

การวินิจฉัยการเกิดโรคไส้ติ่งอักเสบโดยใช้เทคนิคการเรียนรู้ของเครื่อง

Appendicitis Disease Diagnosis using Machine Learning Techniques

นภาพร ดวงแก้ว (Napaporn Duangkaew)¹ อมรรัตน์ หลอดทองกลาง (Amornrad Lodtonglang)²

และพิชัย จอดพิมาย (Pichai Jodpimai)³

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก วิทยาเขตจรัลพงษานุวรณ

¹napaporns2916@gmail.com, ²taewbangna01@gmail.com, ³pichai.j@cpc.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างโมเดลสำหรับวินิจฉัยโรคไส้ติ่งอักเสบ และเพื่อค้นหาปัจจัยที่มีความสำคัญต่อการวินิจฉัยโรคไส้ติ่งอักเสบ ผู้วิจัยทำการสำรวจกลุ่มตัวอย่างทั้งหมด 232 คน แบ่งออกเป็น ผู้ที่เคยเป็นโรคไส้ติ่งอักเสบ จำนวน 63 คน และผู้ที่ไม่เคยเป็นโรคไส้ติ่งอักเสบ จำนวน 169 คน ข้อมูลจากกลุ่มตัวอย่างนี้นำมาใช้สำหรับสร้างโมเดลในการวินิจฉัยโรคโดยใช้เทคนิคการเรียนรู้ของเครื่อง ผลการวิจัยพบว่า ปัจจัยด้านอาการมีความสำคัญและให้ประสิทธิภาพสูงสุด ซึ่งมีค่า Accuracy ร้อยละ 90-95, Precision ร้อยละ 92-95, Recall ร้อยละ 86-94, และ F1-score ร้อยละ 87-94 และพบว่า SVM ให้ประสิทธิภาพสูงสุดในกลุ่มของเทคนิคการเรียนรู้ของเครื่อง โดยเฉพาะอย่างยิ่งเมื่อใช้ SVM ร่วมกับปัจจัยด้านอาการ ซึ่งให้ค่า Accuracy ร้อยละ 95, Precision ร้อยละ 94, Recall ร้อยละ 94, และ F1-score ร้อยละ 94 จากผลการวิจัยชี้ให้เห็นว่า แนวทางที่นำเสนอนี้สามารถใช้เป็นเครื่องมือในการวินิจฉัยโรคไส้ติ่งอักเสบได้อย่างมีประสิทธิภาพ ซึ่งเป็นประโยชน์ต่อทั้งบุคลากรทางการแพทย์และบุคคลทั่วไป

คำสำคัญ: ไส้ติ่งอักเสบ การวินิจฉัยโรค การทำนาย การเรียนรู้ของเครื่อง

Abstract

The objectives of this paper are to establish prediction models for diagnosing appendicitis disease and to find important factors to diagnosis the disease. In this paper, we collected 232 samples consisting of

63 samples being appendicitis disease, and 169 samples not being the disease. The samples were yielded to establish the model using machine learning techniques. The experimental results show that the symptom factors are the most important and provide the highest performance, i.e., Accuracy=90-95%, Precision=92-95%, Recall=86-94%, and F1-score 87-94%. Additionally, SVM yields the highest performance among different machine learning techniques, especially, using SVM with symptom factors. The performance is 95%, 94%, 94%, and 94% for Accuracy, Precision, Recall, and F1-score, respectively. The results indicate that the proposed approach can be used for appendicitis diagnosis which can be benefit to not only medical personnel but also ordinary people.

Keyword: Appendicitis, Diagnosis, Prediction, Machine Learning

1. บทนำ

การเกิดภาวะไส้ติ่งอักเสบเป็นอาการที่เกิดขึ้นได้อย่างเฉียบพลันและพบได้ทุกอายุ แต่พบบ่อยที่สุดอยู่ระหว่าง 10-20 ปี [1] โดยลักษณะอาการของภาวะไส้ติ่งอักเสบประกอบด้วย อาการปวดท้องมาก เริ่มแรกอาจปวดเป็นพัก ๆ รอบสะดือ ต่อมาอาการปวดจะย้ายมาที่ท้องน้อยข้างขวา ลักษณะปวดเสียดตลอดเวลา คลื่นไส้ อาเจียน และมีไข้ อาการจะเป็นอยู่นับชั่วโมงถึงหลายวัน บางคนอาจมีอาการปวดท้องน้อยข้างขวา โดยไม่มีอาการอื่นนำมาก่อนเลยก็ได้ [2], [3] ทั้งนี้หากเกิดความล่าช้าในการรักษาอาจ

ทำให้ใส่ตั้งอีกเสบหรือเกิดภาวะแทรกซ้อน เป็นสาเหตุทำให้เสียชีวิตได้ [4] ดังนั้นการวินิจฉัยโรคอย่างรวดเร็วและแม่นยำ จึงมีความสำคัญอย่างยิ่งในการวินิจฉัยก่อนที่จะนำผู้ป่วยไปรักษา และให้การผ่าตัดอย่างทันที่ ซึ่งสามารถลดภาวะแทรกซ้อนรุนแรงได้อีกด้วย [1] งานวิจัยที่ผ่านมาได้มีการเสนอแนวทางในการวินิจฉัยโรคใส่ตั้งอีกเสบด้วยวิธีที่แตกต่างกัน ได้แก่ การอัลตราซาวด์ [5], การใช้คะแนนพยากรณ์ [6], [7], [8] และการเอกซเรย์คอมพิวเตอร์ [2], [9] อย่างไรก็ตาม ประสิทธิภาพของการวินิจฉัยการเกิดโรคใส่ตั้งอีกเสบยังต้องได้รับการปรับปรุงให้มีความแม่นยำมากยิ่งขึ้น เพื่อลดต้นทุนในการผ่าตัดโดยไม่จำเป็นและลดอันตรายต่อผู้เข้ารับการรักษา

งานวิจัยนี้เสนอแนวทางในการพัฒนาโมเดลสำหรับการวินิจฉัยการเกิดโรคใส่ตั้งอีกเสบ โดยใช้เทคนิคการเรียนรู้ของเครื่อง และศึกษาถึงปัจจัยที่ส่งผลต่อการทำนายการเกิดโรคใส่ตั้งอีกเสบทั้งด้านประชากรศาสตร์ ด้านพฤติกรรม และด้านอาการ

2. ทบทวนวรรณกรรม

มีงานวิจัยที่ได้ทำการศึกษาเกี่ยวกับการภาวะการเกิดโรคใส่ตั้งอีกเสบ ดังนี้

อลิสรา วงศ์สุทธิเลิศ และคณะ [5] ศึกษาประสิทธิภาพในการประเมินภาวะใส่ตั้งอีกเสบด้วยอัลตราซาวด์ในกลุ่มผู้ป่วยเด็ก โดยทำการรวบรวมข้อมูลผู้ป่วยเด็กอายุ 1-16 ปี ที่ได้รับการทำอัลตราซาวด์เนื่องจากสงสัยว่าเป็นใส่ตั้งอีกเสบ จำนวนทั้งหมด 428 คน ผลการศึกษาพบว่าการอัลตราซาวด์สามารถพบผู้ป่วยที่เป็นใส่ตั้งอีกเสบจำนวน 270 คน คิดเป็นร้อยละ 61.1 โดยมีค่า Recall ร้อยละ 71.2 ค่า Precision ร้อยละ 88.1 และค่า Accuracy ร้อยละ 92.6 จากผลการศึกษา สรุปได้ว่า การอัลตราซาวด์มีประสิทธิภาพมากพอในการวินิจฉัยโรคใส่ตั้งอีกเสบ โดยเฉพาะกลุ่มผู้ป่วยเด็กที่มีปัญหาในการวินิจฉัยจากประวัติและการตรวจร่างกาย

เนตรา เนตรโพธิ์แก้ว [6] ศึกษาการประเมินผลของการใช้ระบบคะแนน Alvarado score เพื่อทำการวินิจฉัยโรคใส่ตั้งอีกเสบ โดยใช้ข้อมูลผู้ป่วยที่วินิจฉัยว่า

เป็นโรคใส่ตั้งอีกเสบและได้รับการผ่าตัดใส่ตั้งในภาวะฉุกเฉินทั้งหมด 103 คน ผลการศึกษาพบว่า Alvarado Score ที่มีค่ามากกว่าหรือเท่ากับ 7 ให้ค่า Recall ร้อยละ 91.2 ค่า Precision ร้อยละ 89.25 และ Accuracy ร้อยละ 82.52 จากผลการศึกษาชี้ให้เห็นว่าการวินิจฉัยโรคใส่ตั้งอีกเสบเทียบพลงด้วยระบบคะแนน Alvarado Score สามารถนำไปใช้ในทางปฏิบัติต่อการวินิจฉัยโรคใส่ตั้งอีกเสบเทียบพลง

เนตรา เนตรโพธิ์แก้ว [7] ศึกษาการเปรียบเทียบการประเมินผลของ Alvarado Score และ RIPASA Score ในการวินิจฉัยโรคใส่ตั้งอีกเสบ โดยเก็บรวบรวมข้อมูลผู้ป่วยทั้งหมด 145 คน จากการศึกษาพบว่า การวินิจฉัยโรคใส่ตั้งอีกเสบเทียบพลงด้วย RIPASA Score ที่ค่าจุดตัด 7.5 พบค่า Recall และ Accuracy ที่มากกว่า Alvarado Score แต่ไม่สามารถ Rule In โรคใส่ตั้งอีกเสบได้เลย เมื่อเปลี่ยนค่าจุดตัดใน RIPASA score เป็นค่าจุดตัดที่ 10.75 พบว่า มีความสามารถในการ Rule In และ Rule Out โรคใส่ตั้งอีกเสบได้ใกล้เคียงกับการประเมิน Alvarado Score ที่จุดตัด 7 ซึ่งสามารถนำมาใช้เป็นหลักเกณฑ์ในการพิจารณาผู้ป่วยโรคใส่ตั้งอีกเสบไว้รักษาได้

ปริญญารักษ์ เสนะวัต และกัจจา เจียรวัฒนกก [8] ศึกษาคะแนนการพยากรณ์ภาวะใส่ตั้งอีกเสบในผู้ป่วยใส่ตั้งอีกเสบเทียบพลงโดยใช้อาการทางคลินิก ผู้วิจัยทำการรวบรวมข้อมูลผู้ป่วยที่มีผลพยาธิวิทยาและได้รับการผ่าตัดทั้งหมด 499 คน ผลการศึกษาพบปัจจัยเสี่ยง 5 ปัจจัย ประกอบด้วย เพศชาย อายุมากกว่า 20 ปี ระยะเวลาปวดท้องจนถึงโรงพยาบาลมากกว่า 24 ชั่วโมง อัตราการเต้นของหัวใจมากกว่าหรือเท่ากับ 130 ครั้งต่อนาที และการตรวจท้องพบอาการปล่อยเจ็บบริเวณขวาล่าง โดยระบบคะแนนพยากรณ์นี้ให้ประสิทธิภาพร้อยละ 77.98 แสดงให้เห็นว่าระบบคะแนนนี้ สามารถใช้เพื่อเป็นแนวทางในการตัดสินใจในการรักษา

ชนิกาน มหารักษ์ [2] ศึกษาการวินิจฉัยโรคอื่นในผู้ป่วยที่สงสัยว่าเป็นโรคใส่ตั้งอีกเสบเทียบพลงและมีการตรวจด้วยเอกซเรย์คอมพิวเตอร์ช่องท้อง โดยใช้ข้อมูลผู้ป่วยทั้งหมด 120 คน ที่สงสัยว่าเป็นใส่ตั้งอีกเสบและได้รับการ

เอกซเรย์คอมพิวเตอร์ช่องท้อง ผลการศึกษาพบว่า ผู้ป่วยที่ได้รับการวินิจฉัยเป็นโรคไส้ติ่งอักเสบมีทั้งหมด 68 คน คิดเป็นร้อยละ 56.7 วินิจฉัยเป็นโรคอื่น ๆ จำนวน 43 คน คิดเป็นร้อยละ 35.8 และไม่สามารถวินิจฉัยโรคจากการเอกซเรย์คอมพิวเตอร์ช่องท้อง จำนวน 9 คน คิดเป็นร้อยละ 7.5 จากผลการศึกษาชี้ให้เห็นว่าการเอกซเรย์คอมพิวเตอร์ช่องท้องเป็นการตรวจที่มีประโยชน์ในผู้ป่วยที่สงสัยว่าเป็นโรคไส้ติ่งอักเสบเฉียบพลัน

จากงานวิจัยที่กล่าวมาพบว่ามีข้อเสนอแนะทางในการวินิจฉัยโรคไส้ติ่งอักเสบด้วยวิธีการอัลตราซาวด์ การคำนวณคะแนนพยากรณ์ และการเอกซเรย์คอมพิวเตอร์ งานวิจัยนี้ เสนอแนะทางในการวินิจฉัยโรคโดยการประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่อง

3. วิธีดำเนินการวิจัย

3.1 การเก็บรวบรวมข้อมูล

ผู้วิจัยจัดทำแบบสอบถาม โดยปรับปรุงจากแบบประเมิน Alvarado Score และ RIPASA Score [6], [7] จากนั้นให้ผู้ใช้งานเฟซบุ๊ก (Facebook) ตอบแบบสอบถามผ่านกูเกิลฟอร์ม (Google Form) จำนวน ทั้งหมด 23 ข้อ แบ่งเป็น 22 ข้อ เกี่ยวกับปัจจัยที่ส่งผลต่อการเป็นโรคไส้ติ่งอักเสบ และอีก 1 ข้อ เกี่ยวกับความเป็นโรคไส้ติ่งอักเสบของผู้ตอบแบบสอบถาม (เป็นหรือไม่เป็นโรคไส้ติ่งอักเสบ) ใน 22 ข้อ แบ่งเป็น 3 ด้าน ประกอบด้วย ด้านประชากรศาสตร์ มี 3 ข้อ (อายุ, เพศ, และค่า BMI) ด้านพฤติกรรม มี 12 ข้อ (ประเภทอาหารที่เลือกรับประทาน, การรับประทานอาหารจำพวกจังก์ฟู้ด, การรับประทานอาหารกึ่งสุกกึ่งดิบ, การรับประทานอาหารจำพวกของหวาน, การรับประทานอาหารจำพวกกลุ่มโซเดียมสูง, การรับประทานเมล็ดของผลไม้, การดื่มน้ำ, การดื่มเครื่องดื่มที่มีคาเฟอีน, การดื่มแอลกอฮอล์, การสูบบุหรี่เป็นประจำ, การพักผ่อน, และการออกกำลังกาย) และด้านอาการ มี 7 ข้อ (อาการปวดท้องบริเวณรอบสะดือแล้วย้ายมาปวดบริเวณหน้าท้องด้านขวาตอนล่าง, รู้สึกเหมือนมีก้อนแข็ง ๆ เมื่อกดบริเวณท้องน้อยขวา, อาการไข้

สูง, รู้สึกคลื่นไส้หรืออาเจียน, อาการความเครียดสะสม, รู้สึกเบื่ออาหาร, และอาการท้องผูก)

3.2 การสร้างโมเดล

ผู้วิจัยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) [10] เพื่อสร้างโมเดลในการทำนายการเกิดโรคไส้ติ่งอักเสบ มี 6 เทคนิค ได้แก่ การถดถอยโลจิสติก (Logistic Regression: LR), นาอีฟเบย์ (Naïve Bayes: NB), ต้นไม้ตัดสินใจ (Decision Tree: DT), เพื่อนบ้านที่ใกล้ที่สุด (K-Nearest Neighbors: KNN), ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM), และมัลติเลเยอร์เพอร์เซปตรอน (Multilayer Perceptron: MLP)

ตัวแปรต้นมีทั้งหมด 22 ตัวแปร ได้มาจากคำถาม 22 ข้อ ที่เกี่ยวกับปัจจัยที่ส่งผลต่อการเป็นโรคไส้ติ่งอักเสบ ตัวแปรตามมี 1 ตัวแปร คือ ความเป็นโรคของผู้ตอบแบบสอบถาม ในการสร้างโมเดลจะพิจารณาตัวแปรในแต่ละด้านรวมถึงการนำตัวแปรของแต่ละด้านมาพิจารณาร่วมกัน ซึ่งมีทั้งหมด 7 กลุ่ม ได้แก่ ด้านประชากรศาสตร์ (3 ตัวแปร), ด้านพฤติกรรม (12 ตัวแปร), ด้านอาการ (7 ตัวแปร), ด้านประชากรศาสตร์และด้านพฤติกรรม (15 ตัวแปร), ด้านประชากรศาสตร์และด้านอาการ (10 ตัวแปร), ด้านพฤติกรรมและด้านอาการ (19 ตัวแปร), ทุกตัวแปร (22 ตัวแปร) ดังนั้นมีโมเดลทั้งหมด 42 โมเดล (6 เทคนิค \times 7 กลุ่มตัวแปร)

3.3 การประเมินประสิทธิภาพของโมเดล

ในส่วนนี้กล่าวถึงการประเมินประสิทธิภาพของโมเดล โดยมีรายละเอียดดังนี้

10-Fold Cross-Validation คือ วิธีการตรวจสอบประสิทธิภาพของโมเดลที่สามารถเป็นอิสระต่อชุดข้อมูลหรือไม่ ถ้าชุดข้อมูลมีทั้งหมด 100 ตัวอย่าง จะถูกแบ่งออกเป็น 10 กอง แต่ละกองมี 10 ตัวอย่าง กองที่ 1 ใช้สำหรับประเมินโมเดลด้วย Evaluation Metrics กองที่ 2-10 ใช้สำหรับสร้างโมเดล หลังจากนั้น ใช้กองที่ 2 สำหรับประเมินโมเดล กองที่ 1 และกองที่ 3-10 ใช้สำหรับสร้างโมเดล ทำเช่นนี้ไปจนกว่าทั้ง 10 กองจะถูกใช้ประเมินโมเดล ดังนั้น แต่ละโมเดลจะถูกประเมินทั้งหมด 10 ครั้ง และผลการประเมินสุดท้ายของแต่ละโมเดล คือ ค่าเฉลี่ย

ของการประเมินจากทั้งหมด 10 ครั้ง

Evaluation Metrics คือ เมตริกสำหรับประเมินประสิทธิภาพของโมเดล ประกอบด้วย Accuracy, Precision, Recall, และ F1-score ดังสมการต่อไปนี้

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3-1)$$

$$Precision = \frac{TP}{TP+FP} \quad (3-2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3-3)$$

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3-4)$$

โดยที่ True Positive (TP) คือ ผลบวกจริง หรือผู้ป่วยตรวจสอบพบว่าเป็นโรคจริง True Negative (TN) คือ ผลลบจริง หรือ คนปกติตรวจพบว่าเป็นโรคจริง False Positive (FP) คือ ผลบวกเท็จ หรือ คนปกติแต่ถูกตรวจพบว่าเป็นโรค False Negative (FN) คือ ผลลบเท็จ หรือ ผู้ป่วยที่ตรวจพบว่าเป็นโรค Accuracy คือ อัตราส่วนของจำนวนตัวอย่าง (ทั้งคนปกติและผู้ป่วย) ที่วินิจฉัยถูกต้องต่อจำนวนตัวอย่างทั้งหมด Precision คือ อัตราส่วนของผลบวกจริงต่อจำนวนตัวอย่างที่วินิจฉัยว่าเป็นโรค Recall คือ อัตราส่วนของผลบวกจริงต่อจำนวนผู้ป่วยที่เป็นโรคจริง และ F1-score คือ ค่าเฉลี่ยฮาร์โมนิก (Harmonic Mean) ของค่า Precision และ Recall

4. ผลการดำเนินงาน

ผู้ตอบแบบสอบถามทั้งหมด 232 คน แบ่งเป็นผู้ที่เคยเป็นโรคไส้ติ่งอักเสบ 63 คน คิดเป็น ร้อยละ 27.16 และผู้ที่ไม่เคยเป็นโรคไส้ติ่งอักเสบ 169 คน คิดเป็น ร้อยละ 72.84 ตารางที่ 1 แสดงสถิติปัจจัยในการทำนายโรค คอลัมน์แรกแสดงชื่อปัจจัย คอลัมน์ที่ 2 แสดงค่าที่แปลงเป็นตัวเลขเพื่อใช้สำหรับการคำนวณในขั้นตอนการสร้างโมเดล กรณีที่ตัวแปรใดอยู่ในระดับ Nominal และมี 2 ค่า จะให้ค่าใดค่าหนึ่งเป็น 0 และอีกค่าเป็น 1 เช่น ตัวแปรเพศ จะให้ “ชาย” เป็น 0 และ “หญิง” เป็น 1 กรณีที่ตัวแปรใดอยู่ในระดับ Ordinal จะให้ค่าแรกเป็น 0 และเพิ่มขึ้นทีละ 1 เช่น ตัวแปรอายุ จะให้ “อายุน้อยกว่า 18 ปี” เป็น 0 “18-35 ปี” เป็น 1 “35-60 ปี” เป็น 2 และ “มากกว่า 60 ปี” เป็น 3 คอลัมน์ที่ 3-5 แสดงความถี่ และคอลัมน์สุดท้ายแสดงค่า P-value

ตารางที่ 1: สถิติของปัจจัยในการทำนายโรคไส้ติ่งอักเสบ

ปัจจัย	ความถี่				P-value
	ชื่อปัจจัย	ค่าที่แปลง	จำนวน	เคยเป็นไส้ติ่ง	
ด้านประชากรศาสตร์					
1. อายุ	-	-	-	-	0.042
น้อยกว่า 18 ปี	0	8	2	6	-
18-35ปี	1	165	37	128	-
35-60ปี	2	58	23	35	-
มากกว่า 60 ปี	3	1	0	1	-
2. เพศ	-	-	-	-	0.149
ชาย	0	84	28	56	-
หญิง	1	148	35	113	-
3. ค่า BMI	-	-	-	-	0.020
ต่ำกว่า 20	0	72	12	60	-
20-30	1	143	45	95	-
มากกว่า 30	2	17	3	14	-
ด้านพฤติกรรม					
4. ประเภทอาหารที่เลือกรับประทาน (>= 3 วันต่อสัปดาห์)	-	-	-	-	0.106
ผัก	0	6	0	6	-
เนื้อสัตว์และผัก	1	163	41	122	-
เนื้อสัตว์	2	63	22	41	-
5. การรับประทานอาหารเช้า (>= 3 วันต่อสัปดาห์)	-	-	-	-	1.607
ไม่ใช่	0	135	16	119	-
ใช่	1	97	47	50	-
6. การรับประทานอาหารเช้าถึงที่สุดถึงดึก (>= 3 วันต่อสัปดาห์)	-	-	-	-	0.027
ไม่ใช่	0	198	48	150	-
ใช่	1	34	15	19	-
7. การรับประทานอาหารเช้าของหวาน (>= 3 วันต่อสัปดาห์)	-	-	-	-	0.0005
ไม่ใช่	0	100	15	85	-
ใช่	1	132	48	84	-
8. การรับประทานอาหารเช้ากลุ่มโซเดียมสูง (>= 3 วันต่อสัปดาห์)	-	-	-	-	4.09E-08
ไม่ใช่	0	99	8	91	-
ใช่	1	133	55	78	-
9. การรับประทานเมล็ดของผลไม้นี้ (>= 3 วันต่อสัปดาห์)	-	-	-	-	1.21E-09
ไม่ใช่	0	187	34	153	-
ใช่	1	45	29	16	-
10. การดื่มน้ำ (< 7 แก้วต่อวัน)	-	-	-	-	0.0009
ไม่ใช่	0	86	12	74	-
ใช่	1	146	51	95	-

ตารางที่ 1: สถิติของปัจจัยในการทำนายโรคไส้ตั้งอีกเสบ (ต่อ)

ปัจจัย	ค่าที่ แปลง	ความถี่			P-value
		จำนวน	เคย เป็นไส้ ตั้ง	ไม่เคย เป็นไส้ ตั้ง	
11. การดื่มเครื่องดื่มที่มีคาเฟอีน (≥ 2 แก้วต่อวัน)	-	-	-	-	0.0002
ไม่ใช้	0	94	13	81	-
ใช้	1	138	50	88	-
12. การดื่มแอลกอฮอล์ (≥ 3 วันต่อสัปดาห์)	-	-	-	-	0.451
ไม่ใช้	0	150	50	143	-
ใช้	1	39	13	26	-
13. การอื่นอุจจาระเป็นประจำ	-	-	-	-	0.025
ไม่ใช้	0	183	61	122	-
ใช้	1	49	2	47	-
14. การพักผ่อน (< 6 ชั่วโมงต่อวัน)	-	-	-	-	0.104
ไม่ใช้	0	103	22	81	-
ใช้	1	129	41	88	-
15. การออกกำลังกาย (< 3 ครั้งต่อสัปดาห์)	-	-	-	-	0.00018
ไม่ใช้	0	75	8	67	-
ใช้	1	157	55	102	-
ด้านอาการ	-	-	-	-	-
16. อาการปวดท้องบริเวณรอบ สะดือแล้วย้ายมาปวดบริเวณหน้า ท้องด้านขวาตอนล่าง	-	-	-	-	4.46E-40
ไม่ใช้	0	168	5	163	-
ใช้	1	64	58	6	-
17. รู้สึกเหมือนมีก้อนแข็ง ๆ เมื่อ กดบริเวณท้องน้อยขวา	-	-	-	-	1.62E-29
ไม่ใช้	0	183	18	165	-
ใช้	1	49	45	4	-
18. อาการไข้สูง (> 37 °C)	-	-	-	-	3.31E-21
ไม่ใช้	0	185	24	161	-
ใช้	1	47	38	9	-
19. รู้สึกคลื่นไส้หรืออยากอาเจียน (≥ 2 วัน)	-	-	-	-	1.55E-23
ไม่ใช้	0	188	24	164	-
ใช้	1	44	39	5	-
20. อาการความเครียดสะสม	-	-	-	-	0.016
ไม่ใช้	0	120	24	96	-
ใช้	1	112	39	73	-
21. รู้สึกเบื่ออาหาร (≥ 2 วัน)	-	-	-	-	0.887
ไม่ใช้	0	192	53	139	-
ใช้	1	40	10	30	-

ตารางที่ 1: สถิติของปัจจัยในการทำนายโรคไส้ตั้งอีกเสบ (ต่อ)

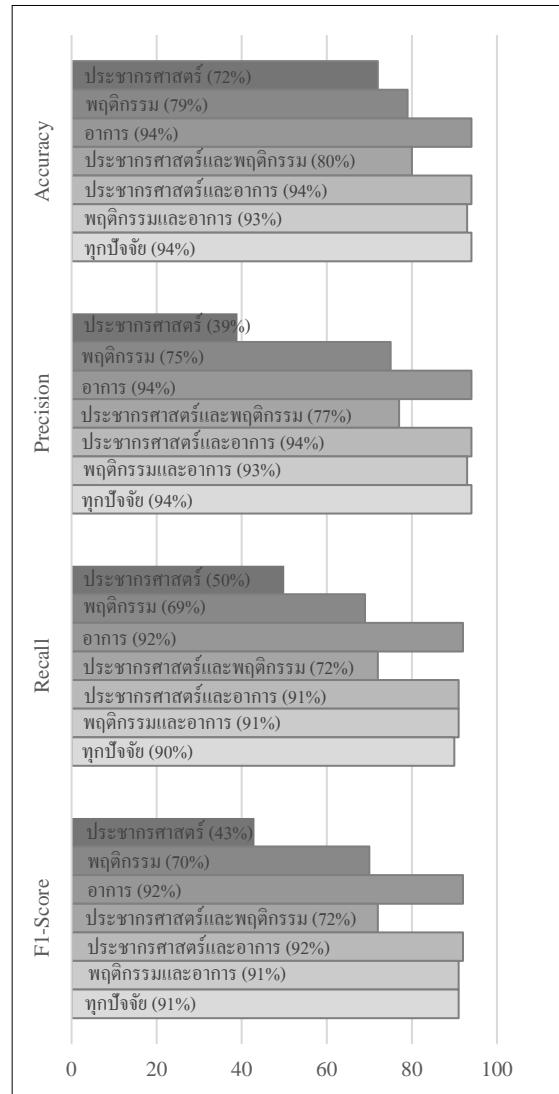
ปัจจัย	ค่าที่ แปลง	ความถี่			P-value
		จำนวน	เคย เป็น ไส้ตั้ง	ไม่เคย เป็น ไส้ตั้ง	
ด้านอาการ	-	-	-	-	-
22. อาการท้องผูก (≥ 3 วันต่อ สัปดาห์)	-	-	-	-	1.20E-11
ไม่ใช้	0	143	16	127	-
ใช้	1	89	47	42	-

จากตารางที่ 1 พบว่ามีทั้งหมด 11 ปัจจัยที่ผลการทดสอบ Chi-Square ให้ค่า P-value ต่ำกว่า 0.01 ที่ระดับความเชื่อมั่นร้อยละ 99 ซึ่งปัจจัยเหล่านี้แสดงถึงความสามารถในการแยกความแตกต่างระหว่างผู้ที่เป็นโรคไส้ตั้งอีกเสบและไม่เป็นโรคไส้ตั้งอีกเสบ ทั้ง 11 ปัจจัย มี 6 ปัจจัย อยู่ในด้านพฤติกรรม ได้แก่ การรับประทานอาหารจำพวกของหวาน, การรับประทานอาหารจำพวกกลุ่มโซเดียม, การรับประทานเมล็ดของผลไม้, การดื่มน้ำ, การดื่มเครื่องดื่มที่มีคาเฟอีน, และการออกกำลังกาย มี 5 ปัจจัย อยู่ในด้านอาการ ได้แก่ อาการปวดท้องบริเวณรอบสะดือแล้วย้ายมาปวดบริเวณหน้าท้องด้านขวาตอนล่าง, รู้สึกเหมือนมีก้อนแข็ง ๆ เมื่อกดบริเวณท้องน้อยขวา, อาการไข้สูง, รู้สึกคลื่นไส้หรืออยากอาเจียน, และอาการท้องผูก แม้ว่าปัจจัยด้านพฤติกรรมที่มีค่า P-value ต่ำกว่า 0.01 มีจำนวนมากกว่าปัจจัยด้านอาการ (6 ต่อ 5) แต่ส่วนมากปัจจัยด้านอาการมีค่า P-value ต่ำกว่า 0.01 อย่างมากเมื่อเทียบกับปัจจัยด้านพฤติกรรม ซึ่งพบว่า 5 ปัจจัยด้านอาการข้างต้นมีค่า P-value อยู่ในช่วง 4.46E-40 ถึง 1.20E-11 ขณะที่ 6 ปัจจัยด้านพฤติกรรมข้างต้นมีค่า P-value อยู่ในช่วง 1.21E-09 ถึง 0.0009 ในส่วนของปัจจัยด้านประชากรศาสตร์ไม่มีปัจจัยไหนที่มีค่า P-value ต่ำกว่า 0.01

ในตารางที่ 2 แสดงประสิทธิภาพการทำงานของโมเดลคอตมันน์แรกแสดงปัจจัยที่ใช้ในการสร้างโมเดล คอตมันน์ที่ 2 แสดงเมตริกในการประเมินประสิทธิภาพของโมเดลประกอบด้วย Accuracy, Precision, Recall, และ F1-Score คอตมันน์ที่ 3-8 แสดงเทคนิคการเรียนรู้ของเครื่องตัวเลขที่ขีดเส้นใต้และอยู่ในช่องที่บแสดงถึงค่าของเมตริก

ตารางที่ 2: ประสิทธิภาพการทำงานของโมเดล

ปัจจัย	เมตริก	LR	NB	DT	KNN	SVM	MLP
ด้าน ประชากรศาสตร์	Accuracy	71	70	70	*72	*72	*72
	Precision	36	54	42	*62	36	36
	Recall	49	51	50	*57	50	50
	F1-Score	41	49	43	*56	42	42
ด้านพฤติกรรม	Accuracy	79	*82	73	77	78	79
	Precision	77	*79	68	72	75	75
	Recall	69	*81	67	71	68	69
	F1-Score	70	*79	67	70	68	69
ด้านอาการ	Accuracy	94	94	90	94	*95	94
	Precision	94	*95	92	*95	94	93
	Recall	91	92	86	91	*94	92
	F1-Score	92	92	87	92	*94	92
ด้าน ประชากรศาสตร์ และด้านพฤติกรรม	Accuracy	81	*83	69	75	80	79
	Precision	78	*79	65	69	77	76
	Recall	74	*81	66	70	71	72
	F1-Score	75	*79	64	69	72	71
ด้าน ประชากรศาสตร์ และด้านอาการ	Accuracy	94	94	93	93	*95	93
	Precision	94	*95	93	94	94	93
	Recall	91	92	90	88	*94	91
	F1-Score	92	92	90	90	*94	91
ด้านพฤติกรรมและ ด้านอาการ	Accuracy	93	92	91	*94	*94	93
	Precision	93	92	90	*96	94	92
	Recall	91	88	90	89	*93	91
	F1-Score	91	89	89	91	*93	92
ทุกปัจจัย	Accuracy	*94	92	91	93	*94	*94
	Precision	94	92	90	*95	93	94
	Recall	91	88	89	89	*92	*92
	F1-Score	*92	89	89	90	*92	*92

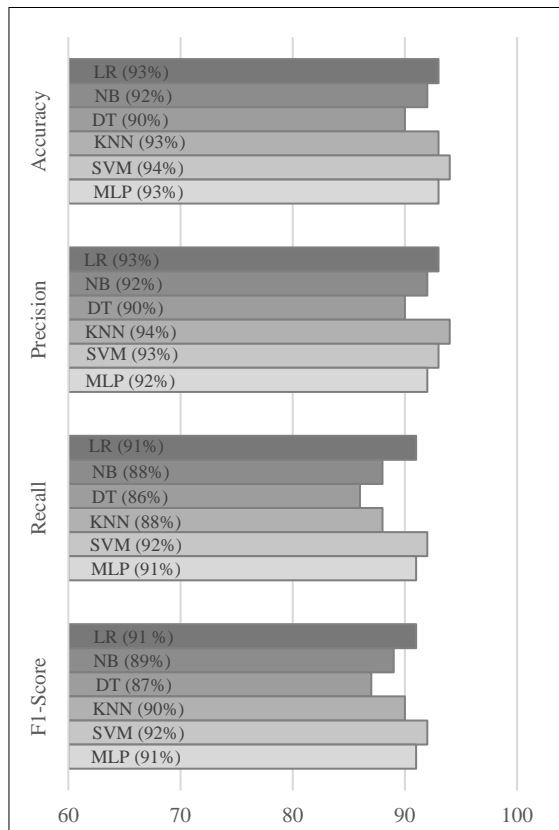


ภาพที่ 1: เปรียบเทียบประสิทธิภาพของปัจจัย

ที่มีค่าดีที่สุดจาก 7 กลุ่มของปัจจัย และตัวเลขที่มีเครื่องหมายดอกจันนำหน้าแสดงถึงค่าของเมตริกที่ดีที่สุดจาก 6 เทคนิคการเรียนรู้ของเครื่อง

จากตารางที่ 2 เมื่อพิจารณาปัจจัยแต่ละด้านอย่างอิสระ โดยไม่มีการรวมปัจจัยจากด้านต่าง ๆ เข้าด้วยกัน พบว่า หากใช้ปัจจัยด้านอาการเพียงอย่างเดียวให้ประสิทธิภาพดีที่สุด (Accuracy ร้อยละ 90-95, Precision ร้อยละ 92-95, Recall ร้อยละ 86-94, และ F1-score ร้อยละ 87-94) รองลงมาเป็นปัจจัยด้านพฤติกรรม (Accuracy ร้อยละ 73-82, Precision ร้อยละ 68-79, Recall ร้อยละ 67-81, และ F1-score ร้อยละ 67-79) และปัจจัยด้านประชากรศาสตร์ (Accuracy ร้อยละ 70-72, Precision ร้อยละ 36-62, Recall ร้อยละ 49-57, และ F1-score ร้อยละ 41-56) ซึ่ง

สอดคล้องกับค่า P-value ของปัจจัย โดยที่ปัจจัยด้านอาการ มีค่า P-value ต่ำที่สุด รองลงมาเป็นด้านพฤติกรรม และด้านประชากรศาสตร์ ตามลำดับ เมื่อพิจารณาจากการนำปัจจัยด้านต่าง ๆ มารวมกัน พบว่า การรวมกันของปัจจัยด้านประชากรศาสตร์และด้านอาการให้ประสิทธิภาพดีที่สุด อย่างไรก็ตาม ประสิทธิภาพจากการรวมกันนี้ไม่ได้ทำให้ประสิทธิภาพสูงขึ้นจากการใช้เพียงปัจจัยด้านอาการอย่างเดียว เนื่องมาจากการค้นพบข้างต้นที่ว่า การใช้ปัจจัยด้านประชากรศาสตร์เพียงอย่างเดียวให้ประสิทธิภาพก่อนข้างต่ำ โดยเฉพาะค่า Precision, Recall, และ F1-Score



ภาพที่ 2: เปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่อง

ในกลุ่มของเทคนิคการเรียนรู้ของเครื่อง SVM ให้ประสิทธิภาพดีที่สุด โดยเฉพาะอย่างยิ่งเมื่อนำ SVM มาสร้างโมเดลด้วยปัจจัยด้านอาการ มีค่า Accuracy ร้อยละ 95, Precision ร้อยละ 94, Recall ร้อยละ 94, และ F1-score ร้อยละ 94 รองลงมาเป็น NB, KNN, MLP, LR, และ DT ตามลำดับ ภาพที่ 1 แสดงผลสรุปการเปรียบเทียบประสิทธิภาพของปัจจัยด้วยค่ากลาง (Median) ของเมตริก และภาพที่ 2 แสดงผลสรุปการเปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่อง

5. การอภิปราย

จากผลการทดลอง พบว่า ปัจจัยที่มีค่า P-value ต่ำ มีความสำคัญต่อการสร้างโมเดลที่มีประสิทธิภาพสูง ซึ่งปัจจัยด้านอาการมักจะมีค่า P-value ที่ต่ำมาก รองลงมาเป็นปัจจัยด้านพฤติกรรม และปัจจัยด้านประชากรศาสตร์

ในการเปรียบเทียบประสิทธิภาพของปัจจัย พบว่า เมื่อนำปัจจัยด้านอาการเพียงอย่างเดียวไปใช้สร้างโมเดลจะให้ประสิทธิภาพสูงสุด รองลงมาเป็นปัจจัยด้านพฤติกรรม และปัจจัยด้านประชากรศาสตร์ ตามลำดับ โดยประสิทธิภาพดังกล่าวสอดคล้องกับค่า P-value ตามที่ได้อธิบายข้างต้น เป็นที่น่าสังเกตว่า แม้ว่าปัจจัยด้านพฤติกรรมไม่ได้ให้ประสิทธิภาพสูงสุด แต่สามารถให้ประสิทธิภาพที่อยู่ในระดับที่น่าพอใจ ซึ่งเป็นประโยชน์อย่างมากต่อประชาชนทั่วไป ในการป้องกันและตรวจสอบความเสี่ยงของการเกิดโรคได้ตั้งอีกเสบจากพฤติกรรมในชีวิตประจำวัน โดยเฉพาะพฤติกรรมการกิน และแม้ว่าปัจจัยด้านอาการจะให้ประสิทธิภาพสูงสุด แต่อาการส่วนใหญ่จะแสดงเมื่อได้ตั้งเริ่มอีกเสบแล้ว เช่น อาการปวดท้อง ดังนั้นปัจจัยด้านอาการจึงเหมาะกับการวินิจฉัยเพื่อทำการรักษา ขณะที่การใช้ปัจจัยด้านประชากรศาสตร์เพียงอย่างเดียว ไม่สามารถให้ประสิทธิภาพที่น่าพอใจได้ นอกจากนี้ เมื่อรวมปัจจัยแต่ละด้านเข้าด้วยกันไม่สามารถทำให้มีประสิทธิภาพมากกว่าปัจจัยด้านอาการเพียงอย่างเดียว

ในการเปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่อง พบว่า เมื่อนำปัจจัยด้านอาการมาสร้างโมเดลด้วยเทคนิค SVM สามารถให้ประสิทธิภาพสูงที่สุด แม้ว่าเทคนิค NB, KNN, MLP, LR, และ DT ให้ประสิทธิภาพที่ต่ำกว่า SVM แต่ประสิทธิภาพยังอยู่ในระดับที่น่าพอใจ นั้นแสดงให้เห็นว่าการใช้เทคนิคการเรียนรู้ของเครื่องในการวินิจฉัยโรคได้ตั้งอีกเสบเป็นวิธีการที่มีประสิทธิภาพ นอกเหนือจากการใช้วิธีคะแนนพยากรณ์ การอัลตราซาวด์ และการเอกซเรย์คอมพิวเตอร์

6. สรุป

งานวิจัยนี้เสนอแนวทางในการพัฒนาโมเดลเพื่อวินิจฉัยโรคได้ตั้งอีกเสบ โดยใช้เทคนิคการเรียนรู้ของเครื่อง ผลการวิจัยพบว่า การใช้ปัจจัยด้านอาการให้ประสิทธิภาพสูงสุด มีค่า Accuracy ร้อยละ 90-95, Precision ร้อยละ 92-95, Recall ร้อยละ 86-94, และ F1-score ร้อยละ 87-94 และพบว่า SVM คือเทคนิคที่ให้

ประสิทธิภาพสูงสุดโดยเฉพาะอย่างยิ่งเมื่อนำ SVM มาสร้างโมเดลด้วยปัจจัยด้านอาการ มีค่า Accuracy ร้อยละ 95, Precision ร้อยละ 94, Recall ร้อยละ 94, และ F1-score ร้อยละ 94

จากผลการวิจัยชี้ให้เห็นว่าแนวทางในการพัฒนาโมเดลโดยใช้เทคนิค SVM ร่วมกับปัจจัยด้านอาการสามารถนำไปใช้ในการวินิจฉัยโรคไส้ติ่งอักเสบได้อย่างมีประสิทธิภาพ ซึ่งเป็นประโยชน์ต่อบุคลากรทางการแพทย์ในการวินิจฉัยโรค และบุคคลทั่วไปในการตรวจสอบความเสี่ยงการเกิดไส้ติ่งอักเสบ ทั้งนี้ เทคนิคการเรียนรู้เชิงลึกและการประมวลผลภาพจะถูกนำมาใช้สำหรับการปรับปรุงประสิทธิภาพของโมเดลในอนาคต

เอกสารอ้างอิง

- [1] พิเชษฐ พืดขุนทด, “ความแม่นยำในการวินิจฉัยโรคไส้ติ่งอักเสบโดยใช้ระบบคะแนน Modified Alvarado ในโรงพยาบาลละหานทราย จังหวัดบุรีรัมย์” *วารสารการแพทย์โรงพยาบาลศรีสะเกษ สุรินทร์ บุรีรัมย์* ปีที่ 26 ฉบับที่ 1 มกราคม-เมษายน พ.ศ. 2554 หน้า 149-156.
- [2] ชนิกา มหารักษ์, “การวินิจฉัยอื่นในผู้ป่วยที่สงสัยโรคไส้ติ่งอักเสบเฉียบพลันและมีการตรวจด้วยเอกซเรย์คอมพิวเตอร์ช่องท้องในโรงพยาบาลพังงา” *วารสารวิชาการแพทย์เขต 11* ปีที่ 33 ฉบับที่ 2 เมษายน-มิถุนายน พ.ศ. 2562 หน้า 211-222.
- [3] คำรงค์ ชูไพฑูรย์, “ปัจจัยที่สัมพันธ์กับการแตกของไส้ติ่งอักเสบในผู้ป่วยโรงพยาบาลปราณบุรี” *วารสารหัวหินสุขใจ ไกลกังวล* ปีที่ 4 ฉบับที่ 3 กันยายน-ธันวาคม พ.ศ. 2562.
- [4] เสาวนีย์ เกิดปากแพก และ สุชาดา วิภาคนันต์, “การจัดการความปวดของผู้ป่วยหลังผ่าตัดไส้ติ่งอักเสบที่ไม่มีภาวะไส้ติ่งอักเสบแตกในห้องผ่าตัด แผนกวิสัญญี โรงพยาบาลกระบี่” *วารสารการแพทย์กระบี่* ปีที่ 2 ฉบับที่ 1 พ.ศ. 2561-2562 หน้า 15-27.
- [5] อลิสรดา วงศ์สุทธิเลิศ, ศรสุภา ลิมเจริญ, และ อติสรณ์ บุญญฤทธิ์, “ประสิทธิภาพในการประเมินภาวะไส้ติ่งอักเสบด้วยอัลตราซาวด์ในกลุ่มผู้ป่วยเด็ก” *คณะแพทยศาสตร์มหาวิทยาลัยบูรพา* พ.ศ. 2557.
- [6] เนตรา เนตรโพธิ์แก้ว, “การประเมินผลของการใช้ระบบคะแนน Alvarado score ในการวินิจฉัยโรคไส้ติ่งอักเสบ” *วารสารสมาคมศัลยแพทย์ทั่วไปแห่งประเทศไทยในพระบรมราชูปถัมภ์* ปีที่ 2 ฉบับที่ 2 พ.ศ. 2560 หน้า 3-16.
- [7] เนตรา เนตรโพธิ์แก้ว, “การประเมินผลของ Alvarado score และ RIPASA score ในการวินิจฉัยโรคไส้ติ่งอักเสบ” *วารสารโรงพยาบาลมหาสารคาม* ปีที่ 15 ฉบับที่ 3 กันยายน-ธันวาคม พ.ศ. 2561 หน้า 55-64.
- [8] ปรียาภรณ์ เสนะวัต และ กิจจา เจียรวัฒนกกน, “คะแนนพยากรณ์ภาวะไส้ติ่งแตกในผู้ป่วยไส้ติ่งอักเสบเฉียบพลันโดยใช้อาการทางคลินิก” *วารสารโรงพยาบาลนครพิงค์* ปีที่ 10 ฉบับที่ 2 กรกฎาคม-ธันวาคม พ.ศ. 2562 หน้า 16-30.
- [9] J. Iqbal, R. Sayani, M. Tahir, and SM. Mustahsan, “Diagnostic Efficiency of Multidetector Computed Tomography in the Evaluation of Clinically Equivocal Cases of Acute Appendicitis with Surgical Correlation,” *Cureus*, vol. 10, no. 3, 2018.
- [10] Miroslav Kubat, *An Introduction to Machine Learning*, Springer International Publishing AG: Switzerland, 2017.