



วารสารพยาบาลตำรวจ
Journal of The Police Nurses

ปัจจัยที่มีผลต่อการทำนายการคลอดก่อนกำหนดในหญิงตั้งครรภ์

ในโรงพยาบาลตำรวจ

FACTORS AFFECTING THE PREDICTION OF GRAVID WOMEN
GIVING PRETERM BIRTHS AT POLICE GENERAL HOSPITAL

วิไลรัตน์ วิสวไพศาล¹

Wilairatn Wisvapaisan

บุญช่วย ศรีธรรมศักดิ์²

Boonchuay Srithammasak

สาธิษฐ์ นากกระแสน์²

Sathit Nakkrasae

¹นักศึกษาปริญญาโท หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาสถาปัตยกรรมศาสตร์ มหาวิทยาลัยรามคำแหง

²อาจารย์ที่ปรึกษาวิทยานิพนธ์ หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาสถาปัตยกรรมศาสตร์ มหาวิทยาลัยรามคำแหง
Ramkhamhaeng University

บทคัดย่อ

การวิจัยนี้เป็นการศึกษาแบบย้อนหลัง (retrospective study) มีวัตถุประสงค์เพื่อวิเคราะห์หาปัจจัยที่มีผลต่อการทำนายการคลอดก่อนกำหนดในหญิงตั้งครรภ์ในโรงพยาบาลตำรวจ โดยใช้เทคนิคการทำเหมืองข้อมูล วิธีการจำแนกประเภทข้อมูล และวิธีต้นไม้ตัดสินใจ โดยศึกษาผ่านโปรแกรมเวกา (WEKA) เพื่อพิจารณาหาปัจจัยต่าง ๆ ที่มีผลต่อการทำนายการเกิดการคลอดก่อนกำหนดของหญิงตั้งครรภ์ที่คลอดมีชีพในโรงพยาบาลตำรวจ ระหว่างวันที่ 1 มกราคม 2553 ถึง 31 ธันวาคม 2554 จำนวน 4,210 คน และปรับเพิ่มข้อมูลด้วยวิธีสุ่มเพิ่มข้อมูล (oversampling) ใช้เทคนิค SMOTE (Synthetic Minority Over-sampling Technique) โดยแบ่งข้อมูลเป็นกลุ่มข้อมูลสำหรับเรียนรู้ และกลุ่มข้อมูลทดสอบวิเคราะห์ และทดสอบด้วยวิธีการตรวจสอบไขว้แบบ 10 กลุ่ม (10 fold cross validation) โดยมีค่าความถูกต้องเท่ากับร้อยละ 90.37 ผลการวิจัยพบว่า ปัจจัยที่มีผลต่อการทำนายการคลอดก่อนกำหนดในหญิงตั้งครรภ์ในโรงพยาบาลตำรวจ คือ ภาวะแทรกซ้อนทางสูติศาสตร์ระหว่างการตั้งครรภ์และการคลอด ได้แก่ ภาวะถุงน้ำคร่ำแตกก่อนเจ็บครรภ์ การมีเลือดออกระหว่างตั้งครรภ์ ภาวะพิษแห่งครรภ์ การติดเชื้อในน้ำคร่ำ และการตั้งครรภ์แฝด

คำสำคัญ : คลอดก่อนกำหนด, การทำเหมืองข้อมูล, ต้นไม้ตัดสินใจ, ความไม่สมดุลของข้อมูล

Abstract

In this retrospective study, the researcher analyzed factors that affected the prediction of gravid women giving preterm births at Police General Hospital using the method of data mining in the course of which are classified data types and decision trees. The researcher studied Waikato Environment for Knowledge Analysis (WEKA) computer programming in preparation for examining factors affecting the prediction of preterm births given by gravid women who gave live births at Police General Hospital in the period between January 1, 2010 and December 31, 2011. The total number of live births was 4,210. Data were adjusted involving the method of oversampling through the use of the synthetic minority over-sampling technique. The data were divided into training data and testing data. Data were analyzed and tested using the method of 10-fold cross-validation. Findings showed that factors affecting the prediction of preterm births at Police General Hospital were various kinds of obstetric complications during pregnancy, premature rupture of membrane, bleeding per vagina, Pre-eclampsia, Amnioitis and Twins, at an accuracy of 90.37 percent.

Keywords : preterm birth, data mining, decision tree, imbalance data

ความเป็นมาและความสำคัญของปัญหา

การคลอดก่อนกำหนดเป็นปัญหาทางสาธารณสุขที่สำคัญ เพราะเป็นสาเหตุของการตายของทารกแรกเกิดที่พบบ่อย (ปัญญา พันธุ์บุรณะ, 2553) จากการศึกษาในประเทศสหรัฐอเมริกาพบว่า ทารกที่เสียชีวิตมากกว่าร้อยละ 50 อยู่ในกลุ่มของการคลอดก่อนกำหนดที่มีน้ำหนักตัวน้อยกว่า 1,500 กรัม (สายฝน ขวาลไพบุลย์, 2553) และหากทารกที่คลอดก่อนกำหนดไม่เสียชีวิตอาจเกิดภาวะทุพพลภาพ เช่น สมอ่งพิการ หรือความสามารถในการมองเห็นผิดปกติจากการได้รับออกซิเจนเพื่อช่วยหายใจเป็นเวลานาน (รุ่งตวรรณ ช้อยจ่อหอ, เยาวลักษณ์ เสรีเสถียร, เอมพร รตินธร, และ เอกชัย โควาวีสารัช, 2552) ทารกคลอดก่อนกำหนดที่มีภาวะทุพพลภาพ ต้องได้รับการดูแลอย่างใกล้ชิดจากบุคลากรทางการแพทย์ที่มีความชำนาญ ใช้อุปกรณ์ทางการแพทย์ที่มีราคาสูงและใช้เวลาอยู่ในโรงพยาบาลนานเฉลี่ยรายละ 2-3 เดือน มีค่ารักษาพยาบาลเฉลี่ย 175,000 บาทต่อราย และประเทศต้องใช้งบประมาณดูแลปีละไม่ต่ำกว่า 5,000 ล้านบาท (ณัฐญา เนตรหิน, 2555)

ในปัจจุบันมีการจัดเก็บข้อมูลต่าง ๆ ผ่านระบบคอมพิวเตอร์มากขึ้น ข้อมูลจำนวนมากถูกจัดเก็บอยู่ในระบบฐานข้อมูล แต่ไม่สามารถนำมาประยุกต์ใช้ได้อย่างมีประสิทธิภาพ จึงได้เกิดเทคโนโลยีในการวิเคราะห์ข้อมูลจากแหล่งเก็บข้อมูลขนาดใหญ่ขึ้น เรียกว่า “การทำเหมืองข้อมูล” ซึ่งเป็นกระบวนการที่กระทำกับข้อมูลจำนวนมาก เพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น ในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ด้านการตลาด เพื่อวิเคราะห์ยอดขาย ด้านวิทยาศาสตร์ ด้านเศรษฐกิจและสังคม (อดุลย์ ยิ้มงาม, ม.ป.ป.)

สำหรับด้านการแพทย์ มีผู้นำเทคนิคการทำเหมืองข้อมูล ขั้นตอนวิธีต้นไม้ตัดสินใจ มาใช้พยากรณ์และจำแนกผู้ป่วย เช่น การวินิจฉัยโรคระบบการหายใจ (ดิษฐพล มั่นธรรม, 2553) การพยากรณ์ความเสี่ยงการเกิดโรคมะเร็งปอด (รักถิ่น เหลาหา, 2553) เป็นต้น

การคลอดก่อนกำหนด เป็นภาวะแทรกซ้อนหนึ่งที่ไม่ทราบสาเหตุของการเกิดที่แน่ชัดโดยประมาณร้อยละ 25 มีสาเหตุมาจากการมารดาและทารก และร้อยละ 75 เป็นการเจ็บครรภ์ก่อนกำหนดที่เกิดขึ้นเอง (ประภัทร วานิชพงษ์พันธุ์, 2554) จากการศึกษาพบว่า ปัจจัยที่เกิดร่วมกับการคลอดก่อนกำหนดมีหลายอย่าง ได้แก่ ปัจจัยด้านมารดา เช่น อายุ น้ำหนัก การสูบบุหรี่ การทำงานหนัก ภาวะเครียด เป็นต้น ประวัติทางสูติศาสตร์ เช่น การแท้ง การคลอดก่อนกำหนด มดลูกผิดปกติ มีเลือดออกในระหว่างตั้งครรภ์ เป็นต้น ภาวะแทรกซ้อนในการตั้งครรภ์ปัจจุบัน เช่น การตั้งครรภ์แฝด ความดันโลหิตสูง เป็นต้น อย่างไรก็ตามมีมารดาจำนวนไม่น้อยที่คลอดก่อนกำหนดโดยไม่พบปัจจัยเสี่ยงใด ๆ ได้มีการศึกษาปัจจัยหรือสาเหตุของการคลอดก่อนกำหนด ด้วยวิธีทางสถิติต่าง ๆ เช่น การแจกแจงความถี่ การทดสอบความสัมพันธ์ เป็นต้น แต่ผลการศึกษายังไม่สามารถพบสาเหตุที่แน่ชัดได้ ผู้วิจัยจึงมีแนวคิดนำเทคนิคการทำเหมืองข้อมูล มาวิเคราะห์หาปัจจัยที่มีผลต่อการทำนายการเกิดการคลอดก่อนกำหนด เพื่อเป็นประโยชน์ในการวินิจฉัยและวางแผนป้องกันหญิงตั้งครรภ์รายอื่นต่อไป

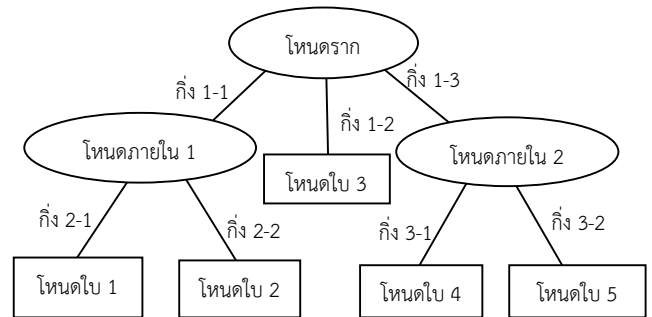
การทบทวนวรรณกรรมหรืองานวิจัยที่เกี่ยวข้อง

1. การทำเหมืองข้อมูล (data mining) เป็นการค้นหาความสัมพันธ์ รูปแบบหรือองค์ความรู้ใหม่ที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่ โดยใช้เทคนิคทางสถิติ ทางคณิตศาสตร์หรือ

เทคนิคทางวิทยาการคอมพิวเตอร์การทำเหมืองข้อมูลเปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลอย่างง่าย ๆ มาสู่การจัดเก็บในฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนถึงการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล

2. การจำแนกประเภทข้อมูล (classification data) เป็นเทคนิคการทำเหมืองข้อมูลหนึ่งที่ทำหน้าที่จัดกลุ่มข้อมูลให้อยู่เป็นกลุ่มตามที่กำหนดโดยอาศัยลักษณะที่คล้ายคลึงกันหรือแตกต่างกัน โดยแบ่งข้อมูลออกเป็น 2 ส่วน คือ ส่วนแรกเป็นชุดข้อมูลเรียนรู้ (training data) เพื่อให้ระบบเรียนรู้ในการจำแนกข้อมูลและได้เป็นตัวแบบ (model) ในการจำแนกข้อมูล ส่วนที่สองเป็นชุดข้อมูลทดสอบ (testing data) ซึ่งนำมาใช้ทดสอบตัวแบบโดยประเมินความถูกต้องในการจำแนกของตัวแบบ และสามารถนำข้อมูลใหม่มาผ่านตัวแบบเพื่อทำนายกลุ่มข้อมูลได้ (บวร น้อยแสง, 2549)

3. ต้นไม้ตัดสินใจ (decision tree) เป็นเทคนิคในการจำแนกข้อมูล เพื่อหาทางเลือกที่ดีที่สุด และสร้างแบบจำลองการพยากรณ์ในรูปแบบของโครงสร้างต้นไม้ ซึ่งประกอบด้วยโหนด (node) แสดงคุณสมบัติ (attribute) ที่ใช้ทดสอบข้อมูล, กิ่ง (branch) แสดงผลลัพธ์ในการทดสอบตามเงื่อนไขและโหนดใบ (leaf node) แสดงกลุ่ม (class) ที่เป็นคำตอบที่ได้กำหนดไว้ดังภาพ 1 ขั้นตอนวิธี (algorithm) ที่ใช้ในการสร้างตัวแบบต้นไม้ตัดสินใจคือ C4.5 ซึ่งถูกพัฒนาขึ้นโดย Quinlan โดยมีหลักการสร้างต้นไม้คือคัดเลือกคุณสมบัติ (attribute) ที่สำคัญที่สุดจากค่า Gain ratio ที่มีค่าสูงที่สุดมาเป็นโหนดราก (root node) (หทัยชนก กรชี, 2551)



ภาพ 1 โครงสร้างต้นไม้ตัดสินใจ

4. โปรแกรมเวกา (WEKA) มาจาก Waikato Environment for Knowledge Analysis เป็นโปรแกรมที่พัฒนามาจากภาษาจาวา เน้นการทำงานด้วยการเรียนรู้ด้วยเครื่อง (machine learning) และการทำเหมืองข้อมูล (รุจิรา ธรรมสมบัติ, 2554)

5. การจัดการความไม่สมดุลของข้อมูล (imbalance data) การจำแนกข้อมูลที่มีกลุ่มเป้าหมาย (class) หนึ่งมากกว่าอีกกลุ่มเป้าหมาย (class) จำนวนมาก จะเกิดปัญหาความไม่สมดุลของข้อมูล ทำให้ผลลัพธ์ที่ได้จากการจำแนกข้อมูลมีความโน้มเอียงไปทางกลุ่มเป้าหมาย (class) ที่มีจำนวนมาก ซึ่งวิธีการแก้ปัญหาคือการจัดการความไม่สมดุลของข้อมูลทำได้หลายวิธี (เขาวนนท์ โสโธ, 2556)

5.1 การปรับความสมดุลของข้อมูล (cost-sensitive learning) เป็นวิธีการพิจารณาค่าใช้จ่ายในการจำแนกผิด (misclassification cost) คือ ฟังก์ชันระหว่างกลุ่มค่าความจริง (actual class) และกลุ่มค่าทำนาย (predicted class) แทนด้วย คอสต์เมทริกซ์ (cost matrix) โดยค่าใช้จ่าย (cost) เป็นเวลาที่ใช้

5.2 การสุ่มเพิ่มจำนวนข้อมูล (oversampling) เป็นการสุ่มเพิ่มจำนวนข้อมูลของกลุ่มที่เป็นเสียงข้างน้อย (minority class) ให้มีปริมาณข้อมูลใกล้เคียงหรือเท่ากับข้อมูลในกลุ่มที่เป็นเสียงข้างมาก (majority class) เช่น เทคนิค

SMOTE (Synthetic Minority Over-sampling Technique) ซึ่งมีวิธีการ คือ

5.2.1 คำนวณหาผลต่างระหว่างชุดข้อมูลที่พิจารณาและชุดข้อมูลที่ใกล้เคียง

5.2.2 คำนวณหาค่าข้อมูลใหม่ด้วยการคูณผลลัพธ์จากข้อ 1 ด้วยค่าที่ได้จากการสุ่มตัวเลขที่อยู่ในช่วง 0 ถึง 1

6. การวัดประสิทธิภาพใช้ วิธีการตรวจสอบไขว้ (k fold cross validation) โดยแบ่งชุดข้อมูลออกเป็น k กลุ่มขนาดเท่ากันและเลือกข้อมูลหนึ่งกลุ่มมาเพื่อเป็นชุดข้อมูลทดสอบ และข้อมูลที่เหลืออีก k-1 กลุ่มถูกใช้เป็นชุดข้อมูลเรียนรู้ ทำการคำนวณจนกระทั่งข้อมูลทุกชุดถูกนำมาใช้เป็นชุดเรียนรู้และชุดทดสอบ ซึ่งการวิจัยนี้ใช้ k=10 และคำนวณหาประสิทธิภาพเฉลี่ย (สิทธิโชค มุกดาสกุลภิบาล, 2551)

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาปัจจัยที่มีผลต่อการทำนายการคลอดก่อนกำหนดในหญิงตั้งครรภ์ในโรงพยาบาลตำรวจ

วิธีดำเนินการวิจัย

ขั้นตอนการทำวิจัยอาศัยตามหลักการทำเหมืองข้อมูลของ CRISP-DM (Cross-Industry Standard Process for Data Mining) ซึ่งเป็นกระบวนการที่บริษัท Daimler Chrysler (Daimler-Benz) บริษัท SPSS (ISL) และบริษัท NCR ได้ร่วมกันพัฒนาขึ้นเพื่อทำเหมืองข้อมูล (เอกสิทธิ์ พชรวงศ์ศักดิ์ และ สิริวรรณ แต่วิจิตร, 2554) ประกอบด้วย

1. การทำความเข้าใจปัญหา พบว่า การคลอดก่อนกำหนดเป็นภาวะหนึ่งที่เกิดขึ้นโดยไม่ทราบสาเหตุที่แน่นอน แต่มีปัจจัยเสี่ยงต่าง ๆ ที่ทำให้เกิดการคลอดก่อนกำหนด ดังนั้น การค้นหาปัจจัยเสี่ยงที่ทำให้เกิดการคลอดก่อนกำหนดจะช่วยให้นักวิจัยการคลอดก่อนกำหนดได้ถูกต้อง

ช่วยป้องกันการคลอดก่อนกำหนดและลดอัตราการคลอดก่อนกำหนดได้

2. ทำความเข้าใจข้อมูลเป็นการทำความเข้าใจข้อมูลที่นำมาใช้ในการวิเคราะห์ โดยการวิจัยนี้ใช้ข้อมูลของหญิงตั้งครรภ์ที่คลอดในโรงพยาบาลตำรวจ ตั้งแต่วันที่ 1 มกราคม 2553 ถึง วันที่ 31 ธันวาคม 2554 จำนวน 4,242 ราย โดยใช้ตัวอย่างเป็นข้อมูลของหญิงตั้งครรภ์ที่คลอดทารกมีชีพจำนวน 4,210 ราย จากแฟ้มเวชระเบียนผู้ป่วยนอกแบบรายงานการคลอดของโรงพยาบาลตำรวจ และแบบรายงานการคลอดก่อนกำหนด

3. เตรียมข้อมูล (data preparation) เป็นการเตรียมข้อมูลให้พร้อมนำไปใช้วิเคราะห์ข้อมูลโดยการรวมข้อมูลและนำไปเตรียมข้อมูลดังนี้

3.1 คัดเลือกข้อมูล โดยเลือกปัจจัยที่เกี่ยวข้องกับการคลอดก่อนกำหนด ซึ่งประกอบด้วย

3.1.1 ปัจจัยส่วนบุคคล ได้แก่ อายุ อาชีพ ส่วนสูง และระดับการศึกษา

3.1.2 ปัจจัยด้านสุขภาพอนามัยของหญิงตั้งครรภ์ ได้แก่ การสูบบุหรี่ ค่าดัชนีมวลกาย การฝากครรภ์ จำนวนครั้งการฝากครรภ์ จำนวนครั้งของการตั้งครรภ์ ลำดับที่การคลอด ประวัติการแท้งบุตร ประวัติการคลอดก่อนกำหนด ประวัติการคลอดทารกตายปริกำเนิด ระดับความเข้มข้นของโลหิต ภาวะแทรกซ้อนทางสูติศาสตร์ระหว่างการตั้งครรภ์และการคลอด ประวัติเจ็บป่วยก่อนการตั้งครรภ์ และประวัติเจ็บป่วยระหว่างการตั้งครรภ์

3.1.3 ปัจจัยด้านทารก ได้แก่ เพศ และความผิดปกติของทารก

3.2 ทำความสะอาดข้อมูล จากการเก็บบันทึกข้อมูลพบว่า มีข้อมูลที่ไม่สมบูรณ์ ได้แก่ มีค่าว่าง (missing data) มีข้อมูลซ้ำซ้อน ซึ่งแก้ไขโดยการแทนที่ข้อมูลที่ขาดหายด้วยค่าศูนย์ “0” และลบข้อมูลที่มีความซ้ำซ้อน

3.3 แปลงข้อมูล ข้อมูลที่ใช้ในการวิจัยมีทั้งเป็นตัวเลข ตัวอักษร ข้อมูลต่อเนื่อง และข้อมูลไม่ต่อเนื่องจำเป็นต้องแปลงข้อมูลให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ได้

4. สร้างตัวแบบ (model) โดยนำข้อมูลที่เก็บบันทึกมาแปลงให้อยู่ในรูปแบบ ARFF (Attribute-Relation File Format) ประกอบด้วย ส่วนหัวของแฟ้ม (relation) คุณสมบัติของข้อมูล (attribute) และส่วนข้อมูล (data) ตามภาพ 2 และนำไปวิเคราะห์ด้วยขั้นตอนวิธีต้นไม้ตัดสินใจ C4.5 เลือกวิธีวัดประสิทธิภาพตัวแบบด้วยการตรวจสอบไขว้แบบ 10 กลุ่ม (10 fold cross validation) ในโปรแกรมเวก้า 3.6.2

```
@relation preterm
@attribute AN numeric
@attribute Age {1,2,3,4,5}
.....
@attribute result {1,2}
@data
2,186192,3,4,2,2,1,1,2,1,4,2,4,1,1,12,1,1,3,1,2,2,1
3,186197,3,4,1,1,1,1,1,1,4,2,3,1,1,1,1,1,2,1,2,3,1
.....
4212,186331,2,3,2,2,1,1,1,1,2,1,5,6,2,1,1,1,2,1,2,3,2
```

ภาพ 2 ตัวอย่างแฟ้มรูปแบบ ARFF (Attribute-Relation File Format)

5. การประเมินตัวแบบ พิจารณาจาก

5.1 ประสิทธิภาพตัวแบบ ประเมินได้จากค่าความถูกต้อง (correctly classified instance) ในการจำแนกข้อมูล

5.2 ค่า Confusion matrix ซึ่งเป็นตารางสรุปจำนวนการจำแนกข้อมูลของตัวแบบได้ถูกต้อง และไม่ถูกต้อง สามารถนำมาคำนวณเพื่อวัดประสิทธิภาพ โดยการหาค่าดังนี้ ค่าความแม่นยำ (accuracy) คำนวณจากค่าของข้อมูลที่มีการจำแนกถูกต้องโดยพิจารณาจากผลรวมของ

การจำแนกข้อมูลทั้งหมด อัตราความถูกต้องเชิงบวก (True positive rate: TP rate) ค่าความระลึก (recall) คำนวณจากค่าของข้อมูลที่ผลลัพธ์ถูกต้องโดยพิจารณาจากข้อมูลของผลลัพธ์เดียวกัน ค่าความถ่วงดุล (F-measure) คำนวณได้จากค่าเฉลี่ยระหว่างค่าความแม่นยำและค่าความระลึก

ผลการวิจัย

ผลการวิเคราะห์ตัวแบบด้วยขั้นตอนวิธีต้นไม้ตัดสินใจ พบว่า ตัวแบบมีความแม่นยำในการทำนายข้อมูลถูกต้อง (correctly classified instances) สูงถึงร้อยละ 94 แต่ไม่สามารถจำแนกกลุ่มที่คลอตก่อนกำหนดได้อย่างถูกต้อง เนื่องจากเกิดความไม่สมดุลของข้อมูล (imbalanced data) ขึ้นโดยมีความแตกต่างของข้อมูลกลุ่มที่คลอตก่อนกำหนดกับข้อมูลของกลุ่มที่คลอตก่อนกำหนดอย่างมาก ดังนั้น ผู้วิจัยจึงเลือกวิธีสุ่มเพิ่มขนาดของข้อมูลเพื่อให้ข้อมูลทั้ง 2 กลุ่มมีความสมดุลกัน ด้วยเทคนิค SMOTE (Synthetic Minority Over-sampling Technique) ก่อนนำไปวิเคราะห์ใหม่ โดยพบว่ามีความแม่นยำในการทำนายข้อมูลถูกต้องคิดเป็นร้อยละ 90.37 อัตราความถูกต้องเชิงบวก (TP Rate) เท่ากับร้อยละ 98.4 ค่าความแม่นยำ (Precision) เท่ากับร้อยละ 98.44 ค่าความระลึก (Recall) เท่ากับร้อยละ 97.82 และค่าความเฉลี่ย (F-measure) เท่ากับร้อยละ 98.09

ผลการวิจัยครั้งนี้พบว่า ปัจจัยที่มีผลต่อการทำนายการคลอตก่อนกำหนดในหญิงตั้งครรภ์ในโรงพยาบาลตำรวจ คือ ภาวะแทรกซ้อนทางสูติศาสตร์ระหว่างการตั้งครรภ์และการคลอด ได้แก่ ภาวะถุงน้ำคร่ำแตกก่อนเจ็บครรภ์ การมีเลือดออกระหว่างตั้งครรภ์ ภาวะพิษแห่งครรภ์ การติดเชื้อในน้ำคร่ำ และการตั้งครรภ์แฝด

ผลการวิเคราะห์ตัวแบบด้วยขั้นตอนวิธีต้นไม้ตัดสินใจ พบว่า ตัวแบบมีความแม่นยำใน

การทำนายข้อมูลถูกต้อง (correctly classified instances) สูงถึงร้อยละ 94 แต่ไม่สามารถจำแนกกลุ่มที่คลอตก่อนกำหนดได้อย่างถูกต้อง เนื่องจากเกิดความไม่สมดุลของข้อมูลขึ้น (imbalance data) โดยมีความแตกต่างของข้อมูลกลุ่มที่คลอตก่อนกำหนดกับข้อมูลของกลุ่มที่คลอตก่อนกำหนดอย่างมาก ดังนั้น ผู้วิจัยจึงได้เลือกวิธีสุ่มเพิ่มขนาดของข้อมูลเพื่อให้ข้อมูลทั้ง 2 กลุ่มมีความสมดุลกันด้วยเทคนิค SMOTE ก่อนนำไปวิเคราะห์ใหม่ โดยพบว่า มีความแม่นยำในการทำนายข้อมูลถูกต้องคิดเป็นร้อยละ 90.37 อัตราความถูกต้องเชิงบวก (TP Rate) เท่ากับร้อยละ 98.4 ค่าความแม่นยำ (precision) เท่ากับร้อยละ 98.44 ค่าความระลึก (recall) เท่ากับร้อยละ 97.82 และค่าความเหวี่ยง (F-measure) เท่ากับร้อยละ 98.09 ผลการวิจัยครั้งนี้พบว่า ปัจจัยที่มีผลต่อการทำนายการคลอตก่อนกำหนดในหญิงตั้งครรภ์ในโรงพยาบาลตำรวจ คือ ภาวะแทรกซ้อนทางสูติศาสตร์ระหว่างการตั้งครรภ์และการคลอตก่อนกำหนด ได้แก่ ภาวะถุงน้ำคร่ำแตกก่อนเจ็บครรภ์ การมีเลือดออกระหว่างตั้งครรภ์ ภาวะพิษแห่งครรภ์ การติดเชื้อในน้ำคร่ำ และการตั้งครรภ์แฝด

การอภิปรายผลการวิจัย

การวิเคราะห์ปัจจัยที่มีผลต่อการทำนายการคลอตก่อนกำหนดด้วยวิธีการทำเหมืองข้อมูลพบว่าปัจจัยที่ทำนายการคลอตก่อนกำหนด คือ ภาวะแทรกซ้อนทางสูติศาสตร์ระหว่างการตั้งครรภ์และการคลอตก่อนกำหนด ได้แก่ ภาวะถุงน้ำคร่ำแตกก่อนเจ็บครรภ์ การมีเลือดออกระหว่างการตั้งครรภ์ การมีสายสะดือพลัดต่ำ มีภาวะพิษแห่งครรภ์ การมีน้ำคร่ำน้อย และการตั้งครรภ์แฝด ซึ่งสอดคล้องกับการศึกษาด้วยวิธีทางสถิติของ สุลลตตินันทพร (2554) ที่พบว่า ปัจจัยเสี่ยงที่ทำให้เกิดการคลอตก่อนกำหนด ได้แก่ การมีสายสะดือพลัดต่ำ ภาวะพิษแห่งครรภ์ และการมีน้ำคร่ำน้อย

และการศึกษาของ ฐิติกันต์ ณ.ปัน (2557) ที่พบว่า การมีเลือดออกระหว่างการตั้งครรภ์เป็นปัจจัยที่มีความสัมพันธ์กับการเจ็บครรภ์คลอตก่อนกำหนดอย่างมีนัยสำคัญทางสถิติ ($p = 0.002$) สำหรับการวิเคราะห์ด้วยเทคนิคเหมืองข้อมูล สอดคล้องกับการศึกษาของ Chen et al. (2011) โดยพบว่า ปัจจัยที่สัมพันธ์กับการคลอตก่อนกำหนด ได้แก่ ภาวะตกเลือดก่อนคลอดและการตั้งครรภ์แฝด

ข้อเสนอแนะ

1. นำผลการศึกษาไปใช้ในการพัฒนาระบบวิเคราะห์ปัจจัยในการคลอตก่อนกำหนดของหญิงตั้งครรภ์เพื่อนำไปใช้ในการคัดกรองหญิงตั้งครรภ์ให้ได้ครอบคลุม สะดวก และรวดเร็วยิ่งขึ้น
2. พัฒนาแผนการดูแลรักษาแก่หญิงตั้งครรภ์รายใหม่ที่มีภาวะแทรกซ้อนทางสูติศาสตร์และมีปัจจัยเสี่ยงให้ได้รับสุขศึกษาอย่างเหมาะสมเพื่อลดการคลอตก่อนกำหนดและอาการแทรกซ้อนจากการคลอตก่อนกำหนด
3. นำเทคนิคเหมืองข้อมูลมาใช้ในการวิเคราะห์หารูปแบบความสัมพันธ์ของภาวะเจ็บป่วยอื่น เช่น ภาวะถุงน้ำคร่ำแตกก่อนเจ็บครรภ์ ภาวะพิษแห่งครรภ์ และการติดเชื้อ เป็นต้น เพื่อทำนายการเกิดภาวะเจ็บป่วยนั้น
4. การวิเคราะห์ข้อมูลด้วยเทคนิคเหมืองข้อมูล ควรใช้จำนวนข้อมูลที่มีมากกว่านี้หรือเก็บข้อมูลในระดับประเทศจะทำให้ไม่เกิดความไม่สมดุลของข้อมูลที่มีความแตกต่างกันมาก ทำให้ได้ผลลัพธ์ที่ดีกว่าและสามารถนำไปใช้ประโยชน์ได้อย่างกว้างขวาง

กิตติกรรมประกาศ

การศึกษานี้สำเร็จลุล่วงได้ด้วยความกรุณาของ อาจารย์ ดร. บุญช่วย ศรีธรรมศักดิ์ ประธานกรรมการที่ปรึกษาวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.สุรียพันธ์ วรพงศธร ที่กรุณาให้คำปรึกษาและ

ความช่วยเหลือแก้ไขข้อบกพร่องต่าง ๆ อาจารย์วีรพันธ์ แซ่ด่าน ที่ให้คำปรึกษาในส่วนการวิเคราะห์ข้อมูล เจ้าหน้าที่ห้องคลอดและเวชระเบียนโรงพยาบาลตำรวจทุกท่านที่ให้ความช่วยเหลือและอำนวยความสะดวกในการเก็บรวบรวมข้อมูลสุดท้ายขอขอบพระคุณมหาวิทยาลัยรามคำแหงที่เป็นแหล่งเรียนรู้และให้ความรู้ในการศึกษาระดับต่าง ๆ แก่คนทุกชนชั้น ทุกเพศ และทุกวัย

เอกสารอ้างอิง

- เชาวรัตน์ โลโท. (2556). แบบจำลองการทำนายผลการรักษาผู้ป่วยมะเร็งปากมดลูกด้วยโครงข่ายประสาทเทียม. *วารสารวิจัย มช. (บค)*, 13(1), 39-50.
- ฐิติกันต์ ณ ปั่น. (2557). ปัจจัยที่มีผลต่อการเจ็บครรภ์คลอดก่อนกำหนดของหญิงตั้งครรภ์ โรงพยาบาลอมก๋อย อ.อมก๋อย จ.เชียงใหม่. *วารสารสาธารณสุขล้านนา*, 10(2), 142-150.
- ณัฐญา เนตรหิน. (2555). *สธ.เผยสถิติหญิงตั้งครรภ์คลอดก่อนกำหนด 6.4-8 หมื่นคน/ปีกว่า 1 หมื่นรายเด็กต้องอยู่ I.C.U..* สืบค้นได้จาก http://www.thanonline.com/index.php?option=com_content&view=article&id=151171:64-8-1icu
- ดิษฐพล มั่นธรรม. (2553). *การประยุกต์อัลกอริทึมต้นไม้ตัดสินใจกับการวินิจฉัยโรคระบบทางเดินหายใจ (วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต)*. มหาวิทยาลัยเกษตรศาสตร์.
- บวร น้อยแสง. (2549). *คลังข้อมูลและเทคนิคการทำเหมืองข้อมูลสำหรับการวิเคราะห์การขยาย (สารนิพนธ์วิทยาศาสตร์มหาบัณฑิต)*. มหาวิทยาลัยศิลปากร.
- ประภัทร วานิชพงษ์พันธุ์. (2554). การดูแลรักษาการเจ็บครรภ์คลอดก่อนกำหนด. ใน ประสงค์ ตันมหาสมุทร, พญหัสจันทรประภาพร, อรรถพล ใจชื่น, ธันยรัตน์ วงศ์วานิชย์, และ ศิริภาพ เลิศบรรณพงษ์ (บก.). *สูติเวชทันยุค OB-GYN in Practice 2011*, หน้า 111-119. กรุงเทพมหานคร: สำนักพิมพ์ พี. อี. ลิฟวิ่ง.
- ปัญญ พันธ์บุรณะ. (2554). การเจ็บครรภ์คลอดก่อนกำหนด. ใน ถวัลย์วงศ์ รตินสิริ, ฐิติมา สุนทรสัจ, และ สมศักดิ์ สุทัศน์วรวิทย์ (บก.). *สูติศาสตร์ฉุกเฉิน (พิมพ์ครั้งที่ 2, หน้า 91-130)*. สมุทรสาคร: พิมพ์ดี.
- รุจิรา ธรรมสมบัติ. (2554). *ระบบสนับสนุนการตัดสินใจในการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือโดยใช้ต้นไม้ตัดสินใจ*. รายงานการวิจัยวิทยาลัยราชพฤกษ์. กรุงเทพฯ: วิทยาลัยราชพฤกษ์.
- รุ่งตวรรณ ช้อยจ่อหอ, เยาวลักษณ์ เสรีเสถียร, เอมพร รตินธร, และ เอกชัย โคววิสารัช. (2552). ผลของการบำบัดทางการพยาบาลต่อการจัดการเพื่อป้องกันภาวะเจ็บครรภ์คลอดก่อนกำหนด การกลับมารักษาซ้ำและอายุครรภ์เมื่อคลอด ในหญิงตั้งครรภ์ที่มีภาวะเจ็บครรภ์คลอดก่อนกำหนด. *Journal of Nursing Science (Supplement)*, 27(2), 39-48.
- รักถิ่น เหลาหา. (2553). *การพยากรณ์ความเสี่ยงการเกิดโรคมะเร็งปอด โดยใช้ทฤษฎีของการทำเหมืองข้อมูล (การศึกษาอิสระวิทยาศาสตร์มหาบัณฑิต)*. มหาวิทยาลัยขอนแก่น.
- สายฝน ขวาลไพบูลย์. (2553). *ตำราคลอดก่อนกำหนด (preterm labour)*. กรุงเทพมหานคร: สำนักพิมพ์ พี. เอ. ลิฟวิ่ง.
- สิทธิโชค มุกดาสกุลภิบาล. (2551). *การวัดประสิทธิภาพของขั้นตอนวิธีตัวจำแนก C4.5, ADTree, และ Naive Bayes ในการจำแนกข้อมูลการชุกซ่อนสิ่งเสพติดสำหรับไปรษณีย์ระหว่างประเทศ (วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต)*. มหาวิทยาลัยเกษตรศาสตร์.
- สุผล ตติยันทพร. (2554). ปัจจัยเสี่ยงของมารดาต่อการคลอดก่อนกำหนด. *Khon Kaen Medical Journal*, 35(2), 7-14.
- หทัยชนก กรชี. (2551). *การคำนวณน้ำหนักของคุณลักษณะข้อมูลสำหรับตัวจำแนก ในอีฟ เบย์ โดยวิธีการจัดกลุ่มข้อมูลและต้นไม้ตัดสินใจ (วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต)*. มหาวิทยาลัยเกษตรศาสตร์.
- อดุลย์ ยิ้มงาม. (ม.ป.ป.). *การทำเหมืองข้อมูล (data mining)*. สืบค้นได้จาก http://computer.bu.ac.th/index.php?option=com_content&task=view&id=75&Itemid=172
- เอกสิทธิ์ พชรวงศ์ศักดิ์ และ สิริวรรณ แต้วิจิตร. (2554). *เอกสารประกอบการอบรม เชิงปฏิบัติการ An Introduction to Data Mining (Workshop with WEKA)*. เอกสารอัดสำเนา, มหาวิทยาลัยเกษตรศาสตร์ (บางเขน).
- Chen, H-Y., Chuang, C. H., Yang, Y-J., & Wu, T-P. (2011). Exploring the risk factors of preterm birth using data mining. *Journal Expert Systems with Applications: An International Journal*, 38, 5384-5387.