่อนกลางคืน

☐ ไม่ปวด ☐ ปวดปานกลาง

☐ ปวดมาก ☐ ปวดมากที่สุด

☐ ปวดรุนแรงมาก

15.5 ขณะลุกนั่ง

☐ ไม่ปวด ☐ ปวดปานกลาง

☐ ปวดมาก ☐ ปวดมากที่สุด

☐ ปวดรุนแรงมาก

15.6 ขณะยืน

รูปที่ 1 ตัวอย่างคำถามในแบบสอบถามอาการเจ็บข้อเข่า

นำข้อมูลที่ได้มาจัดรูปแบบเพิ่มข้อมูลให้อยู่ในรูปแบบที่เหมาะสมเพื่อใช้กับซอฟต์แวร์ WEKA จากนั้นดำเนินการทำ Data cleaning โดยการลบ Record ที่มีค่าที่มีความผิดปกติหรือไม่จำเป็นต้องใช้ ซึ่งเมื่อนำข้อมูลดังกล่าวไปวิเคราะห์อาจทำให้รูปแบบทางสถิติมีความคลาดเคลื่อนหรือผิดพลาดแล้วจัดการกับข้อมูลสูญหายโดยแทนที่ข้อมูลที่สูญหายด้วยค่าเฉลี่ยของเขตข้อมูลสำหรับข้อมูลที่เป็นตัวเลขและแทนด้วยฐานนิยมสำหรับข้อมูลเชิงจำแนกสำหรับ

ข้อมูลที่เกี่ยวข้องกับการแพทย์เช่นข้อมูลส่วนตัวข้อมูลการประเมินความเจ็บปวดข้อเข่าพฤติกรรมเสี่ยงต่อการเกิดข้อเข่าเสื่อมโดยเลือกใช้วิธีการตัดแถวที่ปรากฏ Missing value ซึ่งเป็นวิธีที่นิยมมากสำหรับเทคนิคการจำแนกข้อมูลเลือก attribute ที่ส่งผลต่อการวิเคราะห์ด้วยวิธีการ `filters.supervised.attribute.AttributeSelection` ในโปรแกรม WEKA ซึ่งใช้ในการวิเคราะห์ เพื่อหาตัวแปรที่สำคัญและไม่สำคัญ โดยจะกำจัดตัวแปรที่ไม่สำคัญทิ้ง และเก็บตัวแปรที่สำคัญไปทำการวิเคราะห์ ขั้นตอนการใช้งานในการวิเคราะห์ข้อมูลดำเนินการโดยการคลิกปุ่ม 'open file' ในแท็บ Preprocess แล้วเลือกชุดข้อมูลที่ต้องการจะวิเคราะห์ เพื่อนำข้อมูลไปวิเคราะห์ที่แท็บ 'classify' โดยคลิกที่ 'choose' เพื่อเลือกวิธีที่ต้องการวิเคราะห์ข้อมูลหลังจากกด 'start' โปรแกรมจะทำการทำนายและแสดงผลลัพธ์ในส่วน classifier output สามารถดูผลการทำนายได้ เช่น Correctly Classified Instances จะบอก % ความถูกต้องจากการทำนาย ในงานวิจัยนี้ผลการวิเคราะห์ทำให้ลด factor เหลือเฉพาะ factor ที่สำคัญ 15 factor

3. การหาโมเดล

งานวิจัยนี้ได้นำโมเดลของการจำแนกข้อมูลมา 4 โมเดลเพื่อหาโมเดลที่มีประสิทธิภาพที่สุดมากที่สุดโมเดลที่เลือกมา ได้แก่ Decision tree (J48), Naïve bayes, SMO และ neural network ทดสอบแบบจำลองโดยวิธี 10-fold cross-validation evaluation model ในการวิเคราะห์แบบจำลอง

4. การประเมินประสิทธิภาพของโมเดลเลือกโมเดล

จากโมเดลที่เลือกมาทำการทดลองหาค่าพารามิเตอร์ที่ดีที่สุดสำหรับแต่ละโมเดลโดยใช้โปรแกรม WEKA แสดงผลการประเมินประสิทธิภาพของโมเดลดังตารางที่ 6

ตารางที่ 6 ผลการประเมินประสิทธิภาพของโมเดล

Algorithm	Correct (%)	MAE	RMSE	Precision	Recall
SMO	93.4555	0.2576	0.3235	0.935	0.935
Naïve Bays	95.8115	0.0315	0.1688	0.96	0.958
Neural Network	93.1937	0.0381	0.1326	0.932	0.932
J48	84.0314	0.0934	0.2605	0.84	0.84

จากผลการเปรียบเทียบผลการวิเคราะห์โมเดลของชุดข้อมูลในตารางที่ 7 สรุปได้ว่า Naïve Bays เป็นอัลกอริทึมที่ให้ความถูกต้องมากที่สุด

สรุปและอภิปราย

ผลการศึกษาพบว่า การนำเทคนิคเหมืองข้อมูลเข้ามาช่วยในการแก้ปัญหาโดยใช้ Naöve Bays มาใช้ในการสร้างแบบจำลองเพื่อที่จะทำการวิเคราะห์ความเสี่ยงในการเกิดโรคข้อเข่าเสื่อมซึ่งระบบวิเคราะห์ความเสี่ยงในการเกิดโรคข้อเข่าเสื่อมเพื่อช่วยลดความเสี่ยงในการเกิดโรคข้อเข่าเสื่อมในประเทศไทยให้ลดน้อยลงและเพื่อศึกษาปัจจัยเสี่ยงด้านต่างๆ พฤติกรรมที่ส่งผลต่อการเกิดโรคข้อเข่าเสื่อม โดยให้เทคนิคการวิเคราะห์ในการจำแนกข้อมูลเข้ามาช่วยในการแก้ปัญหาซึ่งจะช่วยในการวิเคราะห์บุคคลนั้นมีความเสี่ยงในการเกิดโรคข้อเข่าเสื่อมและยังช่วยให้บุคคลทั่วไปสามารถเข้าถึงระบบวิเคราะห์ความเสี่ยงในการเกิดโรคข้อเข่าเสื่อมได้ง่ายมากยิ่งขึ้นจากผลการวัดประสิทธิภาพของระบบสรุปได้ว่าระบบการวิเคราะห์ความเสี่ยงในการเกิดโรคข้อเข่าเสื่อมใกล้เคียงกับความจริงโดยใช้โมเดล Naöve Bays มีประสิทธิภาพค่าความถูกต้องมีค่าเท่ากับ 95.8115% อยู่ในระดับที่น่าเชื่อถือได้ นอกจากนี้จากข้อมูลที่ใช้ในการวิเคราะห์โดยใช้หลักการความน่าจะเป็นเหมาะสมกว่าวิธีการอื่นดังนั้นข้อมูลพฤติกรรมของผู้คัดกรองมีผลต่อการวิเคราะห์และสนับสนุนโอกาสการทำนายความเสี่ยงของโรคข้อเข่าเสื่อม ซึ่งวิธีการที่ใช้ปัจจัยร่วม เช่น SMO และ Neural Network ก็ได้ผลดีเช่นเดียวกัน แต่ในข้อมูลชุดนี้วิธีที่เหมาะสมที่สุดได้จากการสร้างแบบจำลองด้วยวิธี Naöve Bays และเป็นวิธีที่มีการประมวลผลได้รวมเร็วกว่าอีก 2 วิธีอีกด้วย การวิเคราะห์ความเสี่ยงต่อโรคข้อเข่าเสื่อมนั้น เพศ อายุ ส่วนสูง น้ำหนัก ไม่เป็นปัจจัยหลักต่อการวิเคราะห์ความเสี่ยงเลย แต่กิจกรรมที่กระทำแล้วเกิดอาการปวด เป็นเหตุบ่งชี้ต่อความเสี่ยงของโรคข้อเข่าเสื่อมได้ เช่นการยืนหรือการเปลี่ยนท่าแล้วมีระดับความเจ็บปวดเกิดขึ้น แสดงว่าเสี่ยงต่อโรคข้อเข่าเสื่อมแล้ว

เอกสารอ้างอิง

- [1] นพรัตน์ พจนัจิราภรณ์. (2554). ระบบสนับสนุนทางการแพทย์เพื่อวิเคราะห์โรคหลอดเลือดแดงโป่งพองโดยใช้เทคนิคเหมืองข้อมูลกรณีศึกษาการรักษาหลอดเลือดแดงโป่งพองผ่านสายสวน. ภาควิชาการจัดการเทคโนโลยีสารสนเทศคณะเทคโนโลยีสารสนเทศมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
- [2] รัตนวลี ภักดีสมัย. (2554). การพัฒนาการดูแลผู้สูงอายุที่มีปัญหาปวดข้อเข่าของโรงพยาบาลส่งเสริมสุขภาพตำบลบ้านปากอ่าวเมืองจันทราจังหวัดร้อยเอ็ด. วารสารพยาบาลศาสตร์และสุขภาพ มหาวิทยาลัยขอนแก่น.
- [3] กรกมล พนาพิทักษ์กุล. (2558). การพยากรณ์ความเสี่ยงที่จะเกิดโรคของแต่ละระดับช่วงอายุโดย FCA กรณีศึกษาโรงพยาบาลส่งเสริมสุขภาพตำบลบ้านสมควรอำเภอกุพารณ์จังหวัดนครศรีธรรมราช. สาขาคณิตศาสตร์ประยุกต์และสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยีอุตสาหกรรมมหาวิทยาลัยสงขลานครินทร์วิทยาเขตสุราษฎร์ธานี.
- [4] วาสนา กรุดไทย. (2553). ปัจจัยที่ทำนายชีวิตผู้ที่มีข้อเข่าเสื่อมที่รักษาทางการแพทย์แผนไทย. พยาบาลศาสตรมหาบัณฑิต สาขาการพยาบาลเวชปฏิบัติชุมชนคณะพยาบาลศาสตร์ มหาวิทยาลัยธรรมศาสตร์.
- [5] สุภาพ อารีเอื้อ. (2553). ปัจจัยที่มีผลต่อพฤติกรรมการจัดการตนเองของผู้ที่เป็นโรคข้อเข่าเสื่อม. คณะแพทยศาสตร์ โรงพยาบาลรามาริธิมมหาวิทยาลัยมหิดลวารสารสาธารณสุขศาสตร์.
- [6] อดุลย์ ยิ้มงาม, "การทำเหมืองข้อมูล (Data Mining)", 2551 [เข้าถึงเมื่อ 15 สิงหาคม 2559] เข้าถึงได้จาก <http://compcenter.bu.ac.th/news-information/data-mining>
- [7] เววดาว ทวีชัย. (2543) พฤติกรรมการดูแลตนเองและความรุนแรงของโรคข้อเข่าเสื่อมในผู้สูงอายุ, วิทยานิพนธ์ปริญญาโทมหาบัณฑิต, มหาวิทยาลัยเชียงใหม่.