

กฎความสัมพันธ์เพื่อการรักษาผู้ป่วยความดันโลหิตสูง

Association Rules for Medical Treatment in Hypertension Patients

ชิษณุพงศ์ บุปผาเต^{1*} และ จารี ทองคำ²

Chisanupong Bubpatate^{1*} and Jaree Thongkam²

นิสิตสาขาเทคโนโลยีสารสนเทศ^{1*} และ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม²

Student at major of Information Technology^{1*} and Faculty of Informatics at MahaSarakhm University²

E-Mail: bboyart616@gmail.com, jaree.thongkam@gmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) เปรียบเทียบความสามารถในการสร้างกฎความสัมพันธ์ในการรักษาผู้ป่วยโรคความดันโลหิตสูง ด้วยเทคนิค Apriori ในแต่ละกระบวนการเตรียมข้อมูล ดังนี้ กระบวนการที่ 1 แยกข้อมูลออกเป็น 2 กลุ่มตามคลาสและสร้างกฎไม่คำนึงถึงคลาส กระบวนการที่ 2 แบ่งแยกข้อมูลออกเป็น 2 กลุ่มตามคลาสและทำการสร้างกฎโดยคำนึงถึงคลาส กระบวนการที่ 3 นำข้อมูลทั้ง 2 กลุ่ม มารวมกันและสร้างกฎไม่คำนึงถึงคลาส และ 2) เปรียบเทียบประสิทธิภาพของกฎความสัมพันธ์ในการรักษาผู้ป่วยโรคความดันโลหิตสูง ด้วยเทคนิค Apriori ในแต่ละกระบวนการเตรียมข้อมูล โดยมีกลุ่มเป้าหมายผู้ป่วยโรคความดันโลหิตสูง จำนวน 6,269 คน เครื่องมือที่ใช้ในงานวิจัยนี้ คือ โปรแกรม Mysql ใช้ในการตรวจสอบข้อมูล และโปรแกรม Weka ใช้ในการวิเคราะห์ประสิทธิภาพของเทคนิค Apriori

ผลการวิจัยพบว่า 1) ผลการเปรียบเทียบความสามารถในการสร้างกฎความสัมพันธ์ ในการรักษาผู้ป่วยโรคความดันสูง ด้วยเทคนิค Apriori ในแต่ละกระบวนการเตรียมข้อมูล พบว่า กระบวนการที่ 1 สามารถสร้างกฎความสัมพันธ์ได้มากที่สุด เฉลี่ยที่ 693.40 กฎ และ 2) ผลการเปรียบเทียบประสิทธิภาพของกฎความสัมพันธ์ ในการรักษาผู้ป่วยโรคความดันโลหิตสูงด้วยเทคนิค Apriori ในแต่ละกระบวนการเตรียมข้อมูล พบว่า กระบวนการที่ 1 มีค่าความเชื่อมั่นเฉลี่ยสูงที่สุดถึง 75.44 %

คำสำคัญ: โรคความดันโลหิตสูง, กฎความสัมพันธ์, เอโพรารี

Abstract

The purposes of this research ,were found that to 1) compare the ability to create relationship rules. In the treatment of hypertensive patients with Apriori technique in each data preparation process. Process 1 separated data into 2 groups according to class and created rules regardless of class. Process 2 separated data into 2 groups according to class and created rules by class consideration. Process 3 combined both groups of data and Create rules, not words to class, and 2) compare the effectiveness of the relationship rules In the treatment of hypertensive patients with Apriori technique in each data preparation process. The target group for patients with

hypertension is 6,269. The tools used in this study are Mysql program for data validation and Weak program for analysis of Apriori technique.

From the objectives of the research, it is found that 1) process 1 can create the most relationship rules, average at 693.40 rules, 2) when comparing the efficiency of the relationship rules In the treatment of hypertension patients with Apriori technique in each data preparation process, it was found that process 1 had the highest confidence level of 75.44%.

Keywords: Hypertension patients, Association Rule, Apriori

บทนำ

โรคความดันโลหิตสูง (Hypertension) [1] คือ โรคไม่ติดต่อ (NCD) ที่มีความสำคัญมากโดยจะตรวจพบได้จากการวัดความดันโลหิตสูงองค์การอนามัยโลกได้กำหนดไว้ตั้งแต่ปี 2542 ว่าผู้ใดก็ตามที่มีความดันมากกว่า 140/90 มิลลิเมตรปรอท เป็นโรคความดันโลหิตสูง [2] ปัจจุบันพบมากในผู้ใหญ่อายุตั้งแต่ 35 ถึง 59 ปีและพบได้สูงขึ้นในวัยผู้สูงอายุตั้งแต่ 60 ปีขึ้นไป ส่วนในวัยเด็กพบได้น้อยกว่าในวัยผู้ใหญ่ โดยปกติคนจะมีระดับความดันโลหิต 120/80 มิลลิเมตรปรอท จะมีความดันโลหิตที่จะคอยผลักดันเลือดไปเลี้ยงอวัยวะส่วนต่าง ๆ ซึ่งอัตราปกติหัวใจของคนนั้นจะเต้นอยู่ประมาณที่ น้อยกว่า 80 ครั้ง ความดันจะเพิ่มขณะที่หัวใจบีบตัวและลดลงขณะที่หัวใจคลายตัว การรักษาโรคความดันสูงจะใชยาในการช่วยลดความดัน เช่น Losartan , Enalapril , Amlodipine, Hydralazine บางครั้งแพทย์จะยาร่วมกันซึ่งเมื่อแพทย์ให้ยาแล้วจะสังเกตอาการตามระยะเวลาที่แพทย์กำหนดจะเห็นว่าถ้าแพทย์หากมีภูมิต้านทานในการรักษาจะทำให้ข้อมูลทาง NCD มีประสิทธิภาพมากขึ้น

ปัจจุบันกระบวนการในการทำ Data Mining อีกหนึ่งอย่างที่ได้ได้รับความนิยมมาก คือ เทคนิคกฎความสัมพันธ์ (Association Rules) ได้ถูกนำมาใช้ในการวิเคราะห์ทางการแพทย์ในการรอดชีวิตมากขึ้น [3] โดยเทคนิคนี้จะทำการค้นหาชุดรายการที่เกิดขึ้นบ่อย (frequent itemset) ในชุดข้อมูล เทคนิคกฎความสัมพันธ์ [4] สามารถจัดอยู่ในรูปแบบ $LHS \Rightarrow RHS$ โดยที่ ค่า LHS และ RHS คือ ค่าชุดของข้อมูลโดยชุดข้อมูลที่อยู่ด้านขวาคือ ชุดที่มีแนวโน้มที่จะเกิดขึ้นได้เมื่อชุดข้อมูลด้านซ้ายเกิดขึ้น [5] อยู่ในรูปแบบที่มีลักษณะเป็นเงื่อนไขที่เข้าใจง่าย [6] ได้มีนักวิจัยหลายท่านได้นำเอาเทคนิค Apriori มาใช้ในการหาความสัมพันธ์เช่น Khaleel, Pradhan และ Dash [7] ใช้เทคนิคการทำเหมืองข้อมูลมาใช้กับข้อมูลทางการแพทย์ในเรื่องจำแนกโรคเฉพาะถิ่นที่เกิดขึ้นบ่อยโดยใช้เทคนิค Apriori จำแนกโรคในท้องถิ่นให้เกิดโรคน้อยที่สุด โดยใช้ค่า support ในการแสดงประสิทธิภาพของเทคนิค Apriori ส่วน Altaf, Shahbaz และ Guergachi [8] ได้มีการพัฒนาความสัมพันธ์สารสนเทศด้านสุขภาพตั้งแต่ปี พ.ศ. 2548 ถึงปี พ.ศ. 2557 ได้มีการสำรวจว่าแทนที่จะมีประสิทธิภาพมากขึ้นถ้าใช้เทคนิค Apriori ที่เป็นชุดข้อมูลที่ซับซ้อนๆ เทคนิคการประยุกต์ใช้การทำข้อมูลกฎความสัมพันธ์สำหรับสารสนเทศสุขภาพ โดยใช้ค่า support และ confidence ในการแสดงประสิทธิภาพของเทคนิค

ดังนั้น ผู้วิจัยจึงได้นำเทคนิคกฎความสัมพันธ์ รวมถึง Apriori มาใช้ในการวิเคราะห์ ค้นหาแบบการรักษาโรคความดันสูงโดยวัดประสิทธิภาพด้วย minimum support และ confidence เพื่อให้ได้กฎความสัมพันธ์ที่มีความน่าเชื่อถือสูงและสามารถเป็นแนวทางในการรักษาโรคความดันโลหิตสูง

1. วัตถุประสงค์การวิจัย

1.1 เพื่อเปรียบเทียบความสามารถในการสร้างกฎความสัมพันธ์ ในการรักษาผู้ป่วยโรคความดันสูง ด้วยเทคนิค Apriori ในแต่ละกระบวนการเตรียมข้อมูล

1.2 เพื่อเปรียบเทียบประสิทธิภาพของกฎความสัมพันธ์ ในการรักษาผู้ป่วยโรคความดันสูง ด้วยเทคนิค Apriori ในแต่ละกระบวนการเตรียมข้อมูล

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 โรคความดันโลหิตสูง

โรคความดันโลหิตสูง [9] หมายถึง โรคที่ผู้ป่วยมีภาวะที่ความดันช่วงบนมีค่า ตั้งแต่ 130 มิลลิเมตรปรอท และความดันช่วงล่างมีค่า ตั้งแต่ 80 มิลลิเมตรปรอทขึ้นไป โดยมากผู้ป่วยจะมีความดันช่วงล่างสูง (Diastolic hypertension) โดยที่ความดันช่วงบนจะสูงหรือไม่ก็ได้ แต่บางรายอาจมีความดันช่วงบนสูงเพียงอย่างเดียว แต่มีค่าความดันช่วงล่างไม่สูงก็ได้เช่นกัน เรียกว่า “ความดันช่วงบนสูงเดี่ยว” (Isolated systolic hypertension – ISHT) ซึ่งนับว่าเป็นอันตรายไม่น้อยไปกว่าความดันช่วงล่างสูงและผู้ป่วยควรได้รับการรักษาอย่างจริงจัง โรคนี้มักไม่มีอาการส่วนใหญ่อาการของโรคความดันโลหิตสูง เป็นอาการที่มีผลข้างเคียงจากโรคหัวใจและจากโรคหลอดเลือดในสมอง หรือเป็นอาการจากโรคที่เป็นปัจจัยที่เสี่ยงจากการที่เป็นโรคเรื้อรังที่รุนแรง ถ้าไม่สามารถควบคุมโรคได้ แต่มักไม่มีอาการ แพทย์บางท่านจึงเรียกโรคความดันโลหิตสูงว่า “เพชฌฆาตเงียบ (Silent killer)” โรคความดันโลหิตสูงมีจำนวนผู้ป่วยเพิ่มขึ้นอย่างต่อเนื่อง

2.2 กฎความสัมพันธ์

กฎความสัมพันธ์ [10] คือ เป็นเทคนิคที่ใช้การค้นหาค่าความสัมพันธ์ของข้อมูลที่มีอยู่ เพื่อช่วยในการวิเคราะห์และตัดสินใจโดยค้นหาค่าความสัมพันธ์ระหว่างแอตทริบิวต์ในข้อมูล ในงานวิจัยนี้การสืบค้นกฎความสัมพันธ์เป็นการวิเคราะห์พฤติกรรมของมนุษย์ว่า อาการของผู้ป่วยที่เป็นโรคความดันโลหิตสูงนั้นมีการค้นพบได้ในหลายสาเหตุของอาการแทรกซ้อน

2.3 เทคนิค Apriori

เทคนิค Apriori [11] คือขั้นตอนวิธีที่ได้รับการยอมรับและเป็นที่ยอมรับในการค้นหากฎความสัมพันธ์ โดยเซตที่มีความถี่มากกว่าค่าสนับสนุนขั้นต่ำของเซตย่อย จะมีความถี่ที่มากด้วยและสามารถตัดเซตของรายการที่มีความถี่ต่ำออก ขั้นตอนวิธีในการค้นหาเซตรายการที่ปรากฏร่วมกันบ่อยและนำมาสร้างกฎความสัมพันธ์ Vijay, Anand และ Dinakaran [12] ได้ใช้เทคนิค Apriori นำมาใช้ในการงานองค์กรด้านอุตสาหกรรม เพื่อใช้หาข้อกำหนดที่เป็นคำถามที่พบบ่อยที่สุดซึ่งจะช่วยในการลดความขัดแย้งของผู้มีส่วนได้ส่วนเสียและได้รับการออกแบบมาเพื่อใช้ทำงานกับฐานข้อมูลที่มีการทำธุรกรรมโดยใช้ค่า support ร้อยละ 50 ตัวแปรในงานวิจัยนี้คือการวัดความต้องการของนักลงทุนผลของการทดลองของเข้าพบว่ามี accuracy ร้อยละ 79.83 โดยที่ใช้ maximum support ร้อยละ 22 ส่วน Karthiyayini และ Jayaprakash [13] ได้ใช้เทคนิค Apriori หาข้อมูลความเป็นไปได้ในการเกิดโรคเรื้อรังคำนวณจากอาการแต่ละอาการของโรคที่พิจารณาการเกิดโรคด้วยเทคนิค Apriori มีความแม่นยำในการคำนวณความเป็นไปได้ของการเกิดโรค ช่วยให้แพทย์สามารถนำไปใช้ในการตัดสินใจที่มีประสิทธิภาพได้ดี

วิธีดำเนินการวิจัย

งานวิจัยนี้ได้นำเอาขั้นตอนกฎความสัมพันธ์มาใช้ในการประยุกต์ใช้เทคนิคลงไปเพื่อพัฒนาประสิทธิภาพของข้อมูล มีขั้นตอนการดำเนินการดังต่อไปนี้

1. เครื่องมือการวิจัย

- 1.1 โปรแกรม Mysql ใช้ในการตรวจสอบข้อมูล
- 1.2 โปรแกรม Weak ใช้ในการวิเคราะห์ประสิทธิภาพของเทคนิค Apriori

2. กลุ่มเป้าหมาย

กลุ่มเป้าหมายของการศึกษานี้ ได้แก่ ข้อมูลผู้ป่วยที่เข้ามารับการรักษาโรคความดันโลหิตสูงนี้ได้รับข้อมูลจากโปรแกรม Hosxp ในโรงพยาบาลพยัคฆภูมิพิสัย อำเภอพยัคฆภูมิพิสัย จังหวัดมหาสารคาม เริ่มตั้งแต่เดือน มกราคม พ.ศ. 2555 จนถึงเดือน ตุลาคม พ.ศ.2560 นำข้อมูลที่ไม่สมบูรณ์ออกจากชุดข้อมูล ทำให้เหลือข้อมูลจำนวน 34,788 ระเบียบ และแปลงข้อมูลให้อยู่ในรูปแบบเพื่อใช้ในการวิเคราะห์ และทำการเลือกตัวแปรที่มีผลต่อการตัวแปรตาม ทำให้ตัวแปรจาก 51 ตัวแปรเหลือเพียง 17 ตัวแปร ประกอบด้วยชุดข้อมูลผู้ป่วยโรคความดันโลหิตสูงและชุดข้อมูลยาที่ใช้ในการรักษาซึ่งมีตัวแปรดังตารางที่ 1

ตารางที่ 1 รายการตัวแปรในงานวิจัย

ตัวแปร	คำอธิบายตัวแปร	ชนิดตัวแปร
Sex	เพศ	Nominal
Age	ระดับอายุผู้ป่วยตอนมารับยา	Nominal
Body Mass Index	ระดับดัชนีมวลกาย	Nominal
Medical treatment	สิทธิ์รักษา	Nominal
Smoking	การสูบบุหรี่	Nominal
Follow up	การติดตาม	Nominal
Occupation	อาชีพ	Nominal
Drink alcohol	การดื่มสุรา	Nominal
HYDRALAZINE	ได้/ไม่ได้ รับยา HYDRALAZINE	Number
ENALAPRIL	ได้/ไม่ได้ รับยา ENALAPRIL	Nominal
ATENOLOL	ได้/ไม่ได้ รับยา ATENOLOL	Nominal
LOSARTAN_POTASSIUM	ได้/ไม่ได้ รับยา LOSARTAN_POTASSIUM	Number
HYDROCHLOROTHIAZIDE	ได้/ไม่ได้ รับยา HYDROCHLOROTHIAZIDE	Nominal
LORAZEPAM	ได้/ไม่ได้ รับยา LORAZEPAM	Nominal
AMLODIPINE	ได้/ไม่ได้ รับยา AMLODIPINE	Nominal
SIMVASTATIN	ได้/ไม่ได้ รับยา SIMVASTATIN	Nominal
Classes	เกิน /ไม่เกิด ระดับมาตรฐาน ระดับความดัน	Nominal

3. ขั้นตอนการดำเนินการวิจัย

ขั้นตอนการดำเนินการวิจัยกระบวนการก่อนการสร้างแบบจำลองในงานวิจัยนี้สามารถแบ่งออกได้ดังนี้

3.1 ตรวจสอบข้อมูล ทำการตรวจสอบชุดข้อมูลที่มีความสมบูรณ์ที่สุดเก็บไว้เพื่อใช้ส่วนข้อมูลที่ไม่สมบูรณ์หรือข้อมูลซ้ำและทำการลบข้อมูลออก

3.2 นำข้อมูลที่ไม่สมบูรณ์ออกจากชุดข้อมูล หลังจากทำการตรวจสอบข้อมูลและลบข้อมูลที่มีความซ้ำซ้อนออกแล้ว ผู้วิจัยจะทำการตรวจสอบชุดข้อมูลที่เป็นค่าว่างทำการนำข้อมูลชุดข้อมูลที่ไม่สมบูรณ์ออกโดยการใช้ Visual Studio และ Microsoft Excel ทำการเขียนชุดคำสั่งเพื่อตรวจสอบชุดข้อมูลที่ไม่สมบูรณ์และทำการนำชุดข้อมูลที่ไม่สมบูรณ์ออกจากชุดข้อมูล

3.3 แปลงข้อมูลให้อยู่ในรูปแบบเพื่อใช้ในการวิเคราะห์ ผู้วิจัยได้นำข้อมูลที่มีความสมบูรณ์ที่สุดนำมาเข้าชุดคำสั่งที่เขียนโดยโปรแกรม Visual Studio เพื่อแปลงข้อมูลให้อยู่ในรูปแบบข้อมูลที่สามารถนำมาวิเคราะห์ใช้เทคนิค Apriori โดยใช้โปรแกรม Weka 3.9.3 ได้ เหลือชุดข้อมูลจะทำการวิเคราะห์ที่ทั้งสิ้นจำนวน 34788 ระเบียบ

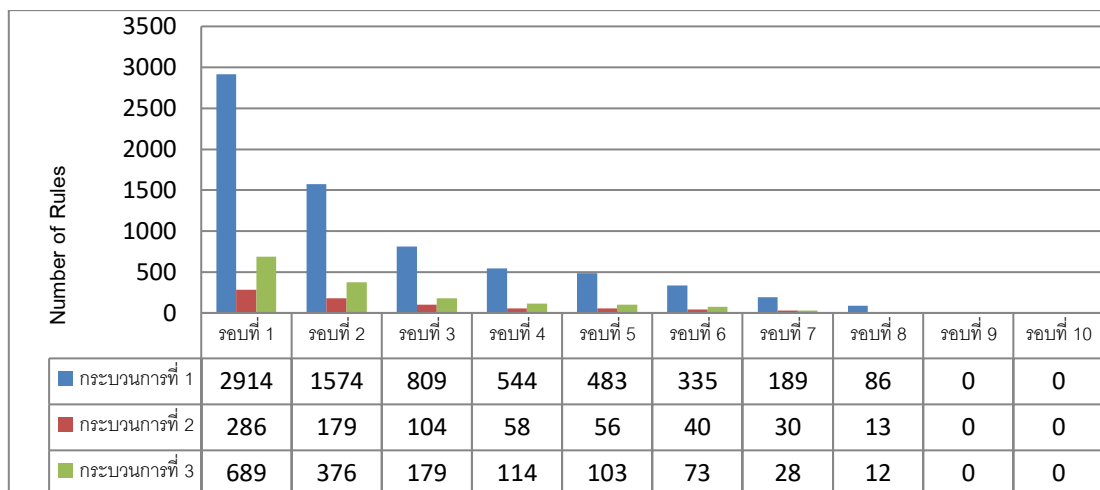
4. การทดลอง

ในงานวิจัยฉบับนี้ได้ทำการทดลองเพื่อคิดค้นขบวนการสร้างกฎความสัมพันธ์ที่มีประสิทธิภาพในการสร้างกฎความสัมพันธ์ ผู้วิจัยได้นำข้อมูลผู้ความดันโลหิตสูงที่มารักษาในคลินิกความดัน โดยกระบวนการแรก แยกข้อมูลออกเป็น 2 กลุ่ม ตามคลาสและสร้างกฎไม่คำนึงถึงคลาส กระบวนการที่ 2 แยกข้อมูลออกเป็น 2 กลุ่มตามคลาสและทำการสร้างกฎโดยคำนึงถึงคลาส (ตัวแปรตาม) กระบวนการที่ 3 นำข้อมูลรวมคลาสและสร้างกฎไม่คำนึงถึงแปรตาม ทดลองการสร้างกฎความสัมพันธ์ด้วยเทคนิค Apriori ซึ่งเป็นเทคนิคที่มีประสิทธิภาพในการสร้างกฎความสัมพันธ์ในงานวิจัยนี้ได้ทำการทดลอง 10 รอบ โดยรอบที่ 1 ผู้วิจัยได้มีการกำหนดค่า Minimum Support ระหว่าง และ UpperBoundMinSupport ในแต่ละรอบดังนี้ รอบที่ 1 ได้ใช้ LowerBoundMinSupport อยู่ที่ 50 UpperBoundMinSupport อยู่ที่ 55 รอบที่ 2 ได้ใช้ LowerBoundMinSupport อยู่ที่ 55.01 UpperBoundMinSupport อยู่ที่ 60 รอบที่ 3 ได้ใช้ LowerBoundMinSupport อยู่ที่ 60.01 UpperBoundMinSupport อยู่ที่ 65 รอบที่ 4 ได้ใช้ LowerBoundMinSupport อยู่ที่ 65.01 UpperBoundMinSupport อยู่ที่ 70 รอบที่ 5 ได้ใช้ LowerBoundMinSupport อยู่ที่ 71.01 UpperBoundMinSupport อยู่ที่ 75 รอบที่ 6 ได้ใช้ LowerBoundMinSupport อยู่ที่ 75.01 UpperBoundMinSupport อยู่ที่ 80 รอบที่ 7 ได้ใช้ LowerBoundMinSupport อยู่ที่ 80.01 UpperBoundMinSupport อยู่ที่ 85 รอบที่ 8 ได้ใช้ LowerBoundMinSupport อยู่ที่ 85.01 UpperBoundMinSupport อยู่ที่ 90 รอบที่ 9 ได้ใช้ LowerBoundMinSupport อยู่ที่ 90.01 UpperBoundMinSupport อยู่ที่ 95 รอบที่ 10 ได้ใช้ LowerBoundMinSupport อยู่ที่ 95.01 UpperBoundMinSupport อยู่ที่ 100 และทำการวัดด้วยจำนวนกฎความสัมพันธ์ และค่า confidence ของกฎในแต่ละรอบแล้วนำมาหาค่าเฉลี่ยของข้อมูลแต่ละชุด

ผลการวิจัย

1. ผลการเปรียบเทียบความสามารถในการสร้างกฎความสัมพันธ์ ในการรักษาผู้ป่วยโรคความดันสูง

ผู้วิจัยได้ดำเนินการสร้างจำนวนกฎความสัมพันธ์โดยเทคนิค Apriori ตามขั้นตอนการวิจัยโดยนำข้อมูลจากการศึกษา และทำการวิเคราะห์ ดังนั้นในส่วนนี้จะเป็นการนำเสนอจำนวนของกฎความสัมพันธ์ที่ได้จากกระบวนการที่แตกต่างกัน และแบ่งผลการทดลองตามช่วงของค่า support โดยสามารถแสดงผลการทดลอง แสดงดังภาพที่ 1

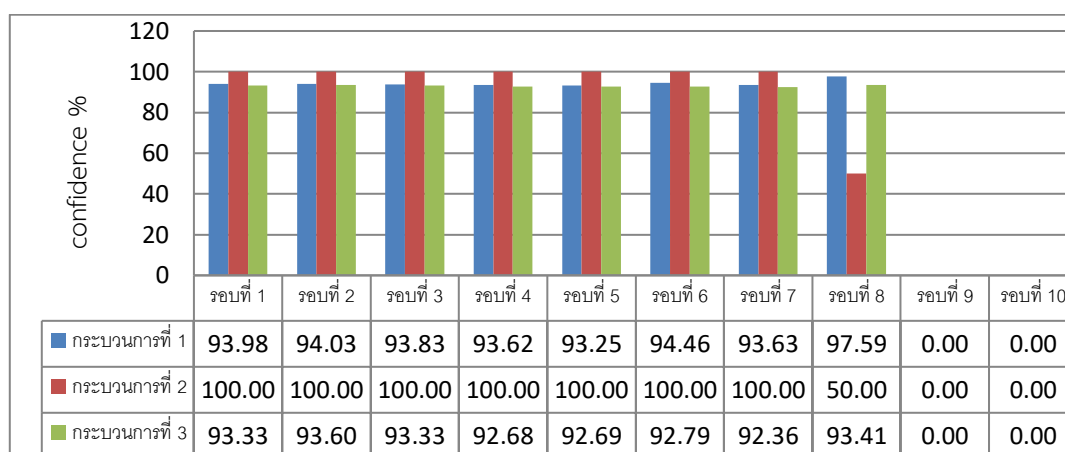


ภาพที่ 1 ระบบแสดงจำนวนกฎความสัมพันธ์ของ 3 กระบวนการในขั้นตอนการเตรียมชุดข้อมูล

จากภาพที่ 1 พบว่า กระบวนการที่ 1 ทำให้เทคนิค Apriori สามารถสร้างกฎได้จำนวนมากที่สุดเฉลี่ยถึง 693.40 กฎ ตามด้วยกระบวนการที่ 3 ซึ่งเป็นการสร้างกฎไม่เรียงถึงแปรตามแบบรวมคลาส ทำให้สามารถสร้างกฎได้จำนวนมีค่าเฉลี่ยที่ 157.40 กฎ และกระบวนการที่ 2 การสร้างกฎตามคลาสและทำการสร้างกฎโดยคำนึงถึงคลาสตัวแปรตาม มีค่าเฉลี่ยที่สุดที่ 76.60 กฎ

2. ผลการเปรียบเทียบประสิทธิภาพของกฎความสัมพันธ์ ในการรักษาผู้ป่วยโรคความดันสูง

ในงานวิจัยนี้ได้เปรียบเทียบประสิทธิภาพของกฎความสัมพันธ์ด้วยค่าความเชื่อมั่นของกฎความสัมพันธ์ที่สร้างด้วยเทคนิค Apriori ทั้งหมด 10 รอบ สามารถแสดงได้ดังภาพที่ 2



ภาพที่ 2 จำนวนค่าความเชื่อมั่นของทั้ง 3 กระบวนการ

จากภาพที่ 2 ผู้วิจัยได้ดำเนินการทดลองกฎความสัมพันธ์ที่พัฒนาขึ้นกับผู้ป่วยที่เข้ามารับการรักษาโรคความดันโลหิตสูง จำนวนผู้ป่วย 6,269 คน พบว่า ค่าความเชื่อมั่นในแต่ละกฎมีความแตกต่างกันตามจำนวนค่ากฎความสัมพันธ์ เมื่อเปรียบเทียบค่าความเชื่อมั่นในช่วงสนับสนุนรอบที่ 8 กระบวนการที่ 1 ทำให้เทคนิค Apriori สร้างกฎตามคลาสและสร้างกฎไม่คำนึงถึงคลาส มีค่าความเชื่อมั่นที่มากที่สุดที่รอบที่ 8 97.59 % ส่วนกระบวนการที่ 3 การสร้างกฎไม่คำนึงถึงแปรตามแบบรวมคลาส มีค่าความเชื่อมั่นที่ 93.41 % และกระบวนการที่ 2 การสร้างกฎตามคลาสและทำการสร้างกฎโดยคำนึงถึงคลาสตัวแปรตาม มีค่าความเชื่อมั่นน้อยที่สุดที่ 50 %

3. กฎความสัมพันธ์ของผู้ป่วยโรคความดันโลหิตสูง

จากการเปรียบเทียบประสิทธิภาพของการสร้างกฎความสัมพันธ์ด้วยค่าความเชื่อมั่นของเทคนิค Apriori สามารถสร้างกฎความสัมพันธ์ที่มีค่าความเชื่อมั่นที่ดีในรอบที่ 7 ซึ่งมีค่า support (80.01-85) และค่า confidence สูงสุดที่ (100%) จากตัวอย่างกฎความสัมพันธ์ กฎข้อที่ 1,2,11 และ ข้อที่ 12 แสดงดังตารางที่ 2 ดังนี้

ตารางที่ 2 กฎความสัมพันธ์ที่ถูกสร้างด้วยเทคนิค Apriori

No.	frequent itemset	Confidence
1	smoking_type_name=0 drinking_type_name=0 ==> class1=0	100%
2	clinic_visit_type_name=1 GEMFIBROZIL=0 ==> class1=0	100%
11	smoking_type_name=0 LOSARTAN_POTASSIUM=0 ==> class1=1	100%
12	clinic_visit_type_name=1 ATENOLOL=0 ==> class1=1	100%

จากตารางที่ 2 สามารถแปลความหมายของกฎความสัมพันธ์ที่คำนึงถึงคลาสผู้ป่วยที่มีความดันโลหิตสูง เป็นกฎข้อที่ 1 และกฎข้อที่ 2 ส่วนคลาสผู้ป่วยที่มีความดันโลหิตไม่สูง ซึ่งเป็นกฎข้อที่ 11 และกฎข้อที่ 12 ได้ดังนี้

กฎข้อที่ 1 ถ้า smoking_type_name เท่ากับ 0 และ drinking_type_name เท่ากับ 0 แล้วถ้า class1 เท่ากับ 0 หมายความว่าผู้ป่วยที่ไม่ได้มีการสูบบุหรี่และดื่มสุรา ได้มีความความดันโลหิตที่สูง

กฎข้อที่ 2 ถ้า clinic_visit_type_name เท่ากับ 1 และ GEMFIBROZIL เท่ากับ 0 แล้วถ้า class1 เท่ากับ 0 หมายความว่า ผู้ป่วยที่ไม่ได้มาตามการติดตามแต่ได้รับยา GEMFIBROZIL ซึ่งเป็นผู้ป่วยที่มีความดันโลหิตที่สูง

กฎข้อที่ 11 ถ้า smoking_type_name เท่ากับ 0 และLOSARTAN_POTASSIUM เท่ากับ 0 แล้วถ้า class1 เท่ากับ 1 หมายความว่าผู้ป่วยที่ไม่สูบบุหรี่ แต่ได้รับยา LOSARTAN_POTASSIUM ทำให้มีความดันโลหิตที่ปกติหลังอยากได้รับยาตัวนี้

กฎข้อที่ 12 clinic_visit_type_name เท่ากับ 1 และ ATENOLOL เท่ากับ 0 แล้วถ้า class1 เท่ากับ 1 หมายความว่าผู้ป่วยที่ไม่ได้มาตามการติดตาม แต่ได้รับยา ATENOLOL ได้มีความดันโลหิตที่ดันที่ปกติหลังได้รับยาตัวนี้

อภิปรายผลการวิจัย

1. การเปรียบเทียบความสามารถในการสร้างกฎความสัมพันธ์ ในการรักษาผู้ป่วยโรคความดันสูง ด้วยเทคนิค Apriori ในแต่ละกระบวนการเตรียมข้อมูล พบว่า กระบวนการที่ 1 ทำให้เทคนิค Apriori สามารถสร้างกฎได้จำนวนมากที่สุด เฉลี่ยถึง 693.40 กฎ ตามด้วยกระบวนการที่ 3 ซึ่งเป็นการสร้างกฎไม่คำนึงถึงแปรตามแบบรวมคลาส ทำให้สามารถสร้างกฎได้จำนวนมีค่าเฉลี่ยที่ 157.40 กฎ และกระบวนการที่ 2 การสร้างกฎตามคลาส และทำการสร้างกฎโดยคำนึงถึงคลาสตัวแปรตาม มีค่าเฉลี่ยที่สุดที่ 76.60 กฎ เหตุที่เป็นเช่นนี้อาจเนื่องมาจากการสร้างกฎความสัมพันธ์แบบไม่คำนึงถึงคลาส ทำให้เกิดกฎใหม่เพิ่มเติมจากกฎแบบคำนึงถึงคลาส ซึ่งสอดคล้องกับจารี ทอคำ และวาทีนิ สุขมาก, และภิรมพศ สุขมาก [14] ได้วิจัยเกี่ยวกับการสร้างกฎความสัมพันธ์ของผู้ป่วยมะเร็งต่อมลูกหมากกับการรักษา พบว่าการสร้างกฎความสัมพันธ์แบบคำนึงถึงคลาส เทคนิค Apriori ได้จำนวนกฎน้อยเช่นกัน

2. จากการเปรียบเทียบค่า support และ confidence ของกฎความสัมพันธ์ที่สร้างจากเทคนิค Apriori มีค่าความเชื่อมั่นสูงสุดโดยการแยกข้อมูลออกเป็น 2 กลุ่มตามคลาสและสร้างกฎไม่คำนึงถึงคลาส ซึ่งสอดคล้องกับ Altaf, Shahbaz และ Guergachi [8] ได้พัฒนากฎความสัมพันธ์ด้วยเทคนิค Apriori โดยใช้ค่า support และ confidence ในการวัดประสิทธิภาพของเทคนิค

ข้อเสนอแนะ

งานวิจัยนี้ได้มีการศึกษาการสร้างกฎความสัมพันธ์โดยใช้เทคนิค Apriori โดยเทคนิคที่ทดสอบแล้วได้นำมาสร้างกฎความสัมพันธ์ที่มีค่า support และ confidence ที่น่าเชื่อถือที่สุดในเอกสารฉบับนี้ หากมีการศึกษาเทคนิควิธีอื่น ๆ เข้ามาเปรียบเทียบเพิ่มเติม ทำให้มีค่า support และ confidence ที่แตกต่างกันในแต่ละเทคนิค โดยการนำข้อมูลจากโปรแกรม Hosxp ในโรงพยาบาลพศกภูมิพิสัย ย้อนหลัง 5 ปี โดยแยกข้อมูลเป็นรายเดือน ในงานวิจัยที่มีข้อมูลใกล้เคียงกันก็สามารถนำไปปรับใช้ให้เหมาะสมกับการสร้างกฎความสัมพันธ์ได้

เอกสารอ้างอิง

- [1] วิชัย เทียนถาวร. (2560). ความดันโลหิตสูง. สืบค้นจาก: https://www.matichon.co.th/columnists/news_557604
- [2] ไวยวรรณ ณะมัย, สมเกียรติ โพสสัย, สิทธิชัย อาชาสินดี, และสุริพร คนละเอียด. (2555). *คู่มือความรู้เพื่อจัดการภาวะความดันโลหิตสูงด้วยตัวเอง*. กรุงเทพฯ: สถาบันวิจัยและประเมินเทคโนโลยีทางการแพทย์.
- [3] Fan, Q., Zhu, C.-j., Xiao, J.-Y., Wang, B.-H., Yin, L., Xu, X.-L., & Rong, F. (2010). An application of Apriori algorithm In SEER breast cancer data. In *International Conference on Artificial Intelligence and Computational Intelligence*, Sanya, (pp. 114-116). doi: 10.1109/AICI.2010.263
- [4] Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. In *ACM SIGMOD International Conference on Management of Data*, (pp. 207-216). Retrieved from <https://doi.org/10.1145/170035.170072>
- [5] Creighton, C., & Hanash, S. (2003). Mining gene expression databases for association rules. *Bioinformatics*, 19(1), 79 - 86. Retrieved from <https://doi.org/10.1093/bioinformatics/19.1.79>
- [6] Han, J., Kamber, M., & Pei, J. (2006). *Data mining concepts and techniques* (3rd ed). Waltham, MA: Morgan Kaufmann.
- [7] Khaleel, M. A., Pradhan, S. K., & G.N.Dash. (2013). Finding Locally Frequent Diseases Using Modified Apriori Algorithm. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(10), 3792-3797.
- [8] Altaf, W., Shahbaz, M., & Guergachi, A. (2017). Applications of association rule mining in health informatics : a survey. *Artificial Intelligence Review*, 47(3), 313-340. Retrieved from <https://doi.org/10.1007/s10462-016-9483-9>
- [9] ความดันโลหิต. (2560). สืบค้นจาก <https://medthai.com/โรคความดันโลหิตสูง>
- [10] Karthikeyan, T., & Ravikumar, N. (2014). A Survey on Association Rule Mining. *International Journal of Advanced Research in Computer and Communication Engineering*, 3(1), 5223-5227.
- [11] Liusheng, H., Huaping, C., Xun, W., & Guoliang, C. (2000). A Fast Algorithm for Mining Association Rules. *Journal of Computer Science and Technology*, 15, 619-624.
- [12] Anand, R. V., & Dinakaran, M. (2017). Handling stakeholder conflict by agile requirement prioritization using Apriori technique. *Computers and Electrical Engineering*, 61, 126-136.
- [13] Karthiyayini, R. & Jayaprakash, J. (2015). Association Technique on Prediction of Chronic Diseases Using Apriori Algorithm. *International Journal of Innovative Research in Science Engineering and Technology*, 4(6), 255-259.
- [14] จาริ ทองคำ, วาทีนิ สุขมาก, และกัมพศ สุขมาก. (2561). การเปรียบเทียบประสิทธิภาพของเทคนิค Apriori และ FP-Growth ในการสร้างกฎความสัมพันธ์ของโรคมะเร็งต่อมลูกหมาก. *วารสารวิทยาการสารสนเทศและเทคโนโลยีประยุกต์*, 1(2), 103-111.