
การพัฒนาตัวแบบทำนายสัญญาณเตือนโรคมะเร็งลำไส้ใหญ่และทวารหนัก Development of a models to predict colorectal cancer warning sign

*วราพร เพ็ชรวิเศษ

*Waraporn Petwised

โรงพยาบาลนครปฐม สำนักงานสาธารณสุขจังหวัดนครปฐม สำนักงานปลัดกระทรวงสาธารณสุข
Nakhonpathom Hospital, Nakhonpathom Provincial Public Health Office, Office of the
Permanent Secretary Ministry of Public Health

*ผู้นิพนธ์หลัก: petwised_joom@hotmail.com

*Corresponding author: e-mail petwised_joom@hotmail.com

Received	Reviewed	Revised	Accepted
27/03/2023	30/03/2023	22/04/2023	28/04/2023

Abstract

This research aims to study the most suitable predictive model and to develop a risk assessment model to be used as a warning sign for colorectal cancer. There are 5 classification algorithms, namely Naive Bayes Classifier, Decision Tree, Random Forest, Deep Learning and Logistic Regression Analysis. By comparing the prediction performance, Accuracy, Precision, Recall, and Mean Square Error (RMSE). This study collected data of 396 patients in Nakhonpathom Hospital with 7 attributes. In the research method, the data was divided into 5 parts or 5 Fold Cross-Validation by randomly dividing the data into 5 parts and selecting 4 parts of data called training data to create a model for both 5 algorithms and randomly selecting 1 part of data called validation data (Testing Data) is used to assess the error of the model. The results of the research showed that the Logistic Regression Analysis method, the accuracy, the precision, the recall and the RMSE were 92.41% 0.90 1.0 and 0.24, respectively, because it is a predictive technique, it will be used when the dependent variable is a quality variable. As for the independent variables or predictive variables, they can be both qualitative (sex) and quantitative variables with continuous values (age and Carcinoembryonic Antigen (CEA), Haemoglobin (HGB), White Blood Cells (WBC), Haematocrit (Hct), Thrombocyte (Plt). Therefore, Logistic Regression Analysis method has the best predictive efficiency.

Keyword: Colorectal cancer, Predictive model, Risk assessment, Logistic regression

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาตัวแบบทำนายที่เหมาะสมที่สุด และนำไปพัฒนารูปแบบการประเมินความเสี่ยงใช้เป็นสัญญาณเตือนการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก ด้วยวิธีการจำแนกกลุ่มมี 5 อัลกอริทึม ได้แก่ Naive Bayes Classifier, Decision Tree, Random Forest, Deep Learning และ Logistic Regression Analysis โดยการเปรียบเทียบประสิทธิภาพการทำนายจะพิจารณาค่าความถูกต้อง (Accuracy) ค่า Precision ค่า Recall และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) การศึกษาครั้งนี้ได้รวบรวมข้อมูลของผู้ป่วยในโรงพยาบาลนครปฐม จำนวน 396 ตัวอย่าง มี 7 แอททริบิวต์ ในส่วนวิธีการวิจัยจะใช้วิธีแบ่งข้อมูลออกเป็น 5 ส่วน หรือ 5 Fold Cross-Validation โดยการสุ่มข้อมูลเพื่อแบ่งข้อมูลออกเป็น 5 ส่วน และทำการเลือกข้อมูล 4 ส่วนเรียกว่าข้อมูลการเรียนรู้ (Training Data) นำมาสร้างตัวแบบ ทั้ง 5 อัลกอริทึม และสุ่มเลือกข้อมูล 1 ส่วนเรียกว่าข้อมูลตรวจสอบ (Testing Data) นำไปประเมินความผิดพลาดของตัวแบบ ผลการวิจัยพบว่าวิธี Logistic Regression Analysis ค่าความถูกต้อง ค่า Precision ค่า Recall และค่า RMSE คือ 92.41% 0.90 1.0 และ 0.24 ตามลำดับ เนื่องจากเป็นเทคนิคเพื่อการทำนายจะใช้เมื่อตัวแปรตาม (Dependent Variable) เป็นตัวแปรคุณภาพ ส่วนตัวแปรอิสระ (Independent Variable) หรือตัวแปรทำนายเป็นได้ทั้งตัวแปรคุณภาพ (เพศ) และตัวแปรเชิงปริมาณที่มีค่าต่อเนื่อง (อายุ และ Carcinoembryonic Antigen (CEA), Haemoglobin (HGB), White Blood Cells (WBC), Haematocrit (Hct), Thrombocyte (Plt) ดังนั้นวิธี Logistic Regression Analysis จึงมีประสิทธิภาพในการทำนายผลดีที่สุด

คำสำคัญ: มะเร็งลำไส้ใหญ่และทวารหนัก, ตัวแบบทำนาย, การประเมินความเสี่ยง, การถดถอยโลจิสติก

บทนำ

มะเร็ง คือ กลุ่มของโรคที่เกิดจากเซลล์ของร่างกายที่มีความผิดปกติที่ DNA ส่งผลให้เซลล์มีการเจริญเติบโต มีการแบ่งตัวเพื่อเพิ่มจำนวนเซลล์ รวดเร็ว และมากกว่าปกติ ดังนั้นจึงอาจทำให้เกิดก้อนเนื้อผิดปกติ และในที่สุดก็จะทำให้เกิดการตายของเซลล์ในก้อนเนื้อนั้น เนื่องจากขาดเลือดไปเลี้ยง ถ้าเซลล์พวกนี้เกิดอยู่ในอวัยวะใดก็จะเรียกชื่อมะเร็งตามอวัยวะนั้น (Petpichetchian, W., 2011) หนึ่งในนั้นก็คือมะเร็งลำไส้ใหญ่และทวารหนัก

องค์การอนามัยโลกพบว่าปี 2561 มีผู้ป่วยมะเร็งรายใหม่จำนวน 18.1 ล้านคนและมีผู้เสียชีวิตจากโรคมะเร็ง 9.6 ล้านคน (National Cancer Prevention and Control Plan Committee Department of Medical Services Ministry of Public Health, 2016) ทั้งนี้โรคมะเร็งลำไส้ใหญ่พบมากเป็นอันดับ 3 มีจำนวน 1.8 ล้านคน รองจากมะเร็งปอด และมะเร็งเต้านม มีผู้เสียชีวิตจากมะเร็งลำไส้ 881,000 คน จาก 9.6 ล้านคน คิดเป็นร้อยละ 9.2 สำหรับประเทศไทยพบว่ามีผู้เสียชีวิตประมาณปีละ 3,000 ราย พบอัตราการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนักเพศชายเป็นอันดับ 3 (รองจากมะเร็งตับและมะเร็งปอด) มีอุบัติการณ์ คือ 8.8 ต่อประชากรแสนคน และเป็นอันดับ 5 ในเพศหญิง (รองจากมะเร็งปากมดลูก มะเร็งเต้านม มะเร็งตับและมะเร็งปอด) โดยมีอุบัติการณ์ 7.6 ต่อประชากรแสนคน อายุที่พบส่วนใหญ่มากกว่า 50 ปีขึ้นไป (Srivatanakul, P & Attasara, P., 2007)

มะเร็งลำไส้ใหญ่และทวารหนักมักมีอาการขับถ่ายผิดปกติ เช่นถ่ายอุจจาระเป็นเลือด ท้องผูก อุจจาระลำเล็ก ส่วนใหญ่พบในระยะที่เป็นมากแล้วซึ่งมีโอกาสเสียชีวิตสูง (Siriraj Piyamaharajkarun Hospital, 2021) โรคมะเร็งลำไส้เป็นโรคมะเร็งที่สามารถทำการตรวจหาได้ตั้งแต่วัยเริ่มแรก ซึ่งสามารถทำการรักษาให้หายได้และเสียค่าใช้จ่ายน้อยกว่าการรักษามะเร็งในระยะลุกลาม การตรวจคัดกรองเป็นวิธีการที่ดีที่สุดหากตรวจพบในระยะเริ่มแรกจะทำการรักษาได้ผลดีและเสียค่าใช้จ่ายน้อยกว่า

การคัดกรองกลุ่มเสี่ยงต่อการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนักเป็นวิธีการที่มีประสิทธิภาพในการลดอุบัติการณ์และอัตราการตายจากโรค สมาคมโรคมะเร็งแห่งสหรัฐอเมริกา (American Cancer Society, 2022) แนะนำให้ตรวจในคนอายุ 50 ปีขึ้นไป ทำได้หลายวิธีแต่ที่ทำการแพร่หลายมีหลัก ๆ 2 วิธี คือการตรวจเลือดออกแฝงในอุจจาระ (FIT Test) เป็นการตรวจหาเลือดแฝงในอุจจาระที่มีปริมาณน้อยมากเพื่อเป็นการคัดกรองเบื้องต้น และตรวจด้วยการส่องกล้องลำไส้ใหญ่ (Colonoscopy) เป็นการตรวจเพื่อหาติ่งเนื้อในลำไส้ใหญ่ (Polyps) ซึ่งเป็นรอยก่อโรค

ปัจจุบันมีการสร้างแบบจำลองถือเป็นวิธีการหนึ่งที่น่าเทคโนโลยีทางด้านคอมพิวเตอร์เข้ามาประยุกต์ใช้ทำนายการเกิดโรคเป็นสัญญาณเตือนภัยสุขภาพ ป้องกันผลกระทบต่อสุขภาพทั้งในประชาชนทั่วไปและกลุ่มเสี่ยง ลดภาระค่าใช้จ่ายในการรักษา จากการค้นคว้าและทบทวนงานวิจัยที่เกี่ยวข้อง ผู้วิจัยจึงมีความสนใจที่จะศึกษาเพื่อเปรียบเทียบประสิทธิภาพของเทคนิควิธี Naive Bayes Classifier, Decision Tree, Random Forest, Deep Learning และ Logistic Regression Analysis โดยนำเทคนิคทั้ง 5 วิธีมาทำนายการเกิดโรคมะเร็งลำไส้ว่าตัวแบบใดมีประสิทธิภาพการสังเคราะห์หรือการเรียนรู้ที่ดีที่สุด เพื่อเปรียบเทียบประสิทธิภาพตัวแบบทำนายในการทำนายวิธีนี้สามารถนำไปใช้ได้ทั้งกลุ่มเสี่ยงและกลุ่มที่ไม่มีความเสี่ยงสามารถใช้ตัวแบบทำนายได้ด้วยตนเอง ซึ่งแตกต่างจากการคัดกรองที่มีการตรวจในกลุ่มเสี่ยงเพื่อค้นหาความผิดปกติที่อาจจะเป็นมะเร็ง วิธีทำนายจะส่งผลต่อการค้นหาผู้ป่วยมะเร็งลำไส้ใหญ่และทวารหนักได้มากขึ้นป้องกันมิให้ผู้ป่วยเป็นโรคมะเร็งในระยะลุกลามเพิ่มอัตราการรอดชีวิต ทั้งนี้ ผู้วิจัยได้ศึกษาข้อมูลการเกิดโรคมะเร็งของจังหวัดนครปฐมซึ่งมีอัตราการเสียชีวิตด้วยโรคมะเร็งและเนื้องอกเป็นอันดับ 1 ของจังหวัด และพบว่ามะเร็งลำไส้ใหญ่และทวารหนักมีอัตราการเสียชีวิตย้อนหลัง 5 ปี 11.26 10.85 6.79 5.4 และ 8.0 ต่อแสนประชากร ในปี 2560 2561 2562 2563 และ 2564 ตามลำดับ (Public Health Office Nakhonpathom, 2022) ซึ่งมีอัตราการเสียชีวิตเป็นอันดับ 3 ในกลุ่มโรคมะเร็ง (รองจาก 1.มะเร็งหลอดคอ, หลอดลมใหญ่ ปอด 2.มะเร็งตับและท่อน้ำดี) ดังนั้นหากมีเครื่องมือที่สามารถทำนายความเสี่ยงการเกิดโรค และพบว่ามีความเสี่ยงควรเข้ารับการตรวจคัดกรองหรือพบแพทย์ทันทีเพราะสำหรับมะเร็งลำไส้หากรู้เร็ว รักษาเร็วและถูกวิธีจะมีโอกาสหายมากขึ้น

วิธีการวิจัย

1. การวิจัยนี้ได้นำข้อมูลส่วนบุคคลประกอบด้วย เพศ อายุ และข้อมูลสุขภาพที่ได้จากห้องปฏิบัติการประกอบด้วย CEA, HGB, WBC, Hct, Plt มาใช้เป็นตัวแปรในการวิเคราะห์ข้อมูล โดยมีขั้นตอนดัง Figure 1

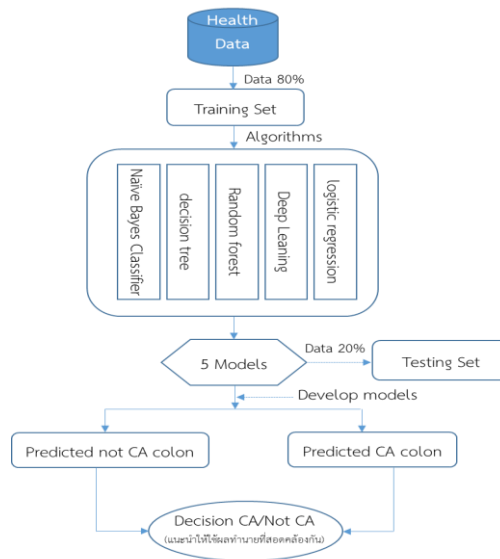


Figure 1 Research process

ขั้นตอนการดำเนินการวิจัย

- กำหนดตัวแปร โดยการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง
- ตรวจสอบข้อมูลให้มีความถูกต้อง และแก้ไขข้อมูลที่มีค่าผิดปกติ (Data Cleaning)
- แบ่งข้อมูลที่ใช้ในการวิจัยเพื่อนำไปใช้ในการทดสอบประสิทธิภาพด้วยวิธี Cross-Validation Test ซึ่งเป็นวิธีที่ได้รับความนิยมในการวิจัย โดยผู้วิจัยกำหนดการแบ่งข้อมูลออกเป็น 5 ส่วน หรือ 5 Fold Cross-Validation โดยการสุ่มข้อมูลเพื่อแบ่งข้อมูลออกเป็น 5 ส่วน และทำการเลือกข้อมูล 4 ส่วนเรียกว่าข้อมูลการเรียนรู้ นำมาสร้างตัวแบบ ทั้ง 5 อัลกอริทึม และสุ่มเลือกข้อมูล 1 ส่วนเรียกว่าข้อมูลตรวจสอบนำไปประเมินความผิดพลาดของตัวแบบ โดยมีการทดสอบจำนวน 5 รอบ ดังนี้

รอบที่ 1 นำข้อมูลส่วนที่ 1 2 3 และ 4 นำมาเรียนรู้สร้างตัวแบบ ข้อมูลส่วนที่ 5 นำมาตรวจสอบตัวแบบ

รอบที่ 2 นำข้อมูลส่วนที่ 1 2 3 และ 5 นำมาเรียนรู้สร้างตัวแบบ ข้อมูลส่วนที่ 4 นำมาตรวจสอบตัวแบบ

รอบที่ 3 นำข้อมูลส่วนที่ 1 2 4 และ 5 นำมาเรียนรู้สร้างตัวแบบ ข้อมูลส่วนที่ 3 นำมาตรวจสอบตัวแบบ

รอบที่ 4 นำข้อมูลส่วนที่ 1 3 4 และ 5 นำมาเรียนรู้สร้างตัวแบบ ข้อมูลส่วนที่ 2 นำมาตรวจสอบตัวแบบ

รอบที่ 5 นำข้อมูลส่วนที่ 2 3 4 และ 5 นำมาเรียนรู้สร้างตัวแบบ ข้อมูลส่วนที่ 1 นำมาตรวจสอบตัวแบบ

4. วัดประสิทธิภาพของตัวแบบทำนาย

5. พัฒนาตัวแบบในระบบสนับสนุนการตัดสินใจ ซึ่งมี 5 อัลกอริทึม การตัดสินใจว่าบุคคลหรือข้อมูลชุดนั้นมีโอกาสเสี่ยงในการเกิดโรคมะเร็งลำไส้โดยใช้ผลทำนายของแต่ละอัลกอริทึมที่สอดคล้องกันส่วนใหญ่

ระเบียบวิธีวิจัย

วิธีวิจัยเป็นการวิจัยประยุกต์ (Applied Research) โดยใช้ปัจจัยที่มีความเสี่ยงในการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนักมาทำนายผ่านอัลกอริทึมทั้ง 5 อัลกอริทึม โดยผู้วิจัยเก็บข้อมูลผู้ป่วยจากการศึกษาเวชระเบียน ทั้งผู้ที่ได้รับการวินิจฉัยเป็นโรคมะเร็งลำไส้ทุกรายและไม่ใช่โรคมะเร็งลำไส้ใหญ่และทวารหนักที่มีข้อมูลทั่วไปและข้อมูลทางห้องปฏิบัติการที่เข้ารับการตรวจและรักษาในโรงพยาบาลนครปฐม ตั้งแต่เดือนมกราคม ปี 2563 ถึงเดือนตุลาคม ปี 2565 ซึ่งมีรายละเอียดดังนี้

1. การเก็บข้อมูลทั่วไป ได้แก่ เพศ อายุ

2. การเก็บข้อมูลทางห้องปฏิบัติการ ได้แก่ CEA, HGB, WBC, Hct, Plt โดยมีความหมายดังนี้
สารบ่งชี้มะเร็ง (Tumor Marker) คือ สารชีวหรือเคมีชีวที่สร้างจากเซลล์มะเร็งและหลั่งออกมาในกระแสเลือดหรือสารคัดหลั่ง (Biological Fluid) ตำแหน่งต่าง ๆ

CEA เป็นสารมักขึ้นสูงผิดปกติในผู้ป่วยมะเร็งชนิด Adenocarcinoma หรือมะเร็งต่อม เช่น ของระบบทางเดินอาหาร มะเร็งตับอ่อน มะเร็งเต้านม มะเร็งปอด มะเร็งรังไข่ มะเร็งตับ มะเร็งปากมดลูก มะเร็งกระเพาะปัสสาวะ มะเร็งไต มะเร็งต่อมไทรอยด์ ฯลฯ โดยเฉพาะมะเร็งลำไส้ใหญ่ จะพบ CEA สูงได้มากและบ่อยกว่ามะเร็งชนิดอื่น ๆ

สำหรับมะเร็งลำไส้ใหญ่ค่าของ CEA ปกติอยู่ระหว่าง 2.5-5 ng/mL ค่าที่เพิ่มขึ้นมากกว่า 5 ng/mL มีความสอดคล้องกับการพยากรณ์โรคที่ไม่ดี เมื่อเทียบกับผู้ป่วยมะเร็งลำไส้ใหญ่ที่มีค่าดังกล่าวในระดับที่ต่ำกว่า (Parinyanitikul, N., 2017)

เลือด ประกอบด้วยเม็ดเลือดลอยอยู่ในพลาสมา มีสีแดงเพราะมีปริมาณเม็ดเลือดแดงเป็นองค์ประกอบ ทำหน้าที่ตัวกลางติดต่อระหว่างเซลล์ของร่างกาย ขนส่งออกซิเจนและอาหารไปยังส่วนต่าง ๆ ของร่างกาย นำคาร์บอนไดออกไซด์ไปขับออกทางปอด และนำของเสียต่าง ๆ เพื่อขับออกทางไต นอกจากนี้เลือดยังเป็นระบบป้องกันด้วยระบบภูมิคุ้มกันของร่างกายและช่วยในการควบคุมอุณหภูมิของร่างกายอีกด้วย

การตรวจความสมบูรณ์ของเม็ดเลือด (Complete Blood Count, CBC) เป็นการตรวจปริมาณและลักษณะของเม็ดเลือดทั้งสามชนิด ได้แก่ เม็ดเลือดแดง เม็ดเลือดขาว และเกล็ดเลือด โดยเป็นการตรวจพื้นฐานทั้งเพื่อการตรวจร่างกายประจำปี และการตรวจเพื่อดูแลรักษาผู้ป่วย หากพบความผิดปกติจะสามารถนำไปปรึกษาแพทย์เพื่อรับการดูแลรักษาต่อไป โดยสามารถจำแนกลักษณะ

ของเม็ดเลือดทั้ง 3 ชนิด ได้ดังนี้

1. เม็ดเลือดแดง มีรูปร่างคล้ายโดนัท เนื่องจากตรงกลางมีรอยบุ๋มแต่ไม่ทะลุตรงกลาง มีขนาดประมาณ 7 ไมครอน ถูกสร้างที่ไขกระดูก ซึ่งเป็นแกนกลางของกระดูกต่าง ๆ ทั่วร่างกาย ภายในเม็ดเลือดแดงสารชื่อ ฮีโมโกลบิน ทำหน้าที่ในการนำออกซิเจนจากปอดไปยังส่วนต่าง ๆ ของร่างกาย และนำคาร์บอนไดออกไซด์จากส่วนต่าง ๆ ของร่างกายไปขับออกที่ปอด ทำให้ในหลอดเลือดแดงที่เม็ดเลือดแดงมีระดับออกซิเจนปริมาณสูง เลือดจึงมีสีแดงสด ส่วนในหลอดเลือดดำที่เม็ดเลือดแดงมีปริมาณออกซิเจนลดลง และมีปริมาณคาร์บอนไดออกไซด์สูงกว่าเลือดจึงมีสีแดงคล้ำ การตรวจเม็ดเลือดแดงจะแสดงผลการตรวจเม็ดเลือดแดง ได้แก่

1.1 ระดับฮีโมโกลบิน (Hemoglobin) เป็นการตรวจความเข้มข้นของระดับฮีโมโกลบินในเลือด เนื่องจากสารฮีโมโกลบินอยู่ในเม็ดเลือดแดง ทำให้ระดับฮีโมโกลบินสัมพันธ์กันโดยตรงกับปริมาณเม็ดเลือดแดงในเลือดและปริมาตรเม็ดเลือดแดงอัดแน่น (Hematocrit) เพศชายมีค่าปกติ 13-17 และเพศหญิงมีค่าปกติ 12-15

- ระดับฮีโมโกลบินที่ลดลงแสดงถึงภาวะโลหิตจาง ซึ่งเกิดได้หลายสาเหตุ เช่น จากการเสียเลือด การขาดสารอาหารโดยเฉพาะธาตุเหล็กหรือวิตามินบี 12 และโฟเลต โรคเลือดทางพันธุกรรมเช่น ธาลัสซีเมีย โรคของไขกระดูกทำให้ผลิตเม็ดเลือดแดงได้ลดลง รวมถึงมะเร็งลำไส้ใหญ่และทวารหนัก (National Cancer Institute Department of Medical Services Ministry of Public Health, n.p.) ดังนั้น หากพบระดับฮีโมโกลบินที่ลดลง ควรปรึกษาแพทย์เพื่อสืบค้นสาเหตุ และให้การรักษาย่างเหมาะสมตามสาเหตุ การให้เลือดเป็นเพียงการแก้ไขภาวะโลหิตจางชั่วคราวในรายที่มีภาวะโลหิตจางรุนแรงเท่านั้น ไม่ได้รักษาสาเหตุของภาวะโลหิตจาง จึงควรสืบค้นหาสาเหตุเพื่อแก้ไขให้หายจากภาวะโลหิตจาง

- ส่วนระดับฮีโมโกลบินสูง พบได้ในภาวะขาดสารน้ำในร่างกาย หรือโรคเลือดข้นจากการขาดออกซิเจนเรื้อรังหรือไขกระดูกทำงานผิดปกติก็ได้

1.2 ปริมาตรเม็ดเลือดแดงอัดแน่น เป็นการตรวจปริมาตรของเซลล์เม็ดเลือดแดงในเลือด ซึ่งมีความสัมพันธ์โดยตรงกับจำนวนเม็ดเลือดแดงในเลือดและระดับฮีโมโกลบิน สาเหตุที่ทำให้ระดับต่ำหรือสูงขึ้น มีสาเหตุเช่นเดียวกับปริมาณฮีโมโกลบินที่ผิดปกติไป ดังที่กล่าวข้างต้น เพศชายมีค่าปกติ 38-50 และเพศหญิงมีค่าปกติ 36-45

2. เม็ดเลือดขาว มีขนาดใหญ่กว่าเซลล์เม็ดเลือดแดง จำนวนเม็ดเลือดขาวปกติประมาณ 4,000 – 10,000 เซลล์ต่อ 1 ไมโครลิตร จำนวนเม็ดเลือดขาวมีการเปลี่ยนแปลงขึ้นลงสัมพันธ์กับภาวะการอักเสบติดเชื้อของร่างกาย ถูกสร้างจากไขกระดูก มีหน้าที่ทำลายเชื้อโรคและสิ่งแปลกปลอมที่เข้าสู่ร่างกาย โดยจะใช้วิธีจับกินเชื้อโรคโดยตรง หรือผลิตภูมิคุ้มกันซึ่งเป็นโปรตีนที่มีคุณสมบัติต่อต้านเชื้อโรคและสิ่งแปลกปลอม

- ปริมาณเม็ดเลือดขาวต่ำลง โดยเฉพาะเม็ดเลือดขาวชนิดนิวโทรฟิลด์ ทำให้มีโอกาสติดเชื้อได้ง่าย สาเหตุที่ทำให้เม็ดเลือดขาวต่ำที่พบบ่อย ได้แก่ ยา สารเคมี รังสี ไวรัส และโรคของไขกระดูก เช่น ไขกระดูกทำงานผิดปกติ หรือไขกระดูกติดเชื้อหรือถูกแทรกซึมจากมะเร็ง เป็นต้น

- ปริมาณเม็ดเลือดขาวสูงขึ้น พบได้บ่อย ซึ่งเป็นภาวะที่ร่างกายตอบสนองต่อภาวะอักเสบหรือติดเชื้อในร่างกาย ปริมาณที่เพิ่มขึ้นมักไม่สูงมาก โดยทั่วไปมักไม่เกิน 30,000 เซลล์ต่อไมโครลิตร

3. เกล็ดเลือด เป็นส่วนของเม็ดเลือด จึงมีขนาดเล็กมากเมื่อเทียบกับเม็ดเลือดแดงและเม็ดเลือดขาว ทำหน้าที่ป้องกันเลือดออกจากหลอดเลือด ช่วยในการหยุดไหลของเลือดหากเกิดบาดแผล มีจำนวนประมาณ ห้าแสนถึงสี่แสนตัวต่อเลือดหนึ่งไมโครลิตร เกล็ดเลือดถูกสร้างจากเซลล์ในไขกระดูกเช่นกัน หากมีปริมาณลดลง โดยเฉพาะระดับที่ลดลงมากกว่าห้าหมื่น จะทำให้เลือดออกมากและนานกว่าเลือดจะหยุดหากเกิดบาดแผล โดยเฉพาะหากระดับเกล็ดเลือดต่ำกว่าสองหมื่น จะทำให้เลือดออกเองโดยไม่ต้องมีบาดแผลหรือการกระทบกระแทกใด โดยเฉพาะอย่างยิ่งบริเวณเยื่อบุต่าง ๆ เช่น เลือดออกตามไรฟัน เลือดกำเดาไหล หรือเลือดออกในอวัยวะสำคัญ ได้แก่ ทางเดินอาหาร หรือสมองได้ ค่าเกร็ดเลือดปกติมาตรฐานอยู่ที่ 130 – 400 cell/ml³

ประชากรที่ศึกษา กลุ่มตัวอย่าง

เวชระเบียนผู้ป่วยที่แพทย์วินิจฉัยวันแรกว่าเป็นมะเร็งลำไส้ใหญ่และทวารหนัก โดยใช้รหัส ICD-10 รหัส C18-21, C785, D12 และเวชระเบียนผู้ป่วยที่ไม่ได้เป็นมะเร็งลำไส้ใหญ่และทวารหนักและได้รับการตรวจทางห้องปฏิบัติการได้แก่ CEA, HGB, WBC, Hct, Plt ที่เข้ารับการตรวจและรักษาในโรงพยาบาลนครปฐม ตั้งแต่เดือนมกราคม ปี 2563 ถึงเดือนตุลาคม ปี 2565

ข้อพิจารณาด้านจริยธรรม

โครงการวิจัยได้รับการรับรองจริยธรรมการทำวิจัยในมนุษย์ จากคณะกรรมการจริยธรรมการวิจัยในมนุษย์โรงพยาบาลนครปฐม ได้รับอนุญาตให้ทำวิจัยในโรงพยาบาลจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์โรงพยาบาลนครปฐม เมื่อได้รับการอนุญาตแล้วผู้วิจัยจึงเก็บข้อมูลส่วนบุคคล ผลการตรวจเลือด ผลการตรวจวินิจฉัยของแพทย์ว่าเป็นมะเร็งลำไส้ใหญ่และทวารหนักและผลการตรวจวินิจฉัยผู้ป่วยที่ไม่ได้เป็นมะเร็งลำไส้ใหญ่และทวารหนักในแบบบันทึกที่ผู้วิจัยสร้างขึ้น โดยเปลี่ยนข้อมูลส่วนบุคคล และผลการตรวจวินิจฉัยของแพทย์ของกลุ่มตัวอย่างเป็นรหัส ข้อมูลส่วนตัวของผู้ป่วยจะเก็บไว้เป็นความลับ ไม่เปิดเผยต่อสาธารณะเป็นรายบุคคลแต่นำเสนอผลการศึกษาในภาพรวมและนำไปใช้ประโยชน์ในการพัฒนาเท่านั้น

การวัดประสิทธิภาพของตัวแบบทำนาย

งานวิจัยนี้ได้มีการวัดประสิทธิภาพและเปรียบเทียบผลการทำนายการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก เพื่อนำมาหาค่า Accuracy ค่า Precision และค่า Recall ได้จากสมการดังต่อไปนี้

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \text{Precision} = \frac{TP}{TP + FP}, \text{Recall} = \frac{TP}{TP + FN}$$

โดยที่ True Positive (TP) คือ จำนวนข้อมูลที่จำแนกถูกว่าเป็นชั้น A

True Negative (TN) คือ จำนวนข้อมูลที่จำแนกถูกว่าเป็นชั้น B

False Positive (FP) คือ จำนวนข้อมูลที่จำแนกผิดว่าเป็นชั้น A แต่ชั้นที่แท้จริงเป็นชั้น B

False Negative (FN) คือ จำนวนข้อมูลที่จำแนกผิดว่าเป็นชั้น B แต่ชั้นที่แท้จริงเป็นชั้น A

การประมวลผล

การประมวลผลครั้งนี้ใช้โปรแกรม WEKA (Waikato Environment For Knowledge Analysis) เป็นโปรแกรมที่สามารถดาวน์โหลดได้จากเว็บไซต์ อยู่ภายใต้การควบคุมของ GPL License โปรแกรมถูกพัฒนาจากภาษา Java เป็นโปรแกรมที่ใช้ในการวิเคราะห์ข้อมูลด้วยเทคนิคเหมืองข้อมูล เพื่อหาความสัมพันธ์รูปแบบและแยกประเภทของข้อมูล ซึ่งผู้วิจัยนำมาวิเคราะห์และประมวลผลเพื่อหาตัวแบบทำนายทั้ง 5 อัลกอริทึม

การพัฒนาตัวแบบในระบบสนับสนุนการตัดสินใจ

ในการพัฒนาระบบสนับสนุนการตัดสินใจเพื่อวิเคราะห์ข้อมูลสัญญาณเตือนในการเกิดโรคมะเร็งลำไส้และทวารหนัก มี 2 ขั้นตอนดังต่อไปนี้

1. การพัฒนาระบบ นำตัวแบบทั้ง 5 วิธีมาพัฒนาเป็นโปรแกรมเพื่อการตัดสินใจว่ามีความเสี่ยงในการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก การออกแบบหน้าจอแสดงผลและการพัฒนาตัวแบบ ผู้วิจัยได้ใช้ Microsoft Excel เป็นส่วนติดต่อกับผู้ใช้งานซึ่งพัฒนาด้วยภาษา Visual Basic For Applications (VBA) เพื่ออำนวยความสะดวกให้กับผู้ใช้งานระบบ

2. การแสดงผลและการตัดสินใจ ผู้ใช้งานป้อนข้อมูลทั่วไป ได้แก่ เพศ อายุ และ ข้อมูลทางห้องปฏิบัติการ ได้แก่ CEA, HGB, WBC, Hct, Plt เมื่อผู้ใช้งานป้อนข้อมูล ระบบจะประมวลผลและแสดงผลการทำนายของทั้ง 5 วิธี เนื่องจากตัวแบบทำนายแต่ละตัวแบบมีความผิดพลาดในการทำนายแตกต่างกัน การประเมินความเสี่ยงผู้วิจัยจึงพัฒนาทั้ง 5 วิธี การตัดสินใจใช้หลักการเสียงข้างมาก คือ 3 ใน 5 แบบทำนายตรงกัน

ผลการวิจัยและอภิปรายผลการวิจัย

ผลการศึกษาการพัฒนาตัวแบบทำนายสัญญาณเตือนโรคมะเร็งลำไส้ใหญ่และทวารหนัก ทำการวิเคราะห์ข้อมูลโดยรวบรวมข้อมูลจากบันทึกเวชระเบียนของผู้ป่วยที่เข้ารับการตรวจและรักษา ในโรงพยาบาลนครปฐม ที่แพทย์วินิจฉัยว่าเป็นโรคมะเร็งลำไส้ใหญ่และทวารหนัก และผู้ป่วยที่ไม่เป็น มะเร็งลำไส้ใหญ่และทวารหนัก ที่มีการตรวจทางห้องปฏิบัติการ ได้แก่ CEA, HGB, WBC, Hct, Plt

ตรวจสอบความครบถ้วนและถูกต้อง จำนวน 396 ตัวอย่าง ประกอบด้วย 2 กลุ่มคือ 1.กลุ่มที่ แพทย์วินิจฉัยว่าเป็นมะเร็งลำไส้ใหญ่และทวารหนักจำนวน 150 ตัวอย่าง 2.กลุ่มที่แพทย์วินิจฉัยว่า ไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักจำนวน 246 ตัวอย่างมาวิเคราะห์และเลือกใช้เทคนิคการทำนาย แบ่งข้อมูลออกเป็น 2 ชุด ดังนี้

ชุดข้อมูลที่ 1 ทำการเลือกข้อมูล 4 ส่วนโดยการสุ่ม เรียกว่าข้อมูลการเรียนรู้ แต่ละส่วนประกอบด้วย กลุ่มที่ 1.กลุ่มที่แพทย์วินิจฉัยว่าเป็นมะเร็งลำไส้ใหญ่และทวารหนักมีจำนวน 4 ส่วนรวม 120 ตัวอย่าง กลุ่มที่ 2.กลุ่มที่แพทย์วินิจฉัยว่าไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักมีจำนวน 4 ส่วนรวม 196 หรือ 197 (ขึ้นอยู่กับรอบที่ทดสอบตัวแบบ) ตัวอย่างนำมาสร้างตัวแบบ 5 อัลกอริทึม

ชุดข้อมูลส่วนที่ 2 ทำการเลือกข้อมูล 1 ส่วนโดยการสุ่ม เรียกว่าข้อมูลตรวจสอบ ประกอบด้วย กลุ่มที่ 1.กลุ่มที่แพทย์วินิจฉัยว่าเป็นมะเร็งลำไส้ใหญ่และทวารหนักจำนวน 30 ตัวอย่าง กลุ่มที่ 2.กลุ่มที่แพทย์วินิจฉัยว่าไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักจำนวน 49 หรือ 50 ตัวอย่าง (ขึ้นอยู่กับรอบที่ทดสอบตัวแบบ) นำไปประเมินความความผิดพลาดของตัวแบบโดยมีผลการศึกษาดังต่อไปนี้

เนื่องจาก เพศ อายุ และผลตรวจทางห้องปฏิบัติการ ได้แก่ CEA, HGB, WBC, Hct, Plt เป็นแอททริบิวต์ ที่มีคุณลักษณะที่จะทำให้ผลลัพธ์การทำนายมีความถูกต้องการทำนายโรคมะเร็งลำไส้ใหญ่และทวารหนักจากชุดข้อมูลตรวจสอบจำนวน 79 ราย ทั้ง 5 อัลกอริทึม ได้ผลดังนี้

Table 1 ค่าความถูกต้องในการจำแนกการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก และรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (root mean square error, RMSE)

Algorithm	Correctly Classified Instance (%)	Incorrectly Classified Instance (%)	Precision	Recall	RMSE
Naive Bayes	89.87	10.13	0.86	1.0	0.30
Decision Tree	75.95	24.05	0.82	0.80	0.49
Random Forest	89.87	10.13	0.91	0.94	0.27
Deep Learning	84.81	15.19	0.83	0.96	0.33
Logistic Regression	92.41	7.59	0.90	1.0	0.24

จาก Table 1 การพิจารณาค่าความถูกต้อง ค่า Precision ค่า Recall และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ผลการวิเคราะห์พบว่าชุดข้อมูลตรวจสอบวิธี Logistic Regression Analysis มีประสิทธิภาพในการทำนายผลดีที่สุด 3 ค่าคือ ค่าความถูกต้อง ค่า Recall และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ถึงแม้ว่าค่า Precision (0.90) จะน้อยกว่าค่า Precision ของวิธี Random Forest (0.91) แต่ไม่แตกต่างกันมากเมื่อเทียบกับ ค่าความถูกต้อง ค่า Recall และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของวิธี Logistic Regression Analysis สูงกว่าวิธี Random Forest

สำหรับผลการทดสอบความถูกต้องของอัลกอริทึม โดยชุดข้อมูลตรวจสอบจำนวน 79 ตัวอย่าง ได้ผลแต่ละอัลกอริทึม ดังนี้

1. อัลกอริทึม Naive Bayes Classifier ผลการทดสอบความถูกต้องของอัลกอริทึม Naive Bayes Classifier พบว่าชุดข้อมูลตรวจสอบสามารถจำแนกผู้ป่วยไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนัก ได้ถูกต้อง 51 ตัวอย่าง คิดเป็นร้อยละ 100 ในขณะที่จำแนกผู้ป่วยเป็นมะเร็งลำไส้ใหญ่และทวารหนัก ได้ถูกต้อง 20 ตัวอย่าง คิดเป็นร้อยละ 71.43

2. อัลกอริทึม Decision Tree ผลการทดสอบความถูกต้องของอัลกอริทึม Decision Tree พบว่าชุดข้อมูลตรวจสอบสามารถจำแนกผู้ป่วยไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักได้ถูกต้อง 41 ตัวอย่าง คิดเป็นร้อยละ 80.39 ในขณะที่จำแนกผู้ป่วยเป็นมะเร็งลำไส้ใหญ่และทวารหนักได้ถูกต้อง 19 ตัวอย่าง คิดเป็นร้อยละ 67.86

3. อัลกอริทึม Random Forest ผลการทดสอบความถูกต้องของอัลกอริทึม Random Forest พบว่าชุดข้อมูลตรวจสอบสามารถจำแนกผู้ป่วยไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักได้ถูกต้อง 48 ตัวอย่าง คิดเป็นร้อยละ 94.12 ในขณะที่จำแนกผู้ป่วยเป็นมะเร็งลำไส้ใหญ่และทวารหนักได้ถูกต้อง 23 ตัวอย่าง คิดเป็นร้อยละ 82.14

4. อัลกอริทึม Deep Learning ผลการทดสอบความถูกต้องของอัลกอริทึม Deep Learning พบว่าชุดข้อมูลตรวจสอบสามารถจำแนกผู้ป่วยไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักได้ถูกต้อง 49 ตัวอย่าง คิดเป็นร้อยละ 96.08 ในขณะที่จำแนกผู้ป่วยเป็นมะเร็งลำไส้ใหญ่และทวารหนักได้ถูกต้อง 18 ตัวอย่าง คิดเป็นร้อยละ 64.29

5. อัลกอริทึม Logistic Regression Analysis ผลการทดสอบความถูกต้องของอัลกอริทึม Logistic Regression Analysis พบว่าชุดข้อมูลตรวจสอบสามารถจำแนกผู้ป่วยไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักได้ถูกต้อง 51 ตัวอย่าง คิดเป็นร้อยละ 100 ในขณะที่จำแนกผู้ป่วยเป็นมะเร็งลำไส้ใหญ่และทวารหนักได้ถูกต้อง 22 ตัวอย่าง คิดเป็นร้อยละ 78.57

การออกแบบการแสดงผลและพัฒนาตัวแบบทำนายการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก

การออกแบบหน้าจอแสดงผลและการพัฒนาตัวแบบการทำนายการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนักผู้วิจัยได้ใช้ Microsoft Excel เป็นส่วนติดต่อกับผู้ใช้งานซึ่งพัฒนาด้วยภาษา VBA เพื่ออำนวยความสะดวกให้กับผู้ใช้งานระบบ โดยนำข้อมูลที่เก็บรวบรวมทั้งหมดมาใส่ไว้ในโปรแกรม แสดงได้ดัง Figure 2 และ Figure 3



Figure 2 หน้าจอแสดงผลตัวแบบทำนายการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก

หลักเกณฑ์การประเมินผลการทำนาย

1. การตัดสินใจผลการทำนายใช้หลักการเสี่ยงข้างมาก คือ 3 ใน 5 วิธีทำนายตรงกัน
2. ผลการทำนายพบว่าผู้ประเมินมีความเสี่ยง (3 ใน 5 วิธีทำนายตรงกัน) ต่อการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก ผู้ประเมินควรเข้ารับการคัดกรอง พบแพทย์ทันที
3. ผลการทำนายพบว่าผู้ประเมินไม่เสี่ยง (3 ใน 5 วิธีทำนายตรงกัน) ต่อการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก ควรพิจารณาว่าผลการทำนายนั้นวิธี Logistic regression analysis หรือ Naive Bayes Classifier หรือทั้งสองวิธี ทั้งนี้ผู้ประเมินควรประเมินความเสี่ยงอื่นๆ เช่น ผู้ที่มีอายุ 50 ปีขึ้นไปทั้งชายและหญิง คนในครอบครัวมีประวัติเป็นมะเร็งลำไส้ใหญ่ มีประวัติพบเนื้องอก (Polyps) ในลำไส้ การรับประทานอาหารที่มีไขมันสูง ผู้ที่มีประวัติดื่มสุราหรือสูบบุหรี่ ผู้ที่เป็นโรคอ้วน การรับประทานอาหารประเภทเนื้อสัตว์หรือประเภทเนื้อแดง ที่ผ่านการปรุงด้วยความร้อนนานๆ มากเกินไป เป็นต้น ผู้ประเมินควรเข้ารับการคัดกรอง พบแพทย์

ปิด

Figure 3 หน้าจอแสดงหลักเกณฑ์การประเมินผลการทำนาย

สรุปผลการวิจัย

จากการวิเคราะห์ข้อมูลเมื่อชุดข้อมูลการเรียนรู้ร้อยละ 80 ใช้ในการสร้างตัวแบบทำนาย และใช้ข้อมูลตรวจสอบ ในการตรวจสอบตัวแบบทำนายพบว่าวิธีทำนายทั้ง 5 อัลกอริทึม สามารถทำนายความถูกต้องได้สูงกว่าร้อยละ 80 โดยเฉพาะวิธี Logistic Regression Analysis มีความถูกต้องสูงถึงร้อยละ 92.40 ดังนั้นคุณลักษณะที่นำมาสร้างตัวแบบจึงมีความเหมาะสมในการทำนายการเกิดโรคมะเร็งลำไส้ใหญ่และทวารหนัก

เมื่อพิจารณาความคลาดเคลื่อนผลการทำนายพบว่าผู้ป่วยที่เป็นมะเร็งลำไส้ใหญ่และทวารหนักวิธี Deep Learning วิธี Decision Tree วิธี Naive Bayes Classifier วิธี Logistic Regression

Analysis และวิธี Random Forest ทำนายว่าไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักคิดเป็นร้อยละ 35.71 32.14 28.57 21.43 และ 17.86 ตามลำดับ เพื่อลดความคลาดเคลื่อนที่เกิดขึ้นจากความผิดพลาดของการทำนายจึงควรพิจารณาผลการทำนายทั้ง 5 อัลกอริทึมและความเสี่ยงอื่น ๆ เช่น ผู้ที่มีอายุ 50 ปีขึ้นไปทั้งชายและหญิง คนในครอบครัวมีประวัติเป็นมะเร็งลำไส้ใหญ่ มีประวัติพบเนื้องอก (Polyps) ในลำไส้ การรับประทานอาหารที่มีไขมันสูง ผู้ที่มีประวัติติ่มสุราหรือสูบบุหรี่ ผู้ที่เป็นโรคอ้วน การรับประทานอาหารประเภทเนื้อสัตว์หรือประเภทเนื้อแดง ที่ผ่านการปรุงด้วยความร้อนนาน ๆ มากเกินไป เป็นต้น ควรเข้ารับการคัดกรอง พบแพทย์ก่อนพบมะเร็งในระยะลุกลาม อย่างไรก็ตาม สำหรับผู้ที่ไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนักพบว่าวิธี Logistic Regression Analysis และ Naive Bayes Classifier ลดความคลาดเคลื่อนในการทำนายสามารถทำนายความถูกต้องได้ร้อยละ 100 ซึ่งวิธี Naive Bayes Classifier ทำนายความถูกต้องสำหรับคนที่ไม่เป็นมะเร็งลำไส้ใหญ่และทวารหนัก ร้อยละ 100 (Salmi, N. & Rustam, Z., 2019)

จากผลการวิจัยถึงแม้จะพบว่าวิธี Logistic Regression Analysis มีประสิทธิภาพในการทำนายผลดีที่สุดแต่ยังต้องใช้อัลกอริทึมอีก 4 วิธีได้แก่ วิธี Naive Bayes Classifier วิธี Decision Tree วิธี Random Forest และ วิธี Deep Learning ว่าผลทำนายมีผลสอดคล้องกับวิธี Logistic Regression Analysis หรือไม่ เพื่อให้มั่นใจในผลการวิเคราะห์ ลดความคลาดเคลื่อน ดังที่กล่าวมา

ข้อเสนอแนะ

1. ตัวแบบทำนายยังมีความคลาดเคลื่อนควรศึกษาคุณลักษณะเพิ่ม เช่น Carbohydrate Antigen 125 (CA125), Carbohydrate Antigen 19-9 (CA 19-9) และประวัติในครอบครัวเป็นมะเร็ง เป็นต้น เพื่อลดความคลาดเคลื่อนและเพิ่มประสิทธิภาพตัวแบบทำนาย
2. ทดสอบประสิทธิภาพตัวแบบทำนายโดยการนำไปใช้กับข้อมูลของผู้ป่วยในโรงพยาบาลแห่งอื่นเพื่อให้เกิดความแม่นยำของตัวแบบ
3. ข้อมูลผลตรวจเลือดไม่ควรถ้วนผู้วิจัยได้ตัดข้อมูลออกทำให้รูปแบบการเรียนรู้การจำแนกขาดหรือสูญเสีย ส่งผลต่อการสร้างตัวแบบทำนายที่มีความแม่นยำสูง
4. สร้างแอปพลิเคชัน ตัวแบบทำนายเพื่อให้ประชาชนทั่วไปเข้าถึงง่าย สะดวก สามารถประเมินความเสี่ยงได้ด้วยตนเอง เช่น เว็บ แอปพลิเคชัน (Web Application) หรือ แอปพลิเคชันสำหรับมือถือ เป็นต้น

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณ โรงพยาบาลนครปฐม อนุญาตให้ข้อมูลจากเวชระเบียนเก็บข้อมูลและเป็นสถานที่ในการวิจัย และผู้มีส่วนเกี่ยวข้องทุกท่านที่ให้ความช่วยเหลือในงานวิจัยนี้ดำเนินการจนสำเร็จลุล่วงไปได้ด้วยดี

References

- American Cancer Society. (2022). *Colorectal Cancer Early Detection, Diagnosis, and Staging*. Retrieved from: <https://www.cancer.org/content/dam/CRC/PDF/Public/8893.00.pdf>
- National Cancer Institute Department of Medical Services Ministry of Public Health. (n.p.). *Colorectal cancer*. Bangkok: National Cancer Institute Department of Medical Services Ministry of Public Health.
- National Cancer Prevention and Control Plan Committee Department of Medical Services Ministry of Public Health. (2016). *National Cancer Control Programme 2018-2022*. Nonthaburi: Department of Medical Services Ministry of Public Health. 4.
- Parinyanitikul, N. (2017). *How to approach asymptomatic patients with rising tumor marker*. Retrieved from: http://www.thethaicancer.com/Webdocument/GP_article/GP_article_006.html
- Petpichetchian, W. (2011). *Best nursing practice in cancer care*. Songkhla: Chanmueang Kanpim. 21.
- Public Health Office Nakhon Pathom. (2022). *Performance summary report form according to government inspection issues, supervision, normal case, round 2, fiscal year 2022, Nakhon Pathom Province*. Nakhon Pathom: Nakhon Pathom Hospital.
- Salmi, N., Rustam, Z. (2019). Naïve Bayes Classifier Models for Predicting the Colon Cancer. *IOP Conference Series: Materials Science and Engineering*, 546(5), 052068
- Siriraj Piyamaharajkarun Hospital. (2021). *Colon cancer: know it, it can be prevented*. Retrieved from: <https://www.siphhospital.com/th/news/article/share/932>
- Srivatanakul, P., Attasara, P. (2007). *Cancer incidence in Thailand Vol.IV, 1998-2000*. Bangkok: Bangkok Medical; 20-21.