

ปัจจัยที่สัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19 กรณีศึกษาโรงพยาบาลระยอง

พรพิมล ชัยวุฒิศักดิ์^{1*}, ภูริทัต เมฆวิไลพันธ์²

รับบทความ: 20 มีนาคม 2568; ส่งแก้ไข: 16 พฤษภาคม 2568; ตอรับ: 23 พฤษภาคม 2568

บทคัดย่อ

บทนำ : การแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) ถือเป็นวิกฤตด้านสาธารณสุขที่ส่งผลกระทบต่ออย่างรุนแรงทั่วโลก รวมถึงประเทศไทยที่มีจำนวนผู้ติดเชื้อเพิ่มขึ้นอย่างต่อเนื่อง ส่งผลให้ระบบบริการสาธารณสุขเผชิญกับข้อจำกัด ทั้งในด้านจำนวนเตียง ทรัพยากรทางการแพทย์ และแนวทางการรักษาที่เปลี่ยนแปลงตามสถานการณ์ การคัดกรองและประเมินความเสี่ยงของผู้ป่วยจึงเป็นขั้นตอนสำคัญในการจัดลำดับความสำคัญเพื่อการดูแลรักษาที่เหมาะสม อย่างไรก็ตาม ข้อมูลที่ใช้อ้างอิงส่วนใหญ่มาจากต่างประเทศ ซึ่งอาจไม่สะท้อนบริบทของประเทศไทยได้อย่างถูกต้อง

วัตถุประสงค์ : เพื่อวิเคราะห์ปัจจัยที่สัมพันธ์กับการเสียชีวิตของผู้ป่วยโควิด-19 โดยใช้ตัวแบบที่เหมาะสมสำหรับการทำนายโอกาสเสียชีวิต

วิธีดำเนินการวิจัย : เป็นการวิจัยเชิงปริมาณโดยอาศัยข้อมูลผู้ป่วยโรคโควิด-19 ที่เข้ารับการรักษาในโรงพยาบาลระยอง และจำหน่ายออกจากโรงพยาบาลตั้งแต่วันที่ 1 เมษายน พ.ศ. 2564 ถึงวันที่ 31 มกราคม พ.ศ. 2565 จากฐานข้อมูลผู้ป่วยในระบบสารสนเทศโรงพยาบาล ปัจจัยที่ศึกษาแบ่งเป็น 3 กลุ่ม ได้แก่ (1) คุณลักษณะส่วนบุคคล (เช่น เพศ อายุ สัญชาติ ดัชนีมวลกาย) (2) อาการแรกเริ่ม (เช่น ผลเอกซเรย์ทรวงอก ความดัน อัตราหายใจ และระดับออกซิเจนในเลือด) และ (3) ปัจจัยด้านการรักษา (เช่น การใช้ยา การใช้เครื่องช่วยหายใจ โรคประจำตัว และโรคแทรกซ้อน) การวิเคราะห์ข้อมูลโดยอาศัยสถิติเชิงพรรณนาและวิธีจำแนกกลุ่ม 3 วิธี ประกอบด้วย การวิเคราะห์การถดถอยลอจิสติกทวิภาค ต้นไม้ตัดสินใจ และโครงข่ายประสาทเทียมแบบหลายลำดับชั้น

ผลการวิจัย : ตัวแบบโครงข่ายประสาทเทียมแบบหลายลำดับชั้นที่มีชั้นซ่อน 3 ชั้น ใช้ฟังก์ชัน ReLU และ Sigmoid เป็นวิธีที่ให้ประสิทธิภาพในการทำนายที่ดีที่สุดด้วยค่าความแม่นยำเท่ากับ 0.95 โดยปัจจัยสำคัญที่มีอิทธิพลต่อการเสียชีวิตของผู้ป่วยโควิด-19 ได้แก่ การใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ การป่วยเป็นโรคไต อายุ จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา และยาไฮโดรคอร์ติโซน

สรุป : ปัจจัยที่สัมพันธ์เชิงบวกกับการเสียชีวิต ได้แก่ อายุ โรคประจำตัว เช่น เบาหวาน ความดันโลหิตสูง ระดับออกซิเจนในเลือดต่ำ และการได้รับการรักษาล่าช้า โดยเฉพาะในผู้ป่วยที่ต้องใช้เครื่องช่วยหายใจ ผลการศึกษาสามารถใช้พัฒนาแนวทางคัดกรองและดูแลผู้ป่วยกลุ่มเสี่ยงสูง

คำสำคัญ : ปัจจัยที่สัมพันธ์ต่อการเสียชีวิต โรคโควิด-19 การวิเคราะห์การถดถอยลอจิสติก ต้นไม้ตัดสินใจ โครงข่ายประสาทเทียมแบบหลายลำดับชั้น

¹ ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ 10520

² ภาควิชาสถิติ (นักศึกษาระดับปริญญาโท) คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ 10520 *อีเมลล์: porpimol.ch@kmitl.ac.th

Factors Associated with Deaths of Covid-19 Patients: Case study of Rayong Hospital

Pornpimol Chaiwuttisak^{1*}, Phoositat Mekwilaiphan²

Received: March 18, 2025; Received revision: May 16, 2025 ; Accepted: May 23, 2025

Abstract

Background : The outbreak of Coronavirus Disease 2019 (COVID-19) has become a significant public health crisis with widespread global impacts. In Thailand, the number of confirmed cases has continued to rise, leading to major challenges in the healthcare system, including limited hospital bed capacity, shortages of medical resources, and constantly evolving treatment guidelines. Screening and risk assessment of patients have therefore become crucial steps for prioritizing treatment and allocating resources effectively. However, most reference data are derived from international studies, which may not accurately reflect the context of Thailand's population and healthcare system.

Objective : To analyze the factors associated with mortality among COVID-19 patients using appropriate predictive models for estimating the probability of death.

Material and methods : This quantitative study utilized data from COVID-19 patients admitted to and discharged from Rayong Hospital between April 1, 2021, and January 31, 2022. Data were retrieved from the hospital's Health Information System (HIS). The factors studied were categorized into three groups: (1) Personal characteristics (e.g., gender, age, nationality, body mass index), (2) Initial clinical presentation (e.g., chest X-ray results, blood pressure, respiratory rate, and blood oxygen saturation), and (3) Treatment-related factors (e.g., medication usage, mechanical ventilation, underlying diseases, and complications recorded in medical records). Data were analyzed using descriptive statistics and three classification methods: binary logistic regression, decision trees, and multilayer perceptron (MLP) neural networks.

Result : The MLP neural network model, consisting of three hidden layers using ReLU and Sigmoid activation functions, provided the highest predictive performance with an accuracy of 0.95. The most influential factors associated with mortality among COVID-19 patients were invasive mechanical ventilation, chronic kidney disease, age, the number of days from diagnosis to discharge, and the use of hydrocortisone.

Conclusion : Factors positively associated with mortality included advanced age, underlying conditions such as diabetes and hypertension, low blood oxygen saturation, and delayed access to treatment—particularly among patients requiring mechanical ventilation. The findings of this study can be used to develop screening protocols and care strategies for high-risk patients.

Keywords : factors associated with death, COVID-19, logistic regression, decision tree, multilayer neural networks

¹ Department of Statistics, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520.

² Department of Statistics (Graduate student), School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520.

*E-mail: pornpimol.ch@kmitl.ac.th

บทนำ

การระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) ได้เริ่มต้นขึ้นในเดือน ธันวาคม พ.ศ. 2562 โดยพบครั้งแรกในนคร อู่ฮั่น เมืองหลวงของมณฑลหูเป่ย์ สาธารณรัฐ ประชาชนจีน ซึ่งสาเหตุที่เป็นไปได้มากที่สุดในการ ติดต่อสู่คนคือการสัมผัสกับเนื้อสัตว์ ประเภทต่าง ๆ ที่วางขายในตลาด หลังจากพบ การระบาดของเชื้อไวรัสสายพันธุ์ใหม่ในเมือง อู่ฮั่นประเทศจีน องค์การอนามัยโลก (The World Health Organization: WHO) ได้ออกมาระบุว่า ไวรัสชนิดดังกล่าว คือ SARS-CoV-2 เรียกว่า Corona Virus Disease 2019 (COVID-19) ตามการประกาศชื่ออย่างเป็นทางการที่ใช้เรียกโรคทางเดินหายใจที่เกิดจาก ไวรัสโคโรนาสายพันธุ์ใหม่ 2019 และพบการ แพร่เชื้อจากคนสู่คนผ่านละอองฝอยขนาดเล็ก (Aerosol) องค์การอนามัยโลกได้ประกาศให้ การระบาดนี้เป็นภาวะฉุกเฉินทางสาธารณสุข ระหว่างประเทศ

วันที่ 11 มีนาคม พ.ศ. 2563 องค์การ อนามัยโลกประกาศโรคติดเชื้อไวรัสโคโรนา 2019 เป็นการระบาดใหญ่ (Coronavirus Pandemic) มีจำนวนผู้ป่วยเพิ่มขึ้นนอกเหนือจากประเทศจีน เป็นจำนวนมากอย่างรวดเร็ว พบผู้ป่วยรายใหม่ ทั่วโลกเพิ่มขึ้นวันละประมาณ 3,000 - 4,000 ราย มีผู้เสียชีวิตเพิ่มขึ้นวันละประมาณ 200 - 300 ราย^[1]

ประเทศไทยพบผู้ติดเชื้อรายแรกที่เข้ารับ การรักษาเมื่อวันที่ 13 มกราคม พ.ศ. 2563 เป็น ชาวจีนที่ติดเชื้อจากประเทศจีน และได้เดินทางมา ประเทศไทย ส่วนผู้ป่วยที่ติดเชื้อในประเทศไทย รายแรกมีการรายงานเมื่อ 31 มกราคม พ.ศ. 2563^[2] วันที่ 1 มีนาคม พ.ศ. 2563 กระทรวง สาธารณสุข รายงานพบผู้เสียชีวิตรายแรกของ ประเทศไทย เป็นชายไทย อายุ 35 ปี ซึ่งเสียชีวิตด้วย กลุ่มอาการทำหน้าที่ผิดปกติของหลายอวัยวะ^[3]

การระบาดในระลอกแรกนั้นอยู่ในช่วงเดือน มีนาคม - เมษายน พ.ศ. 2563 และสามารถควบคุม โรคได้ภายใน 2 เดือน และเกิดการระบาดระลอก ใหม่ในประเทศไทย เป็นการระบาดของเชื้อสาย พันธุ์ใหม่คือสายพันธุ์ G614 ซึ่งเป็นสายพันธุ์ที่พบ ในสหภาพพม่า เมื่อเดือนธันวาคม พ.ศ. 2563 ตรวจพบคนไทยติดเชื้อรายแรก และแพร่เชื้อจาก ตลาดกลางกุ้ง จังหวัดสมุทรสาคร^[4]

สำหรับจังหวัดระยอง เมื่อวันที่ 19 มีนาคม พ.ศ. 2563 พบผู้ป่วยติดเชื้อไวรัส โคโรนา รายแรกมีประวัติเดินทางไปทำงานที่ กรุงเทพมหานคร^[5] วันที่ 28 ธันวาคม พ.ศ. 2563 พบผู้ป่วยโควิด-19 เสียชีวิตรายแรก เป็นชายอายุ 45 ปี มีโรคประจำตัวคือ โรคหัวใจ ขาดเลือด โรคเบาหวานไม่ได้ทานยา^[6]

การระบาดของไวรัสโควิดกลายพันธุ์ โดยมีรหัสที่เรียกว่า โควิดสายพันธุ์ B.1.1.529 หรือโอไมครอน (Omicron) ถูกรายงานว่าพบ ครั้งแรก ในตอนใต้ของทวีปแอฟริกาใต้ที่ ประเทศบอตสวานา^[7] จากงานวิจัยพบว่า เชื้อ ไวรัสโควิด-19 สายพันธุ์โอไมครอน มีการกลาย พันธุ์ถึง 50 ตำแหน่ง และมีการกลายพันธุ์บน โปroteinหนาม (Spike Protein) จำนวน 32 ตำแหน่ง ซึ่งเป็นเหตุให้ไวรัสสามารถแพร่ กระจายและติดต่อยุ่ได้เร็วขึ้น และยังทำให้ตัว ไวรัสมีความสามารถที่จะหลบหลีกภูมิคุ้มกันได้ อีกด้วย^[8] โดยส่วนใหญ่ของผู้ติดเชื้อสายพันธุ์ โอไมครอนจะไม่แสดงอาการ หรือแสดงอาการ เพียงเล็กน้อย เมื่อเปรียบเทียบกับอาการของ ผู้ติดเชื้อสายพันธุ์เดลต้าพบว่าความรุนแรงของ เชื้อสายพันธุ์โอไมครอนจะมีความรุนแรงน้อยกว่า แต่เนื่องจากเชื้อไวรัสโคโรนา 2019 สาย พันธุ์โอไมครอน สามารถแพร่กระจายได้อย่าง รวดเร็ว ก็อาจก่อให้เกิดความล้มเหลวทางด้าน สาธารณสุขได้เช่นกัน^[9] เมื่อวันที่ 6 มกราคม พ.ศ. 2565 กระทรวงสาธารณสุขได้ประกาศ

ยกระดับเตือนภัยโควิด-19 จากเดิมระดับ 3 เป็นระดับ 4 หลังจากที่จำนวนของผู้ติดเชื้อเพิ่มขึ้นอย่างมีนัยสำคัญ และมีการแพร่ระบาดของเชื้อโควิด-19 สายพันธุ์โอไมครอนในหลาย ๆ จังหวัด จากสถานการณ์การระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 ในประเทศ คาดการณ์ว่าจะพบจำนวนผู้ติดเชื้อเพิ่มสูงขึ้นเรื่อย ๆ แต่จำนวนผู้เสียชีวิตอาจเพิ่มสูงขึ้นเล็กน้อยหรือคงที่ และจะมีการแพร่ระบาดของเชื้อโควิด-19 สายพันธุ์โอไมครอนอย่างต่อเนื่องและอาจครอบคลุมทุกจังหวัดของประเทศ และอาจประสบกับปัญหาเดียวกันกับการระบาดของเชื้อในระลอกเดือนเมษายนที่ผ่านมา^[10]

จากการทบทวนวรรณกรรมที่เกี่ยวข้องกับการสร้างตัวแบบในการทำนายโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 พบว่าในปี ค.ศ. 2020 การศึกษาของ Das et al.^[11], Muhammad et al.^[14], Li et al.^[15], Agbelusi and Olayemi^[16] และ ในปี ค.ศ. 2021 การศึกษาของ Sánchez-Montañés et al.^[12], Ferreira et al.^[13], Khan. et al.^[17] และ Pourhomayoun and Shakibi^[18] ได้ทำการสร้างตัวแบบในการทำนายโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 เปรียบเทียบโดยใช้วิธีการเรียนรู้ของเครื่อง (Machine Learning) และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น ในการศึกษาของ Das. et al.^[11] และ Sánchez-Montañés. et al.^[12] ผลการศึกษาพบว่าวิธีการวิเคราะห์การถดถอยลอจิสติก มีประสิทธิภาพดีที่สุด เปรียบเทียบประสิทธิภาพด้วยค่าความแม่นยำในการศึกษาของ Das. et al.^[11] มีค่าความแม่นยำสูงสุดเท่ากับ 0.96 และเปรียบเทียบประสิทธิภาพด้วยค่าพื้นที่ใต้โค้งลักษณะเฉพาะดำเนินการตัวรับซึ่งแสดงความสัมพันธ์ระหว่างอัตราบวกจริงและอัตราบวกเท็จในสถานการณ์ จากการศึกษาของ Sánchez-Montañés et al.^[12] มีค่าพื้นที่ใต้

โค้งลักษณะเฉพาะดำเนินการตัวรับมากที่สุดเท่ากับ 0.89 ในการศึกษาของ Ferreira et al.^[13] และ Muhammad et al.^[14] ผลการศึกษาพบว่าวิธีต้นไม้ตัดสินใจ มีประสิทธิภาพดีที่สุด เมื่อเปรียบเทียบประสิทธิภาพด้วยค่าความแม่นยำในการศึกษาของ Ferreira et al.^[13] และ Muhammad et al.^[14] มีค่าความแม่นยำเท่ากับ 0.90 และ 0.95 ตามลำดับ และในการศึกษาของ Li et al.^[15], Agbelusi and Olayemi^[16], Khan et al.^[17] และ Pourhomayoun and Shakibi^[18] ผลการศึกษาพบว่าวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้นมีประสิทธิภาพ ดีที่สุด โดยเปรียบเทียบประสิทธิภาพด้วยค่าความแม่นยำในการศึกษาของ Li et al.^[15], Agbelusi and Olayemi^[16], Khan et al.^[17] และ Pourhomayoun and Shakibi^[18] มีค่าความแม่นยำเท่ากับ 0.85, 0.85, 0.97 และ 0.90 ตามลำดับ จากการทบทวนวรรณกรรมที่เกี่ยวข้องผู้วิจัยจึงดำเนินการสร้างตัวแบบในการทำนายโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 ด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก วิธีต้นไม้ตัดสินใจ และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น

ด้วยสถานการณ์ในช่วงการระบาดของโรคโควิด-19 ระลอกใหม่ระหว่างเดือนเมษายน พ.ศ. 2564 ถึง มกราคม พ.ศ. 2565 ซึ่งจำนวนผู้ป่วยที่เพิ่มสูงขึ้น ประกอบกับแนวทางการรักษาโรคโควิด-19 มีการปรับเปลี่ยนอยู่ตลอดเวลา และจำนวนเตียงที่มีอยู่อย่างจำกัด กระบวนการคัดกรองผู้ป่วย เข้ารับการรักษาในโรงพยาบาลจึงเป็นสิ่งสำคัญ ความรู้หรือข้อมูลต่าง ๆ ที่ใช้ในการอ้างอิงส่วนใหญ่จะเป็นการศึกษากลุ่มตัวอย่างในต่างประเทศซึ่งอาจไม่เพียงพอ ดังนั้นผู้วิจัยเห็นความสำคัญในการศึกษาปัจจัยที่มีความสัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19 ในโรงพยาบาล

ระยอง โดยศึกษาปัจจัยส่วนบุคคล ปัจจัยด้านสุขภาพ ปัจจัยด้านอาการแรกเริ่ม และปัจจัยด้านการรักษา เพื่อนำผลการทำนายนำมาใช้เป็นแนวทางในการเฝ้าระวังผู้ป่วยที่มีความเสี่ยงสูงในการเสียชีวิตและทำการรักษาอย่างมีประสิทธิภาพมากขึ้น

วัตถุประสงค์

- 1) เพื่อศึกษาปัจจัยที่สัมพันธ์กับการเสียชีวิตของผู้ป่วยโควิด-19
- 2) เพื่อวิเคราะห์ตัวแบบที่เหมาะสมในการทำนายโอกาสเสียชีวิตของผู้ป่วยโควิด-19

วิธีดำเนินการวิจัย

จากการศึกษาและทบทวนวรรณกรรมที่เกี่ยวข้องทั้งในประเทศ และต่างประเทศพบว่างานวิจัยที่ศึกษาเกี่ยวกับการสร้างตัวแบบในการทำนายโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 โดยส่วนมากสร้างตัวแบบในการทำนายด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก วิธีต้นไม้ตัดสินใจ และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น และเมื่อเปรียบเทียบประสิทธิภาพของตัวแบบโดยพิจารณาจากค่าความแม่นยำ พบว่า ทั้ง 3 วิธี มีความแม่นยำสูงสุด เมื่อเทียบกับวิธีการจำแนกกลุ่มวิธีอื่น ผู้วิจัยจึงเลือกวิธีจำแนกกลุ่มทั้ง 3 วิธีดังกล่าวในการนำไปสร้างและเปรียบเทียบประสิทธิภาพของตัวแบบสำหรับการวิเคราะห์หาปัจจัยที่สัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19 กรณีศึกษาโรงพยาบาลระยอง เพื่อเปรียบเทียบประสิทธิภาพการทำนายของตัวแบบที่สร้างขึ้นในแต่ละวิธี ตลอดจนศึกษาปัจจัยที่สัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19

1. การเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลผู้ป่วยโรคโควิด-19 ที่เข้ารับการรักษาในโรงพยาบาลระยอง และจำหน่ายออกจากโรงพยาบาลตั้งแต่วันที่ 1 เมษายน พ.ศ. 2564

ถึงวันที่ 31 มกราคม พ.ศ. 2565 โดยมีการเก็บรวบรวมข้อมูลลักษณะประชากรทั่วไป อาทิ เพศ อายุ สัญชาติ และดัชนีมวลกาย จากฐานข้อมูลผู้ป่วยในระบบสารสนเทศโรงพยาบาล(Hospital Information Systems: HIS) และเก็บรวบรวมข้อมูลอาการแรกเริ่มของผู้ป่วยแต่ละราย อันได้แก่ ผลภาพรังสีทรวงอก ความดันโลหิต อัตราการเต้นของหัวใจ ปริมาณออกซิเจนในเลือด อุดมภูมิร่างกาย อัตราการหายใจ ข้อมูลด้านการรักษาพยาบาลทั้งในเรื่องของการใช้ยา จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา การใช้เครื่องช่วยหายใจ โรคประจำตัว และโรคแทรกซ้อนที่แพทย์ผู้ให้การรักษาสรุปไว้ในเวชระเบียนผู้ป่วย รวมถึงประเภทการจำหน่ายของผู้ป่วยทุกราย

2. ตัวแปรที่ใช้ในการวิจัย

งานวิจัยนี้ จำแนกตัวแปรอิสระออกเป็นปัจจัยด้านต่าง ๆ ทั้งหมด 4 ด้าน โดยมีปัจจัยส่วนบุคคล (ดังตารางที่ 1) ปัจจัยสุขภาพ (ดังตารางที่ 2) ปัจจัยอาการแรกเริ่ม (ดังตารางที่ 3) และปัจจัยการรักษา (ดังตารางที่ 4) โดยที่ตัวแปรตามคือ ประเภทการจำหน่ายของผู้ป่วยโดยจำแนกเป็นรอดชีวิต และเสียชีวิต (ดังตารางที่ 5)

ตารางที่ 1 รายละเอียดของตัวแปรอิสระด้านปัจจัยส่วนบุคคล

ตัวแปรอิสระ	ความหมาย	สเกลข้อมูล
X_1	เพศ	นามบัญญัติ
X_2	อายุ (หน่วย : ปี)	อัตราส่วน
X_3	สัญชาติ	นามบัญญัติ
X_4	ดัชนีมวลกาย (หน่วย : เมตร/กิโลกรัม ²)	อัตราส่วน

ตารางที่ 2 รายละเอียดของตัวแปรอิสระด้านปัจจัยสุขภาพ

ตัวแปรอิสระ	ความหมาย	สเกลข้อมูล
X ₅	โรคหลอดเลือดสมอง	นามบัญญัติ
X ₆	โรคปอดอุดกั้นเรื้อรัง	นามบัญญัติ
X ₇	โรคเบาหวาน	นามบัญญัติ
X ₈	โรคหัวใจ	นามบัญญัติ
X ₉	โรคเอดส์	นามบัญญัติ
X ₁₀	โรคไต	นามบัญญัติ
X ₁₁	โรคตับ	นามบัญญัติ
X ₁₂	โรคปอด	นามบัญญัติ
X ₁₃	โรคมะเร็ง	นามบัญญัติ
X ₁₄	วัคซีนแอสตราเซนเก้กา (หน่วย : เข็ม)	อัตราส่วน
X ₁₅	วัคซีนซิโนแวค (หน่วย : เข็ม)	อัตราส่วน
X ₁₆	วัคซีนไฟเซอร์ (หน่วย : เข็ม)	อัตราส่วน
X ₁₇	วัคซีนซิโนฟาร์ม (หน่วย : เข็ม)	อัตราส่วน

ตารางที่ 3 รายละเอียดของตัวแปรอิสระด้านปัจจัยอาการแรกเริ่ม

ตัวแปรอิสระ	ความหมาย	สเกลข้อมูล
X ₁₈	ภาพรังสีทรวงอกรูปแบบ Patchy Opacity	นามบัญญัติ
X ₁₉	ภาพรังสีทรวงอกรูปแบบ Ground Glass Opacity	นามบัญญัติ
X ₂₀	ภาพรังสีทรวงอกรูปแบบ Reticulonodular Opacity	นามบัญญัติ
X ₂₁	การกระจายตัวของรอยโรคในปอดเด่นที่ Peripheral	นามบัญญัติ
X ₂₂	การกระจายตัวของรอยโรคเด่นที่ Perihilar	นามบัญญัติ
X ₂₃	การกระจายตัวของรอยโรคเด่นที่ปอดข้างซ้าย	นามบัญญัติ

X ₂₄	การกระจายตัวของรอยโรคเด่นที่ปอดข้างขวา	นามบัญญัติ
X ₂₅	การกระจายตัวของรอยโรคเด่นที่ปอดทั้ง 2 ข้าง	นามบัญญัติ
X ₂₆	การกระจายตัวของรอยโรคเด่นที่ปอดส่วนล่าง	นามบัญญัติ
X ₂₇	การกระจายตัวของรอยโรคเด่นที่ปอดส่วนบน	นามบัญญัติ
X ₂₈	การกระจายตัวของรอยโรคเด่นที่ปอดส่วนกลาง	นามบัญญัติ
X ₂₉	ความดันโลหิตสูงสุดขณะหัวใจห้องล่างบีบตัว (หน่วย : มิลลิเมตรปรอท)	อัตราส่วน
X ₃₀	ความดันโลหิตที่ต่ำสุดขณะหัวใจห้องล่างคลายตัว (หน่วย : มิลลิเมตรปรอท)	อัตราส่วน
X ₃₁	อัตราการหายใจ (หน่วย : ครั้ง/นาที)	อัตราส่วน
X ₃₂	อัตราการเต้นของหัวใจ (หน่วย : ครั้ง/นาที)	อัตราส่วน
X ₃₃	ปริมาณออกซิเจนในเลือด(หน่วย : เปอร์เซ็นต์)	อัตราส่วน
X ₃₄	อุณหภูมิร่างกาย (หน่วย : องศาเซลเซียส)	อันตรภาค

ตารางที่ 4 รายละเอียดของตัวแปรอิสระด้านปัจจัยการรักษา

ตัวแปรอิสระ	ความหมาย	สเกลข้อมูล
X ₃₅	เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ	นามบัญญัติ
X ₃₆	เครื่องช่วยหายใจแบบไม่ใส่ท่อช่วยหายใจ	นามบัญญัติ

X ₃₇	การให้ออกซิเจนผ่านหน้ากาก หรือสายออกซิเจน	นามบัญญัติ
X ₃₈	ยาฟาวิพิราเวียร์	นามบัญญัติ
X ₃₉	ยาคาซิรีวิแมบ และยาอิมตีวีแมบ	นามบัญญัติ
X ₄₀	ยาเรมเดซิเวียร์	นามบัญญัติ
X ₄₁	ยาฟ้าทลายโจร	นามบัญญัติ
X ₄₂	ยาแพ็กซ์โลวิด	นามบัญญัติ
X ₄₃	ยาเดกซาเมทาโซน	นามบัญญัติ
X ₄₄	ยาเมทิลเพรดนิโซโลน	นามบัญญัติ
X ₄₅	ยาไฮโดรคอร์ติโซน	นามบัญญัติ
X ₄₆	จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา (หน่วย : วัน)	อัตราส่วน

ตารางที่ 5 รายละเอียดของตัวแปรตาม

ตัวแปรตาม	ความหมาย	สเกลข้อมูล
y	ประเภทการจำหน่าย	นามบัญญัติ

3 การเตรียมข้อมูล

3.1 การทำความสะอาดข้อมูล

เป็นขั้นตอนการจัดการกับข้อมูลที่สูญหาย โดยใช้คำสั่ง dropna โมดูลของ pandas โปรแกรมภาษาไพธอน (Python Programming Language) พบว่าข้อมูลไม่พบค่าสูญหาย

3.2 การแปลงข้อมูล

เนื่องจากตัวแปรอิสระบางตัว และตัวแปรตามเป็นตัวแปรเชิงคุณภาพ มีความจำเป็นต้องทำการแปลงข้อมูล ก่อนที่จะนำไปสร้างตัวแบบในการทำนาย ผู้วิจัยจึงใช้วิธีการ

แทนค่ากลุ่มโดยใช้คำสั่ง LabelEncoder ใน preprocessing โมดูลของ scikit-learn ในโปรแกรมภาษาไพธอน

3.3 การแปลงข้อมูลให้อยู่ในคะแนนมาตรฐาน

การแปลงข้อมูลให้อยู่ในคะแนนมาตรฐานเพื่อปรับให้ค่าของข้อมูลในตัวแปรอิสระแต่ละตัว ให้มีค่าเฉลี่ยเท่ากับ 1 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0 โดยใช้คำสั่ง StandardScaler ใน preprocessing โมดูลของ scikit-learn ในโปรแกรมภาษาไพธอน

3.4 การสุ่มเพิ่มจำนวนข้อมูล

เพื่อแก้ไขปัญหาข้อมูลไม่สมดุล (Imbalanced Data) ดังนั้นผู้วิจัยจึงนำวิธีสุ่มเพิ่มจำนวนข้อมูลด้วยเทคนิค Synthetic Minority Over-sampling Technique หรือเรียกสั้น ๆ ว่า SMOTE มาใช้ปรับสมดุลข้อมูลเป็นการเพิ่มจำนวนข้อมูลกลุ่มน้อยให้มีจำนวนเพิ่มขึ้น โดยการเพิ่มข้อมูลในกลุ่มน้อยนั้นทำให้การกระจายของกลุ่มข้อมูลมีความสมดุลมากขึ้น โดยทำการสุ่มค่าข้อมูลที่อยู่ในกลุ่มข้อมูลน้อยขึ้นมา 1 ค่า หลังจากนั้นพิจารณาค่าเพื่อนบ้านใกล้สุด K ตัว (K-Nearest Neighbor) แล้วคำนวณค่าระยะทาง (Euclidean Distance) ระหว่างค่าที่สุ่มกับค่าข้อมูลใกล้เคียงแต่ละค่า เพื่อหาค่าระยะทางที่น้อยที่สุดระหว่างค่าที่สุ่มกับค่าข้อมูลใกล้เคียง จากนั้นจึงสร้างข้อมูลเทียมระหว่างค่าข้อมูลที่สุ่มกับค่าข้อมูลใกล้เคียงตัวที่ให้ค่าระยะทางที่น้อยที่สุด ซึ่งก็คือเพิ่มจำนวนข้อมูลของผู้ป่วยโรคโควิด-19 มีประเภทการจำหน่ายเป็นเสียชีวิต ให้มีอัตราส่วนเท่ากับข้อมูลกลุ่มมาก ซึ่งก็คือจำนวนข้อมูลของผู้ป่วยโรคโควิด-19 มีสถานการณ์จำหน่ายเป็นรอดชีวิต โดยใช้คำสั่ง SMOTE ใน

over_sampling โมดูล ของ Imbalanced-learn ในโปรแกรมภาษาไพธอน

3.5 การแบ่งข้อมูลเป็นชุดข้อมูลฝึกหัด และชุดข้อมูลทดสอบ

ในงานวิจัยนี้จะทำการแบ่งข้อมูลจากการสุ่มเพิ่มจำนวนข้อมูลทั้งหมดออกเป็น ชุดข้อมูลเรียนรู้ร้อยละ 70 และชุดข้อมูลทดสอบร้อยละ 30 โดยใช้คำสั่ง train_test_split ใน selection โมดูล ของ scikit-learn ในโปรแกรมภาษาไพธอน

4. การวิเคราะห์การจำแนกกลุ่ม

ในงานวิจัยนี้ ได้มีการนำวิธีการจำแนกกลุ่มของชุดข้อมูลมาเปรียบเทียบกัน ทั้งหมด 3 วิธี ประกอบไปด้วย

4.1 วิธีการวิเคราะห์การถดถอยลอจิสติก

วิธีการวิเคราะห์การถดถอยลอจิสติก เป็นการศึกษาเพื่อทำนายว่าจะเกิดเหตุการณ์หนึ่งขึ้นหรือไม่ หรือมีโอกาสเกิดขึ้นมากน้อยเพียงใด โดยมีการกำหนดค่าตัวแปรตัวหนึ่งหรือหลายตัวที่คาดว่าจะส่งผลต่อการเกิดเหตุการณ์นั้น ๆ งานวิจัยนี้ใช้การวิเคราะห์การถดถอยลอจิสติกในการศึกษาเนื่องจากตัวแปรตามเป็นตัวแปรเชิงกลุ่มโดยตัวแปรตามมีลักษณะเป็นตัวแปรแบ่งสอง (Dichotomous Variable) สำหรับการวิเคราะห์การถดถอยลอจิสติกทวิภาค (Binary Logistic Regression) เพื่อวิเคราะห์หาปัจจัยเสี่ยงที่สัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19 ซึ่งเป็นการหาความสัมพันธ์ระหว่างตัวแปรตาม 1 ตัวและตัวแปรอิสระมากกว่า 1 ตัว โดยตัวแปรตามสามารถมีค่าได้เพียง 2 ค่า คือ 1 และ 0 ซึ่งจะแสดงถึงการเสียชีวิตหรือรอดชีวิต ส่วนตัวแปรอิสระที่ใช้ในการวิเคราะห์สามารถเป็นได้ทั้งตัวแปรเชิงปริมาณและตัวแปรเชิงคุณภาพ โดยทำการตรวจสอบว่าตัวแปรอิสระไม่มีความสัมพันธ์

กันหรือไม่เกิดปัญหาความสัมพันธ์เชิงเส้นพหุ โดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน และคัดเลือกตัวแปรอิสระด้วยวิธี Backward Stepwise โดยเป็นการคัดเลือกตัวแปรอิสระที่สามารถอธิบายความผันแปรของตัวแปรตามได้อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 และตรวจสอบความเหมาะสมของตัวแบบ โดยพิจารณาจากค่าความเป็นไปได้ ทดสอบสัมประสิทธิ์การถดถอย (Regression Coefficient) ด้วยสถิติทดสอบของวัตต์ และพิจารณาค่าสถิติทดสอบระดับความสัมพันธ์ McFadden's R square

สำหรับการวิเคราะห์การถดถอยลอจิสติกทวิภาค ที่ค่าตัวแปรตาม หรือ y สามารถเป็นไปได้เพียง 2 ค่าคือ ไม่เกิดเหตุการณ์ ($y=0$) หรือเกิดเหตุการณ์ ($y=1$) มีความสัมพันธ์กับตัวแปรอิสระไม่อยู่ในรูปเชิงเส้น โดยมีตัวแบบในการทำนายข้อมูลประชากรตามสมการที่ 1 ดังนี้

$$P_y = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}} \quad (1)$$

และตัวแบบในการทำนายข้อมูลตัวอย่างตามสมการที่ 2 ดังนี้

$$P_y = \frac{e^{b_0 + b_1 x_1 + \dots + b_p x_p}}{1 + e^{b_0 + b_1 x_1 + \dots + b_p x_p}} \quad (2)$$

โดยที่ P_y แทนความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ

x_i แทนตัวแปรอิสระ i

β_i แทนค่าสัมประสิทธิ์ของตัวแปร i ของประชากร

b_i แทนค่าสัมประสิทธิ์ของตัวแปร i ของตัวอย่าง

e แทนค่าคงที่ ($e = 2.71828$)

เมื่อพิจารณาความสัมพันธ์ระหว่างตัวแปรอิสระ และตัวแปรตามของการวิเคราะห์การถดถอยลอจิสติกทวิภาค ที่ไม่เป็นรูปเชิงเส้น จึงต้องมีการปรับให้ความสัมพันธ์อยู่ในรูปเชิงเส้นในรูปของ Odds ซึ่งหมายถึง อัตราส่วนระหว่างโอกาสที่จะเกิดเหตุการณ์ที่สนใจ ($y=1$) และโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจ ($y=0$) ดังแสดงในสมการที่ 3

$$Odds = \frac{P_y}{Q_y} \quad (3)$$

โดยที่ P_y แทนความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ

Q_y แทนความน่าจะเป็นของการไม่เกิดเหตุการณ์ที่สนใจ

และ Odds Ratio จะหมายถึงอัตราส่วนของค่า Odds เปรียบเทียบกันระหว่าง 2 ค่าของตัวแปรอิสระ ซึ่งจะบ่งบอกถึงโอกาสในการเกิดเหตุการณ์ที่สนใจว่าเป็นกี่เท่าเมื่อเทียบระหว่างค่าของตัวแปรอิสระ 2 ค่าตัวอย่างเช่น Odds Ratio ของโอกาสที่ผู้ป่วยจะเสียชีวิตเป็นดังสมการที่ 4

$$Odds = \frac{P_{Died}}{P_{Survival}} \quad (4)$$

1) ค่า Odds Ratio มีค่าเท่ากับ 1 หมายความว่าโอกาสที่ผู้ป่วยจะเสียชีวิตและรอดชีวิตมีโอกาสดังกัน

2) ค่า Odds Ratio มีค่ามากกว่า 1 หมายความว่าโอกาสที่ผู้ป่วยจะเสียชีวิตมากกว่าโอกาสที่ผู้ป่วยจะรอดชีวิต

3) ค่า Odds Ratio มีค่าน้อยกว่า 1 หมายความว่าโอกาสที่ผู้ป่วยจะรอดชีวิตมากกว่าโอกาสที่ผู้ป่วยจะเสียชีวิต

จากสมการที่ 4 หากค่าฟังก์ชันตอบสนอง ลอจิต (Logit Response Function: Logit)

พบว่า ลอจิตของข้อมูลประชากรอยู่ในรูปเชิงเส้นดังสมการที่ 5

$$\log(Odds) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (5)$$

และลอจิตของข้อมูลตัวอย่างอยู่ในรูปเชิงเส้นดังสมการที่ 6

$$\log(Odds) = b_0 + b_1 x_1 + \dots + b_p x_p \quad (6)$$

4.2 วิธีต้นไม้ตัดสินใจ

ในช่วงปลายของช่วง ค.ศ. 1970 J. Ross Quinlan ได้คิดค้นขั้นตอนวิธี (Algorithm) สำหรับสร้างต้นไม้ตัดสินใจที่มีชื่อว่า ID3 (Iterative Dichotomiser) ต่อมาเขาได้พัฒนาต่อยอด ID3 ไปเป็น C4.5 ซึ่งได้กลายมาเป็นขั้นตอนวิธีพื้นฐานที่ใช้สำหรับเปรียบเทียบประสิทธิภาพการทำงานของขั้นตอนวิธีต่าง ๆ ทางด้านการเรียนรู้แบบมีผู้สอน (Supervised Learning) โดยการสร้างต้นไม้ตัดสินใจใช้ค่ามาตรฐานอัตราส่วนเกน (Gain Ratio) เพื่อเลือกลักษณะเฉพาะที่จะใช้เป็นรากหรือโหนด ถ้าให้ชุดของข้อมูล M ประกอบด้วยค่าที่เป็นไปได้คือ $\{m_1, m_2, \dots, m_n\}$ และให้ความน่าจะเป็นที่จะเกิดค่า m_i มีค่าเท่ากับ $P(m_i)$ จะได้ว่า ค่าเกนสารสนเทศ (Information Gain) หรือที่เรียกอีกอย่างหนึ่งว่า ค่าเอนโทรปี (Entropy) ของ M เขียนแทนด้วย $I(M)$ คำนวณได้ดังสมการที่ 1

$$I(M) = -\sum_{i=1}^n P(m_i) \log_2 P(m_i) \quad (1)$$

ถ้าให้ข้อมูลสอน คือ T และลักษณะเฉพาะที่เป็นโหนด คือ x และมีค่าทั้งหมดที่เป็นไปได้ n ค่า โหนดปัจจุบันจะแบ่งตัวอย่าง T ออกตามกิ่งเป็น $\{t_1, t_2, \dots, t_n\}$ ตามค่าที่เป็นไปได้ของ x ดังนั้นจึงสามารถคำนวณค่าเกนสารสนเทศ หลังจากแบ่งตามลักษณะเฉพาะ x ได้ดังสมการที่ 7

$$I_x(T) = \sum_{i=1}^n \frac{|t_i|}{|T|} I(t_i) \quad (7)$$

ค่ามาตรฐานเกน (Gain) ของลักษณะเฉพาะ x ได้ดังสมการที่ 8

$$Gain(x) = I(T) - I_x(T) \quad (8)$$

ถ้าให้ T คือ ชุดของตัวอย่าง เมื่อแบ่งตัวอย่างนี้ตามลักษณะเฉพาะ x จะได้ชุดของตัวอย่างย่อยในแต่ละกิ่ง คือ $\{t_1, t_2, \dots, t_n\}$ จำนวน n ชุด ตามค่าที่เป็นไปได้ในลักษณะเฉพาะ x จะสามารถคำนวณค่าสารสนเทศของการแบ่งแยก (Split Information) ของลักษณะเฉพาะแต่ละตัวได้ดังสมการที่ 9

$$Split\ Information(T) = -\sum_{i=1}^n \frac{|t_i|}{|T|} \log_2 \frac{|t_i|}{|T|} \quad (9)$$

ค่ามาตรฐานอัตราส่วนเกน (Gain Ratio) ของลักษณะเฉพาะ x ได้ดังสมการที่ 10

$$Gain\ Ratio(x) = \frac{Gain(x)}{Split\ Information(T)} \quad (10)$$

ท้ายสุดจึงเลือกค่ามาตรฐานอัตราส่วนเกนสูงสุดเป็นลักษณะเฉพาะเริ่มต้น และเลือกลักษณะเฉพาะถัดไปตามค่ามาตรฐานอัตราส่วนเกนน้อยลงตามลำดับ กำหนดให้ค่าเอนโทรปีสูงสุดเป็นลักษณะเฉพาะเริ่มต้น และเลือกลักษณะเฉพาะถัดไปตามค่าเอนโทรปีน้อยลงตามลำดับ

4.3 วิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น

โครงข่ายประสาทเทียมแบบหลายลำดับชั้นเป็นภาพแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างเป็นแบบหลาย ๆ ชั้น ใช้สำหรับงานที่มีความซับซ้อนโดยมีกระบวนการฝึกฝนเป็นแบบมีผู้สอน และใช้ขั้นตอนการส่งค่าย้อนกลับ (Back Propagation) สำหรับการฝึกฝนกระบวนการส่งค่าย้อนกลับ

ประกอบด้วย 2 ส่วนย่อยคือ การส่งผ่านไปข้างหน้า (Forward Pass) การส่งผ่านย้อนกลับ (Backward Pass) สำหรับการส่งผ่านไปข้างหน้า ข้อมูลจะผ่านเข้าโครงข่ายประสาทเทียมที่ชั้นข้อมูลเข้า และจะส่งผ่านจากอีกชั้นหนึ่งไปสู่อีกชั้นหนึ่งจนกระทั่งถึงชั้นข้อมูลออก ส่วนการส่งผ่านย้อนกลับ คำนวณน้ำหนักการเชื่อมต่อจะถูกปรับเปลี่ยนให้สอดคล้องกับกฎการแก้ข้อผิดพลาด (Error-correction) คือ ผลต่างของผลตอบที่แท้จริง (Actual Response) กับผลตอบเป้าหมาย (Target Response) เกิดเป็นสัญญาณผิดพลาด (Error Signal) ซึ่งสัญญาณผิดพลาดนี้จะถูกส่งย้อนกลับเข้าสู่โครงข่ายประสาทเทียมในทิศทางตรงกันข้ามกับการเชื่อมต่อ และค่าน้ำหนักของการเชื่อมต่อจะถูกปรับจนกระทั่งผลตอบที่แท้จริงเข้าใกล้ผลตอบเป้าหมาย สัญญาณที่มีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น มี 2 ประเภทคือ

- 1) สัญญาณฟังก์ชัน (Function Signal) เป็นสัญญาณเข้าที่มาจากโหนดในชั้นก่อนหน้า และจะส่งผ่านไปข้างหน้าจากโหนดหนึ่งไปสู่อีกโหนดหนึ่ง
- 2) สัญญาณผิดพลาด เป็นสัญญาณย้อนกลับที่เกิดขึ้นที่โหนดในชั้นข้อมูลออกของโครงข่ายประสาทเทียม และถูกส่งผ่านย้อนกลับจากชั้นหนึ่งไปสู่อีกชั้นหนึ่ง

โดยโครงข่ายประสาทเทียมแบบหลายลำดับชั้น มีการเชื่อมโยงกัน 3 ชั้น ประกอบด้วยชั้นข้อมูลเข้า (Input Layer) ถัดมาเป็นชั้นซ่อน (Hidden Layer) และชั้นสุดท้ายคือชั้นข้อมูลออก (Output Layer) โดยส่วนประกอบที่ถูกบรรจุอยู่ในเซลล์ประสาทแต่ละตัวประกอบด้วย 2 ฟังก์ชันย่อย คือ ฟังก์ชันผลรวม (Summation Function) และฟังก์ชันกระตุ้น (Activation Function)

ฟังก์ชันผลรวมทำหน้าที่ในการคำนวณผลรวมของข้อมูลที่ได้จากชั้นข้อมูลเข้า ส่วนฟังก์ชันกระตุ้นทำหน้าที่ปรับเปลี่ยนค่าของข้อมูลที่ได้จากฟังก์ชันผลรวมให้อยู่ในช่วงที่ต้องการ ฟังก์ชันกระตุ้นที่นิยม ยกตัวอย่างเช่น ฟังก์ชันเชิงเส้น (Linear Function) ฟังก์ชันซิกมอยด์ (Sigmoid Function) และฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ (Hyperbolic Tangent Function)

โดยในแต่ละชั้นไม่จำเป็นต้องเป็นฟังก์ชันกระตุ้นเดียวกันก็ได้ ชั้นซ่อนนั้นมีหน้าที่สำคัญคือ จะพยายามแปลงข้อมูลที่เข้ามาในชั้น (Layer) ให้สามารถแยกแยะความแตกต่างโดยใช้เส้นตรงเส้นเดียว (Linearly Separable) และก่อนที่ข้อมูลจะถูกส่งไปถึงชั้นข้อมูลออก ในบางครั้งอาจจำเป็นต้องใช้ชั้นซ่อนมากกว่า 1 ชั้น ในการแปลงข้อมูลให้อยู่ในภาพเส้นตรงเส้นเดียว ในการคำนวณหาข้อมูลออก (Output) ในปัญหาการจำแนกทำได้โดยการใส่ข้อมูลเข้า (Input) เข้าไปในโครงข่ายประสาทเทียมที่ได้ทำการหาไว้แล้ว จากนั้นให้ทำการเปรียบเทียบค่าของข้อมูลออกในชั้นข้อมูลออก และให้ทำการเลือกค่าของข้อมูลออก ที่มีค่าเซลล์ประสาท (Neuron) ที่สูงกว่า และทำการรับค่าของทำนายที่ตรงกับเซลล์ประสาทที่เลือก และให้นำมาเปรียบเทียบกับค่าที่ยอมรับได้ หากอยู่ในช่วงที่รับได้กล่าวคือ ค่าความคลาดเคลื่อน (Error) น้อยกว่าค่าความคลาดเคลื่อนที่กำหนด ก็ให้ทำการรับข้อมูลชุดถัดไป แต่หากมากกว่าค่าที่ยอมรับได้ ให้ทำการปรับค่าน้ำหนักและค่าความเอนเอียง (Bias) ตามขั้นตอนที่ได้กล่าวไว้ข้างต้น เมื่อทำการปรับน้ำหนักเรียบร้อยแล้ว ให้ทำการรับข้อมูลชุดถัดไป และทำตามขั้นตอนซ้ำอีกรอบจนกระทั่งถึงข้อมูลชุดสุดท้าย และเมื่อทำข้อมูลชุดสุดท้ายเสร็จ จะนับเป็น 1 รอบของการคำนวณ (1 Epoch) จากนั้นจะทำการหาค่าผิดพลาดรวมเฉลี่ย จากค่าความผิดพลาดที่ได้เก็บ

เอาไว้ เพื่อใช้ในการตรวจสอบว่าค่าความผิดพลาดโดยเฉลี่ยในการจำแนกนั้น มีค่าน้อยกว่าค่าผิดพลาดที่ยอมรับได้หรือไม่ ถ้าใช่แสดงว่าโครงข่ายประสาทเทียมที่สร้างขึ้นนั้นสามารถให้ผลลัพธ์ที่ถูกต้องของทุก ๆ ข้อมูลแล้ว จึงทำการจบการเรียนรู้ได้ แต่ถ้าไม่ใช่ ให้กลับไปทำตามขั้นตอนแรกโดยเริ่มรับข้อมูลชุดที่ 1 ใหม่

โดยกำหนดค่าอัตราการเรียนรู้ (Learning Rate) เป็น 0.1 ค่าโมเมนตัม (Momentum) เป็น 0.9 จำนวนรอบการสอน (Training Time) 1,000 รอบ ชั้นซ่อนจำนวน 3 ชั้น โดยชั้นแรกมีจำนวนโหนด 500 โหนด ชั้นถัดมามีจำนวน 1,000 โหนด และชั้นสุดท้ายมีจำนวน 500 โหนด ใช้ฟังก์ชันเรคตีไฟด์ลินเนียญิต เป็นฟังก์ชันกระตุ้นในทุกชั้นซ่อน และฟังก์ชันซิกมอยด์เป็นฟังก์ชันกระตุ้นในชั้นข้อมูลออก

5 ค่าความสำคัญของตัวแปร (Feature Importance)

ค่าความสำคัญของตัวแปรเป็นการศึกษาว่าตัวแปรอิสระตัวใดที่ส่งผลต่อตัวแบบในการทำนายมาก หรือน้อยกว่ากัน โดยในวิธีการวิเคราะห์การถดถอยลอจิสติก จะพิจารณาจากค่าสัมประสิทธิ์การถดถอย วิธีนี้ไม่ตัดสินใจ พิจารณาจากค่าเอนโทรปี และวิธีโครงข่ายประสาทเทียมคำนวณค่าความสำคัญจากวิธี Permutation Importance โดยใช้ คำสั่ง `permutation_importance` ใน `inspection` โมดูลของ `scikit-learn` ในโปรแกรมภาษา ไพธอน โดยหลักการทำงานคือ พิจารณาตัวแปรอิสระทีละตัว โดยทำการสลับตำแหน่งของข้อมูลในตัวแปรอิสระที่พิจารณา โดยที่ตำแหน่ง

ของตัวแปรอิสระอื่น ๆ ยังคงเดิม แล้วนำไปสร้างตัวแบบในการทำนาย ถ้าประสิทธิภาพของตัวแบบลดลงจากตัวแบบก่อนหน้าที่ไม่มีการสลับตำแหน่งตัวแปรอิสระแสดงว่า ตัวแปรอิสระนั้นมีความสำคัญมาก โดยประสิทธิภาพของตัวแบบที่ลดลงมากกว่า แปลว่าตัวแปรอิสระนั้นมีความสำคัญมากกว่า

6 การเปรียบเทียบประสิทธิภาพของตัวแบบ

การเปรียบเทียบประสิทธิภาพของตัวแบบที่เหมาะสม โดยนำแบบจำลองที่สร้างขึ้นด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก วิธีต้นไม้การตัดสินใจ และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น นำมาเปรียบเทียบประสิทธิภาพ โดยพิจารณาจากค่าความแม่นยำ ค่าความเที่ยง ค่าเรียกคืน ค่าความถ่วงดุล และค่าพื้นที่ใต้โค้ง โดยถ้าค่ายิ่งมากแสดงว่า ตัวแบบมีประสิทธิภาพในการทำนายสูง

ผลการศึกษา

1. ลักษณะทั่วไปของกลุ่มตัวอย่าง

จากการตรวจสอบข้อมูลผู้ป่วยที่ได้รับการวินิจฉัยร่วมว่าเป็นโรคโควิด-19 ที่เข้ารับการรักษาในโรงพยาบาลระยะของจำนวนทั้งสิ้น 1,006 ราย จำแนกเป็นชายคิดเป็นร้อยละ 45.13 และหญิงคิดเป็นร้อยละ 54.87 อายุเฉลี่ยอยู่ที่ 60 ปี ผู้ป่วยร้อยละ 89.36 มีสัญชาติไทย ที่เหลือเป็นผู้ป่วยชาวต่างชาติ โดยที่ผู้ป่วยทั้งหมดมีดัชนีมวลกายเฉลี่ยอยู่ที่ 26.25 กิโลกรัม/เมตร² ดังตารางที่ 6

ตารางที่ 6 จำนวนและร้อยละของตัวอย่างสำหรับตัวแปรอิสระด้านปัจจัยส่วนบุคคล

	ตัวแปรอิสระ	จำนวน (%)
X ₁	เพศ	
	หญิง	552 (54.87%)
	ชาย	454 (45.13%)
X ₂	อายุ (mean ± SD)	55.99 ± 17.24
X ₃	สัญชาติ	
	ไทย	899 (89.36%)
	ต่างชาติ	107 (10.64%)
X ₄	ดัชนีมวลกาย (mean ± SD)	26.25 ± 5.90

จากตารางที่ 7 โรคเบาหวานเป็นโรคที่พบมากที่สุดร้อยละ 27.04 รองลงมาคือโรคไตร้อยละ 13.62 โดยโรคปอดเป็นโรคที่พบน้อยที่สุดอยู่ที่ร้อยละ 0.89 ผู้ป่วยส่วนมากได้รับวัคซีนแอสตราเซนเนกา โดยจำแนกเป็นได้รับวัคซีน 1 เข็มคิดเป็นร้อยละ 5.27 ได้รับวัคซีน 2 เข็มคิดเป็นร้อยละ 1.69

ตารางที่ 7 จำนวนและร้อยละของตัวอย่างสำหรับตัวแปรอิสระด้านปัจจัยสุขภาพ

	ตัวแปรอิสระ	จำนวน (%)
X ₅	โรคหลอดเลือดสมอง	43 (4.27%)
X ₆	โรคปอดอุดกั้นเรื้อรัง	14 (1.39%)
X ₇	โรคเบาหวาน	272 (27.04%)
X ₈	โรคหัวใจ	84 (8.35%)
X ₉	โรคเอดส์	14 (1.39%)

	ตัวแปรอิสระ	จำนวน (%)
X ₁₀	โรคไต	137 (13.62%)
X ₁₁	โรคตับ	31 (3.08%)
X ₁₂	โรคปอด	9 (0.89%)
X ₁₃	โรคมะเร็ง	19 (1.89%)
X ₁₄	วัคซีนแอสตราเซนเก้กา (mean ± SD)	0.09 ± 0.34
X ₁₅	วัคซีนซิโนแวค (mean ± SD)	0.06 ± 0.26
X ₁₆	วัคซีนไฟเซอร์ (mean ± SD)	0.01 ± 0.08
X ₁₇	วัคซีนซิโนฟาร์ม (mean ± SD)	0.01 ± 0.12

ลักษณะภาพรังสีทรวงอกที่พบส่วนใหญ่เป็นแบบ Patchy opacity คิดเป็นร้อยละ 50.30 รอยโรคมักกระจายตัวอยู่ที่ปอดส่วน peripheral คิดเป็นร้อยละ 13.32 โดยพบความผิดปกติของปอดทั้ง 2 ข้าง มากที่สุดคิดเป็นร้อยละ 54.37 และปอดส่วนล่างคิดเป็นร้อยละ 38.97 ผู้ป่วยมีความดันโลหิตเฉลี่ย 130/79 มิลลิเมตรปรอท อัตราการหายใจเฉลี่ย 22 ครั้ง/นาที อัตราการเต้นของหัวใจเฉลี่ย 92 ครั้ง/นาที เฮอร์เซ็นต์ปริมาณออกซิเจนในเลือดเฉลี่ย เท่ากับ 95% และอุณหภูมิร่างกายเฉลี่ยอยู่ที่ 36.72 องศาเซลเซียส ดังตารางที่ 8

ตารางที่ 8 จำนวนและร้อยละของตัวอย่างสำหรับตัวแปรอิสระด้านปัจจัยแรกครับ

	ตัวแปรอิสระ	จำนวน (%)
X ₁₈	ภาพรังสีทรวงอกแบบ Patchy Opacity	506 (50.30%)
X ₁₉	ภาพรังสีทรวงอกแบบ Ground Glass Opacity	333 (33.10%)
X ₂₀	ภาพรังสีทรวงอกแบบ Reticulonodular Opacity	56 (5.57%)
X ₂₁	การกระจายตัวของรอยโรคในปอดเด่นที่ Peripheral	134 (13.32%)
X ₂₂	การกระจายตัวของรอยโรคเด่นที่ Perihilar	6 (0.60%)
X ₂₃	การกระจายตัวของรอยโรคเด่นที่ปอดข้างซ้าย	125 (12.43%)
X ₂₄	การกระจายตัวของรอยโรคเด่นที่ปอดข้างขวา	111 (11.03%)
X ₂₅	การกระจายตัวของรอยโรคเด่นที่ปอดทั้ง 2 ข้าง	547 (54.37%)
X ₂₆	การกระจายตัวของรอยโรคเด่นที่ปอดส่วนล่าง	392 (38.97%)
X ₂₇	การกระจายตัวของรอยโรคเด่นที่ปอดส่วนบน	73 (7.26%)
X ₂₈	การกระจายตัวของรอยโรคเด่นที่ปอดส่วนกลาง	287 (28.53%)
X ₂₉	ความดันโลหิตสูงสุดขณะหัวใจห้องล่างบีบตัว (mean ± SD)	130.56 ± 25.07
X ₃₀	ความดันโลหิตที่ต่ำสุดขณะหัวใจห้องล่างคลายตัว (mean ± SD)	79.82 ± 15.54
X ₃₁	อัตราการหายใจ (mean ± SD)	22.21 ± 3.91
X ₃₂	อัตราการเต้นของหัวใจ (mean ± SD)	92.38 ± 18.33
X ₃₃	ปริมาณออกซิเจนในเลือด (mean ± SD)	95.02 ± 7.09
X ₃₄	อุณหภูมิร่างกาย (mean ± SD)	36.72 ± 0.67

ตารางที่ 9 จำนวนและร้อยละของตัวอย่างสำหรับตัวแปรอิสระด้านปัจจัยการรักษา

	ตัวแปรอิสระ	จำนวน (%)
X ₃₅	เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ	119 (11.83%)
X ₃₆	เครื่องช่วยหายใจแบบไม่ใส่ท่อช่วยหายใจ	135 (13.42%)
X ₃₇	การให้ออกซิเจนผ่านหน้ากาก หรือสายออกซิเจน	198 (19.68%)
X ₃₈	ยาฟาวิพิราเวียร์	877 (87.18%)
X ₃₉	ยาคาซิรวิแมบ และยาอิมตีวิแมบ	19 (1.89%)
X ₄₀	ยาเรมเดซิเวียร์	71 (7.06%)
X ₄₁	ยาฟ้าทลายโจร	43 (4.27%)
X ₄₂	ยาแพ็กซ์โลวิด	1 (0.1%)
X ₄₃	ยาเดกซาเมทาโซน	730 (72.56%)
X ₄₄	ยาเมทิลเพรดนิโซโลน	414 (41.15%)
X ₄₅	ยาไฮโดรคอร์ติโซน	25 (2.49%)
X ₄₆	จำนวนวันนับตั้งแต่วันตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา (mean ± SD)	15.52 ± 6.56

ผู้ป่วยส่วนใหญ่ให้ออกซิเจนผ่านหน้ากากหรือสายออกซิเจน คิดเป็นร้อยละ 19.68 รองลงมาใช้เครื่องช่วยหายใจแบบไม่ใส่ท่อช่วยหายใจ คิดเป็นร้อยละ 13.42 และใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ คิดเป็นร้อยละ 11.83 ยาที่พบมากที่สุดที่ใช้ในการรักษาผู้ป่วยคือ ยาฟาวิพิราเวียร์ คิดเป็นร้อยละ 87.18 รองลงมาคือยาเดกซาเมทาโซนคิดเป็นร้อยละ 72.56 จำนวนวันนับตั้งแต่วันตรวจพบเชื้อจนถึงวันจำหน่ายการรักษาเฉลี่ย 16 วัน ดังตารางที่ 9

ตารางที่ 10 จำนวนและร้อยละของตัวอย่างสำหรับตัวแปรตาม

	ตัวแปรตาม	จำนวน (%)
ประเภทการจำหน่าย		1,006 (100.00%)
รอดชีวิต		814 (80.91%)
เสียชีวิต		192 (19.09%)

จากตารางที่ 10 พบว่า ผู้ป่วยที่ได้รับการวินิจฉัยร่วมว่าเป็นโรคโควิด-19 ที่เข้ารับการรักษาในโรงพยาบาลระยองจำนวนทั้งสิ้น 1,006 ราย มีเพียง 192 รายเท่านั้น หรือคิดเป็นร้อยละ 19.09 ที่มีประเภทการจำหน่ายเป็นเสียชีวิต ซึ่งเกิดปัญหาข้อมูลไม่สมดุล ดังนั้นผู้วิจัยจึงนำวิธีสุ่มเพิ่มจำนวนข้อมูลด้วยเทคนิค SMOTE มาใช้ปรับสมดุลข้อมูล เป็นการเพิ่มจำนวนข้อมูลกลุ่มน้อย ซึ่งก็คือจำนวนข้อมูลของผู้ป่วยโรคโควิด-19 มีประเภทการจำหน่ายเป็นเสียชีวิต ก่อนนำเข้าสู่กระบวนการสร้างตัวแบบในการทำนายต่อไป โดยผลการสุ่มเพิ่มจำนวนข้อมูลด้วยเทคนิค SMOTE แสดงในตารางที่ 11

ตารางที่ 11 ผลการสุ่มเพิ่มจำนวนข้อมูลด้วยเทคนิค SMOTE

ตัวแปรตาม	จำนวน (%)
ประเภทการจำหน่าย	1,152 (100.00%)
รอดชีวิต	576 (50.00%)
เสียชีวิต	576 (50.00%)

จากตารางที่ 11 พบว่า หลังจากสุ่มเพิ่มจำนวนข้อมูลด้วยเทคนิค SMOTE จำนวนของข้อมูลกลุ่มน้อย (กลุ่มเสียชีวิต) มีจำนวนเพิ่มขึ้น และมีอัตราส่วนของข้อมูลเท่ากับข้อมูลในกลุ่มมาก (กลุ่มรอดชีวิต) เป็น 576 รายทั้งสองกลุ่ม รวมทั้งสิ้น 1,152 ราย

สำหรับการสร้างตัวแบบทำนายด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น ผู้วิจัยได้ดำเนินการปรับค่าของตัวแปรอิสระของทุกตัวแปรให้อยู่ในค่ามาตรฐาน นั่นคือ ค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 ยกเว้นวิธีต้นไม้ตัดสินใจที่ไม่ได้ปรับค่าของตัวแปรอิสระให้อยู่ในค่ามาตรฐานเนื่องจาก การปรับค่าของตัวแปรอิสระ จะส่งผลกระทบต่อขั้นตอนการเลือกลักษณะเฉพาะที่จะใช้เป็นรากหรือโหนด ในวิธีต้นไม้ตัดสินใจ

สำหรับการวิเคราะห์ข้อมูล จะแบ่งข้อมูลออกเป็น 2 ส่วนดังนี้

ส่วนที่ 1 : ชุดข้อมูลชุดเรียนรู้ เพื่อนำไปสร้างตัวแบบ มีร้อยละ 70 ของชุดข้อมูลจากการสุ่มเพิ่มจำนวนข้อมูลด้วยเทคนิค SMOTE ซึ่งจะได้ข้อมูลในส่วนที่ 1 จำนวน 806 ราย

ส่วนที่ 2 : ชุดข้อมูลทดสอบ เพื่อนำไปทดสอบความถูกต้องของตัวแบบ มีร้อยละ 30 ของชุดข้อมูลจากการสุ่มเพิ่มจำนวนข้อมูลด้วยเทคนิค SMOTE ซึ่งจะได้ข้อมูลในส่วนที่ 2 จำนวน 346 ราย

2. การตรวจสอบความเหมาะสมของตัวแบบด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก

จากเงื่อนไขของวิธีการวิเคราะห์การถดถอยลอจิสติกจะต้องกำจัดตัวแปรอิสระที่มีความสัมพันธ์กันเองก่อน จากค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (ดังตารางที่ 12) พบว่าไม่มีตัวแปรอิสระที่มีความสัมพันธ์กันเอง ($r > 0.8$)^[20] แสดงว่าข้อมูลชุดนี้สามารถวิเคราะห์การถดถอยลอจิสติกได้เนื่องจากตัวแปรอิสระไม่มีความสัมพันธ์กันมากหรือไม่มีปัญหาปัญหาความสัมพันธ์เชิงเส้นพหุเกิดขึ้นจากตารางที่ 13 แสดงผลการวิเคราะห์ตัวแบบการถดถอยลอจิสติกพบว่า ตัวแปรอิสระ ที่มีค่าสัมประสิทธิ์เป็นบวก ซึ่งหมายความว่า การที่ผู้ป่วยที่มีอายุ (X_2) สูงขึ้นเป็นปัจจัยสำคัญที่สุดที่สัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19 กล่าวคือเมื่อผู้ป่วยมีอายุสูงขึ้น ส่งผลให้โอกาสที่ผู้ป่วยจะเสียชีวิตเพิ่มมากขึ้นเช่นกัน ปัจจัยที่มีความสำคัญรองลงมาก็คือ ผู้ป่วยที่มีสัญชาติไทย (X_3) มีโอกาสเสียชีวิตสูงกว่าผู้ป่วยที่เป็นชาวต่างชาติ นอกจากนี้ การใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ (X_{35}) ผู้ป่วยที่มีโรคไต (X_{10}) เป็นโรคประจำตัว ผู้ป่วยที่ได้รับยาเมทิลเพรดนิโซโลน (X_{44}) ผู้ป่วยที่มีโรคหัวใจ (X_8) เป็นโรคประจำตัว ผู้ป่วยที่ใช้เครื่องช่วยหายใจแบบไม่ใส่ท่อช่วยหายใจ (X_{36}) ผู้ป่วยที่มีโรคมะเร็ง (X_{13}) เป็นโรคประจำตัว ผู้ป่วยที่มีความผิดปกติของภาพรังสีทรวงอกรูปแบบ Patchy Opacity (X_{18}) อัตราการหายใจ (X_{31}) และผู้ป่วยที่ได้รับยาไฮโดรคอร์ติโซน (X_{45}) ส่งผลให้โอกาสที่ผู้ป่วยจะเสียชีวิตเพิ่มมากขึ้นตามลำดับ ในขณะที่ตัวแปรอิสระที่มีค่าสัมประสิทธิ์เป็นลบ ได้แก่ จำนวนวันระหว่างวันพบเชื้อและวันจะหน่วยการรักษา (X_{46}) พบว่า จำนวนวันที่เพิ่มมากขึ้นนั้น

ส่งผลให้โอกาสการเสียชีวิตลดลง ในทำนองเดียวกัน ผู้ป่วยที่รับฉีดวัคซีนซิโนแอสตราเซนเนก้า (X_{14}) ก่อให้เกิดเชื่อพบว่าช่วยลดโอกาสการเสียชีวิตได้ เช่นเดียวกับผู้ป่วยที่ได้รับยาเดกซาเมทาโซน (X_{43}) ในการรักษาพบว่ามีโอกาสเสียชีวิตลดลง รวมถึงกลุ่มผู้ป่วยมีลักษณะภาพรังสีทรวงอกที่พบความผิดปกติของปอดข้างขวา (X_{24}) หรือผู้ป่วยที่มีโรคเบาหวาน (X_7) เป็นโรคประจำตัว พบว่าโอกาสเสียชีวิตลดลงในกลุ่มผู้ป่วยดังกล่าว

ตารางที่ 12 ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระที่เป็นข้อมูลเชิงปริมาณ

ตัวแปรอิสระ	X_2	X_4	X_{14}	X_{15}	X_{16}	X_{17}	X_{29}	X_{30}	X_{31}	X_{32}	X_{33}	X_{34}	X_{46}
X_2	1.000												
X_4	-0.182	1.000											
X_{14}	0.093	0.019	1.000										
X_{15}	-0.065	0.003	0.127	1.000									
X_{16}	-0.030	0.031	-0.013	0.244	1.000								
X_{17}	-0.080	-0.035	-0.012	-0.012	-0.003	1.000							
X_{29}	0.197	0.075	0.089	0.048	-0.018	-0.008	1.000						
X_{30}	-0.118	0.107	0.022	0.058	0.021	0.009	0.598	1.000					
X_{31}	0.100	0.252	-0.096	-0.109	-0.028	-0.032	0.000	-0.013	1.000				
X_{32}	-0.100	0.053	-0.061	-0.014	0.017	-0.001	-0.025	0.189	0.299	1.000			
X_{33}	-0.115	-0.175	0.095	0.073	0.029	0.031	0.042	0.051	-0.326	-0.100	1.000		
X_{34}	-0.113	0.000	-0.008	0.043	0.013	-0.020	0.090	0.057	-0.014	0.134	0.067	1.000	
X_{46}	-0.088	0.150	0.058	0.095	0.007	-0.025	0.197	0.230	-0.098	0.000	0.073	0.067	1.000

ตารางที่ 13 ผลการวิเคราะห์ตัวแบบการถดถอยลอจิสติกที่มีการคัดเลือกตัวแปรแบบ Backward stepwise

	ตัวแปรอิสระ	coef	Std err	z	p> z
	Intercept	-2.3734	0.282	-8.429	0.000
X_2	อายุ	1.5329	0.190	8.083	0.000
X_3	สัญชาติ	1.2438	0.459	2.709	0.007
X_7	โรคเบาหวาน	-0.4733	0.161	-2.945	0.003
X_8	โรคหัวใจ	0.7133	0.142	5.012	0.000
X_{10}	โรคไต	1.0027	0.139	7.232	0.000
X_{13}	โรคมะเร็ง	0.3997	0.109	3.659	0.000
X_{14}	วัคซีนแอสตราเซนเนก้า	-0.8361	0.241	-3.473	0.001
X_{18}	ภาพรังสีทรวงอกรูปแบบ Patchy Opacity	0.3409	0.154	2.212	0.027
X_{24}	การกระจายตัวของรอยโรคเด่นที่ปอดข้างขวา	-0.4972	0.167	-2.976	0.003
X_{31}	อัตราการหายใจ	0.3385	0.154	2.201	0.028
X_{35}	เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ	1.2081	0.146	8.281	0.000
X_{36}	เครื่องช่วยหายใจแบบไม่ใส่ท่อช่วยหายใจ	0.4201	0.147	2.855	0.004
X_{43}	ยาเดกซาเมทาโซน	-0.6407	0.203	-3.154	0.002
X_{44}	ยาเมทิลเพรดนิโซโลน	0.9584	0.201	4.769	0.000
X_{45}	ยาไฮโดรคอร์ติโซน	0.3075	0.113	2.718	0.007
X_{46}	จำนวนวันนับตั้งแต่วันตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา	-0.9561	0.151	-6.346	0.000

ตารางที่ 14 ค่าความเป็นไปได้ และค่าสถิติทดสอบระดับความสัมพันธ์

Log likelihood		Pseudo R Square
Full Model	Constant-Only Model	
-165.30	-558.04	0.7038

จากตารางที่ 14 ผลการวิเคราะห์ตัวแบบการถดถอยลอจิสติกที่มีการคัดเลือกตัวแปรแบบ Backward stepwise พบว่า มีตัวแปรอิสระทั้งหมด 16 ตัวแปร มีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 ค่าความเป็นไปได้ เมื่อพิจารณาความเหมาะสมค่า -2Log-Likelihood มีค่าเท่ากับ 330.60 ค่า Null deviance หรือค่า Log-Likelihood สำหรับ Constant-Only Model (L0) เท่ากับ -558.04 และค่า Residual deviance หรือค่า Log Likelihood สำหรับ Full Model (L1) เท่ากับ -165.3 ค่า McFadden’s R Square เท่ากับ 0.7038 นั้นหมายความว่า ตัวแบบสามารถทำนายความผันแปรของโอกาสที่ผู้ป่วยโควิด-19 ที่จะเสียชีวิต และรอดชีวิตได้ร้อยละ 70.38 สามารถเขียนสมการ Logit ได้ดังนี้

$$\begin{aligned} \log(\text{Odds}) = & -2.3734 + 1.5329X_2 + 1.2438X_3 - 0.4733X_7 + 0.7133X_8 + 1.0027X_{10} \\ & + 0.3997X_{13} - 0.8361X_{14} + 0.3409X_{18} - 0.4972X_{24} + 0.3385X_{31} \\ & + 1.2081X_{35} + 0.4201X_{36} - 0.6407X_{43} + 0.9584X_{44} + 0.3075X_{45} - 0.9561X_{46} \end{aligned}$$

3. การวิเคราะห์ความสำคัญของตัวแปร (Feature Importance)

จากการวิเคราะห์ความสำคัญของตัวแปร ด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก (LR) วิธีต้นไม้ตัดสินใจ (DT) และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น (NN) พบว่าวิธีการวิเคราะห์การถดถอยลอจิสติกมีตัวแปรอิสระที่มีค่าความสำคัญสูงสุด 5 อันดับแรกที่ส่งผลต่อโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 ได้แก่ อายุ สัญชาติ การใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ โรคไต และการใช้ยามetiloprednizolone ตามลำดับ ในส่วนของวิธีต้นไม้ตัดสินใจพบว่า ตัวแปรอิสระ ได้แก่ จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา อายุ การใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ อัตราการหายใจ และปริมาณออกซิเจนในเลือด เป็นตัวแปรอิสระที่มีความสำคัญในการทำนายโอกาสการเสียชีวิตของผู้ป่วย และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น พบว่าตัวแปรอิสระที่มีค่าความสำคัญสูงสุด 5 อันดับแรกมีความคล้ายคลึงกันกับทั้งสองวิธีที่ผ่านมา โดยที่ใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ มีความสำคัญมากที่สุด รองลงมาคือ โรคไต อายุ จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา และการใช้ยาไฮโดรคอร์ติโซน ตามลำดับ ดังแสดงในตารางที่ 15

ตารางที่ 15 ผลการวิเคราะห์ค่าความสำคัญของตัวแปรอิสระ (Feature Importance) ด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก (LR) วิธีต้นไม้ตัดสินใจ (DT) และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น (NN)

ตัวแปรอิสระ	LR	DT	NN
X_{35} เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ	1.2081	0.1432	0.0873
X_{10} โรคไต	1.0027	0.0081	0.0551
X_2 อายุ	1.5329	0.1823	0.0496
X_{46} จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา	-0.9561	0.2533	0.0300

	ตัวแปรอิสระ	LR	DT	NN
X ₄₅	ยาไฮโดรคอร์ติโซน	0.3075	0.0000	0.0208
X ₈	โรคหัวใจ	0.7133	0.0189	0.0203
X ₃₉	ยากาซิรีวิแมบ และยาอิมตีวิแมบ	0.0000	0.0111	0.0117
X ₇	โรคเบาหวาน	-0.4733	0.0043	0.0109
X ₁₃	โรคมะเร็ง	0.3997	0.0000	0.0094
X ₄₃	ยาเดกซามาทาโซน	-0.6407	0.0271	0.0089
X ₆	โรคปอดอุดกั้นเรื้อรัง	0.0000	0.0000	0.0084
X ₁₄	วัคซีนแอสตราเซนเนก้า	-0.8361	0.0097	0.0082
X ₃	สัญชาติ	1.2438	0.0000	0.0082
X ₃₃	ปริมาณออกซิเจนในเลือด	0.0000	0.0610	0.0077
X ₃₆	เครื่องช่วยหายใจแบบไม่ใส่ท่อช่วยหายใจ	0.4201	0.0099	0.0072
X ₁₅	วัคซีนซิโนแวค	0.0000	0.0000	0.0055
X ₄₀	ยาเรมเดซิเวียร์	0.0000	0.0000	0.0050
X ₁₁	โรคตับ	0.0000	0.0000	0.0047
X ₄₄	ยาเมทิลเพรดนิโซโลน	0.9584	0.0063	0.0047
X ₂₉	ความดันโลหิตสูงสุดขณะหัวใจห้องล่างบีบตัว	0.0000	0.0368	0.0045
X ₅	โรคหลอดเลือดสมอง	0.0000	0.0000	0.0037
X ₁₆	วัคซีนไฟเซอร์	0.0000	0.0000	0.0030
X ₃₂	อัตราการเต้นของหัวใจ	0.0000	0.0115	0.0030
X ₃₁	อัตราการหายใจ	0.3385	0.0681	0.0027
X ₂₃	การกระจายตัวของรอยโรคเด่นที่ปอดข้างซ้าย	0.0000	0.0086	0.0022
X ₃₇	การให้ออกซิเจนผ่านหน้ากาก หรือสายออกซิเจน	0.0000	0.0000	0.0017
X ₉	โรคเอดส์	0.0000	0.0000	0.0017
X ₂₄	การกระจายตัวของรอยโรคเด่นที่ปอดข้างขวา	-0.4972	0.0317	0.0015
X ₁₈	ภาพรังสีทรวงอกรูปแบบ Patchy Opacity	0.0000	0.0040	0.0012
X ₂₇	การกระจายตัวของรอยโรคเด่นที่ปอดส่วนบน	0.0000	0.0000	0.0012
X ₃₄	อุณหภูมิร่างกาย	0.0000	0.0414	0.0012
X ₂₀	ภาพรังสีทรวงอกรูปแบบ Reticulonodular Opacity	0.0000	0.0000	0.0010
X ₂₁	การกระจายตัวของรอยโรคในปอดเด่นที่ Peripheral	0.0000	0.0031	0.0010
X ₁	เพศ	0.0000	0.0126	0.0007
X ₁₂	โรคปอด	0.0000	0.0000	0.0007
X ₁₇	วัคซีนซิโนฟาร์ม	0.0000	0.0000	0.0007
X ₁₉	ภาพรังสีทรวงอกรูปแบบ Ground Glass Opacity	0.3409	0.0000	0.0007
X ₃₀	ความดันโลหิตที่ต่ำสุดขณะหัวใจห้องล่างคลายตัว	0.0000	0.0209	0.0007
X ₄₂	ยาแพ็กซ์โลวิด	0.0000	0.0000	0.0007

ตัวแปรอิสระ	LR	DT	NN
X_{22} การกระจายตัวของรอยโรคเด่นที่ Perihilar	0.0000	0.0000	0.0005
X_4 ดัชนีมวลกาย	0.0000	0.0197	0.0002
X_{41} ยาฟ้าทลายโจร	0.0000	0.0000	0.0002
X_{25} การกระจายตัวของรอยโรคเด่นที่ปอดทั้ง 2 ข้าง	0.0000	0.0000	0.0000
X_{26} การกระจายตัวของรอยโรคเด่นที่ปอดส่วนล่าง	0.0000	0.0065	0.0000
X_{28} การกระจายตัวของรอยโรคเด่นที่ปอดส่วนกลาง	0.0000	0.0000	0.0000
X_{38} ยาฟาวิพิราเวียร์	0.0000	0.0000	0.0000

หมายเหตุ เรียงลำดับค่าความสำคัญด้วยวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น (NN) จากมากไปน้อย

4. การทำนายการเสียชีวิตของผู้ป่วยโควิด-19

การทำนายการเสียชีวิตของผู้ป่วยโควิด-19 จากชุดข้อมูลทดสอบ จำนวน 346 ราย โดยวิธีการวิเคราะห์การถดถอยลอจิสติก วิธีต้นไม้ตัดสินใจ และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น ได้ผลดังนี้

4.1 วิธีการวิเคราะห์การถดถอยลอจิสติก

ตารางที่ 16 เมตริกซ์ความสับสนจากการทำนายการเสียชีวิตของผู้ป่วยโควิด-19 ด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก

		ค่าทำนาย	
		รอดชีวิต	เสียชีวิต
ค่าจริง	รอดชีวิต	136	21
	เสียชีวิต	16	173

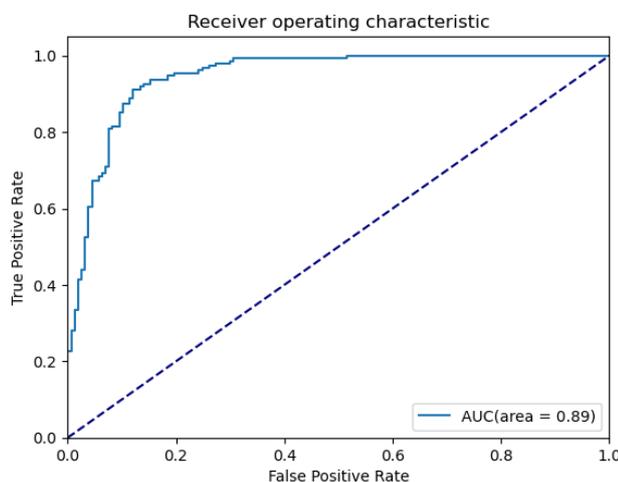
ตารางที่ 16 มีข้อมูล 346 ราย ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 309 ราย โดยมีจำนวนข้อมูลที่จำแนกถูกกว่ารอดชีวิต 136 ราย และข้อมูลที่จำแนกถูกว่าเสียชีวิต 173 ราย และตัวแบบทำนายข้อมูลไม่ถูกต้อง 37 ราย โดยมีจำนวนข้อมูลที่จำแนกผิดว่ารอดชีวิต ซึ่งผลที่แท้จริงเสียชีวิต 16 ราย และข้อมูลที่จำแนกผิดว่าเสียชีวิต ซึ่งผลที่แท้จริงแล้วรอดชีวิต 21 ราย

ตารางที่ 17 ค่าประสิทธิภาพของตัวแบบจำแนกรายละเอียดในแต่ละค่าตอบจากการทำนายเสียชีวิตของผู้ป่วยโควิด-19 ด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก

Class	Precision	Recall	F1- score	Accuracy
0	0.89	0.87	0.88	0.89
1	0.89	0.92	0.90	
macro avg	0.89	0.89	0.89	
weighted avg	0.89	0.89	0.89	

ตารางที่ 17 ประสิทธิภาพการทำนายของตัวแบบมีค่าความแม่นยำเท่ากับ 0.89 สำหรับค่าทำนายรอดชีวิตมีค่าความเที่ยงเท่ากับ 0.89 ค่าเรียกคืนเท่ากับ 0.87 และค่าความถ่วงดุลเท่ากับ 0.88

ส่วนค่าทำนายเสียชีวิตมีค่าความเที่ยงเท่ากับ 0.89 ค่าเรียกคืนเท่ากับ 0.92 และค่าความถ่วงดุลเท่ากับ 0.90



รูปที่ 1 AUC ของวิธีการวิเคราะห์การถดถอยลอจิสติก

จากรูปที่ 1 พบว่า กราฟ ROC curve มีลักษณะโค้งงอเข้าใกล้มุมบนซ้าย และมีค่าพื้นที่ใต้โค้งเท่ากับ 0.89 ซึ่งมีค่าเข้าใกล้ 1 แสดงว่าตัวแบบวิธีการวิเคราะห์การถดถอยลอจิสติก มีความเหมาะสม

4.2 วิธีต้นไม้ตัดสินใจ

ตารางที่ 18 เมตริกซ์ความสับสนจากการทำนายการเสียชีวิตของผู้ป่วยผู้ป่วยโควิด-19 ด้วยวิธีต้นไม้ตัดสินใจ

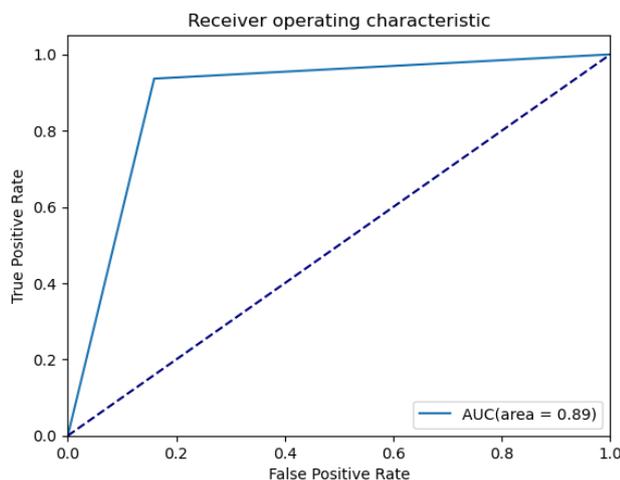
		ค่าทำนาย	
		รอดชีวิต	เสียชีวิต
ค่าจริง	รอดชีวิต	132	25
	เสียชีวิต	12	177

ตารางที่ 18 มีข้อมูล 346 ราย ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 309 ราย โดยมีจำนวนข้อมูลที่จำแนกถูกว่ารอดชีวิต 132 ราย และข้อมูลที่จำแนกถูกว่าเสียชีวิต 177 ราย และตัวแบบทำนายข้อมูลไม่ถูกต้อง 37 ราย โดยมีจำนวนข้อมูลที่จำแนกผิดว่ารอดชีวิต ซึ่งผลที่แท้จริงเสียชีวิต 12 ราย และข้อมูลที่จำแนกผิดว่าเสียชีวิต ซึ่งผลที่แท้จริงแล้วรอดชีวิต 25 ราย

ตารางที่ 19 ค่าประสิทธิภาพของตัวแบบจำแนกรายละเอียดในแต่ละคำตอบจากการทำนายการเสียชีวิตของผู้ป่วยผู้ป่วยโควิด-19 ด้วยวิธีต้นไม้ตัดสินใจ

Class	Precision	Recall	F1- score	Accuracy
0	0.92	0.84	0.88	0.89
1	0.88	0.94	0.91	
macro avg	0.90	0.88	0.89	
weighted avg	0.89	0.89	0.89	

ตารางที่ 19 ประสิทธิภาพการทำนายของตัวแบบมีความแม่นยำเท่ากับ 0.89 สำหรับค่าทำนายรอดชีวิตมีความเที่ยงเท่ากับ 0.92 ค่าเรียกคืนเท่ากับ 0.84 และค่าความถ่วงดุลเท่ากับ 0.88 ส่วนค่าทำนายเสียชีวิตมีความเที่ยงเท่ากับ 0.88 ค่าเรียกคืนเท่ากับ 0.94 และค่าความถ่วงดุลเท่ากับ 0.91



รูปที่ 2 AUC ของวิธีต้นไม้ตัดสินใจ

จากรูปที่ 2 พบว่า กราฟ ROC curve มีลักษณะโค้งเข้าใกล้มุมบนซ้าย และมีค่าพื้นที่ใต้โค้งเท่ากับ 0.89 ซึ่งมีค่าเข้าใกล้ 1 แสดงว่าตัวแบบวิธีต้นไม้ตัดสินใจ มีความเหมาะสม

4.3 วิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น

ตารางที่ 20 เมตริกซ์ความสับสนจากการทำนายการเสียชีวิตของผู้ป่วยผู้ป่วยโควิด-19 ด้วยวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น

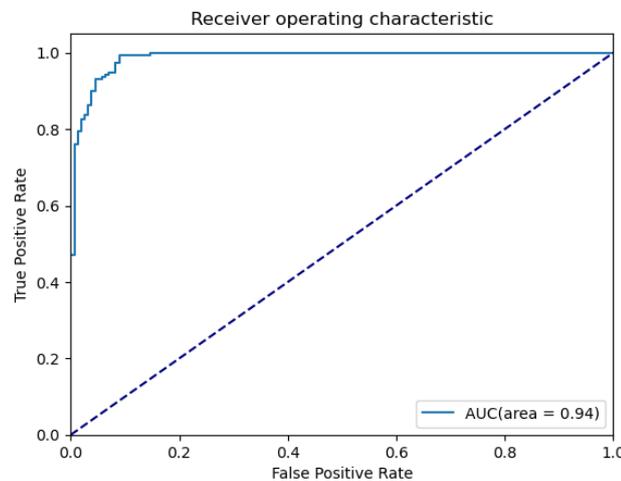
		ค่าทำนาย	
		รอดชีวิต	เสียชีวิต
ค่าจริง	รอดชีวิต	139	18
	เสียชีวิต	0	189

ตารางที่ 20 มีข้อมูล 346 ราย ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 328 ราย โดยมีจำนวนข้อมูลที่จำแนกถูกว่ารอดชีวิต 139 ราย และข้อมูลที่จำแนกถูกว่าเสียชีวิต 189 ราย และตัวแบบทำนายข้อมูลไม่ถูกต้อง 18 ราย โดยทั้งหมดเป็นการจำแนกผิดว่าเสียชีวิต ซึ่งผลที่แท้จริงแล้วรอดชีวิต 18 ราย

ตารางที่ 21 ค่าประสิทธิภาพของตัวแบบจำแนกรายละเอียดในแต่ละคำตอบจากการทำนายการเสียชีวิตของผู้ป่วยผู้ป่วยโควิด-19 ด้วยวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น

Class	Precision	Recall	F1- score	Accuracy
0	1.00	0.89	0.94	0.95
1	0.91	1.00	0.95	
macro avg	0.96	0.94	0.95	
weighted avg	0.95	0.95	0.95	

ตารางที่ 21 ประสิทธิภาพการทำนายของตัวแบบมีค่าความแม่นยำเท่ากับ 0.95 สำหรับการทำนายรอดชีวิตมีค่าความเที่ยงเท่ากับ 1.00 ค่าเรียกคืนเท่ากับ 0.89 และค่าความถ่วงดุลเท่ากับ 0.94 ส่วนการทำนายเสียชีวิตมีค่าความเที่ยงเท่ากับ 0.91 ค่าเรียกคืนเท่ากับ 1.00 และค่าความถ่วงดุลเท่ากับ 0.95



รูปที่ 3 AUC ของวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น

จากรูปที่ 3 พบว่า กราฟ ROC curve มีลักษณะโค้งเข้าใกล้มุมบนซ้าย และมีค่าพื้นที่ใต้โค้งเท่ากับ 0.94 ซึ่งมีค่าเข้าใกล้ 1 แสดงว่าตัวแบบวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น มีความเหมาะสม

5 ผลการเปรียบเทียบประสิทธิภาพในการทำนาย

การเปรียบเทียบประสิทธิภาพในการทำนายการเสียชีวิตของผู้ป่วยผู้ป่วยโควิด-19 ระหว่างวิธีการวิเคราะห์การถดถอยลอจิสติก วิธีต้นไม้ตัดสินใจ และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น โดยพิจารณาจากค่าความแม่นยำ ค่าความเที่ยง ค่าเรียกคืน และค่าความถ่วงดุล

ตารางที่ 22 ผลการเปรียบเทียบประสิทธิภาพในการทำนายของวิธีการจำแนกกลุ่มทั้ง 3 วิธี

วิธีจำแนกกลุ่ม	Accuracy	Precision	Recall	F1-score
วิธีการวิเคราะห์การถดถอยลอจิสติก	0.89	0.89	0.92	0.90
วิธีต้นไม้ตัดสินใจ	0.89	0.88	0.94	0.91
วิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น	0.95	0.91	1.00	0.95

ตารางที่ 22 พบว่า วิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น มีค่าความแม่นยำมากที่สุดเท่ากับ 0.95 มีค่าความเที่ยงเท่ากับ 0.91 ค่าเรียกคืนเท่ากับ 1.00 และค่าความถ่วงดุลเท่ากับ 0.95 ดังนั้นวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น มีประสิทธิภาพในการทำนายดีที่สุด

ผลการศึกษาปัจจัยที่สัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19 กรณีศึกษาโรงพยาบาลระยองโดยพิจารณาจากค่าความแม่นยำ พบว่าวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้นมีประสิทธิภาพในการทำนายผลดีที่สุดซึ่งสอดคล้องกับงานวิจัยของ Agbelusi and Olayemi^[16] ได้พัฒนาตัวแบบในการทำนายอัตราการเสียชีวิตของผู้ป่วยติดเชื้อไวรัสโคโรนาในประเทศไนจีเรีย โดยใช้เทคนิคการทำเหมืองข้อมูล 4 วิธี ประกอบด้วยวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น วิธีนาอ์ฟเบสส์ วิธีต้นไม้ตัดสินใจ และกฎการตัดสินใจ ผลการศึกษาแสดงให้เห็นว่าวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น มีประสิทธิภาพในการทำนายสูงที่สุด โดยมีค่าความแม่นยำที่ร้อยละ 85 และยังสอดคล้องกับงานวิจัยของ Khan et al.^[17] ได้นำตัวแบบการเรียนรู้เชิงลึกที่มีการกำหนดชั้นซ่อนทั้งหมด 6 ชั้น โดยมีฟังก์ชันเรคตีไฟด์ลินเนียยูนิต เป็นฟังก์ชันกระตุ้น และฟังก์ชันซิกมอยด์ เป็นกระตุ้นในชั้นท้ายสุด รวมไปถึงตัวแบบการเรียนรู้ของเครื่อง ได้แก่ วิธีต้นไม้ตัดสินใจ วิธีการวิเคราะห์การถดถอยลอจิสติก วิธีป่าสุ่ม

วิธีเอ็กซ์ทริมกาเตียนบูตติง และวิธีเพื่อนบ้านใกล้สุด K ตัว เพื่อใช้ทำนายโอกาสเสียชีวิตในผู้ป่วยโควิด-19 โดยใช้ข้อมูลผู้ป่วยที่รับการยืนยันว่าติดเชื้อไวรัสโคโรนาจากทั้งหมด 146 ประเทศพบว่าตัวแบบการเรียนรู้เชิงลึกมีประสิทธิภาพในการทำนายสูงที่สุด ด้วยค่าความแม่นยำ 0.97

ส่วนปัจจัยที่สัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19 จากการทำนายด้วยตัวแบบวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น พบว่าการใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ เป็นปัจจัยที่สำคัญที่สุดที่ส่งผลต่อการเสียชีวิตของผู้ป่วยโควิด-19 ซึ่งสอดคล้องกับการศึกษาของ Moulaei et al.^[19] ได้ศึกษาผู้ป่วยโควิด-19 จำนวน 1,500 ราย ของโรงพยาบาล Ayatollah Taleghani เมือง Abadan ประเทศอิหร่าน โดยใช้ขั้นตอนวิธีการเรียนรู้ของเครื่อง ในการสร้างตัวแบบทำนายพบว่าจากตัวแปรทำนายทั้งหมด 38 ตัวแปรภาวะหายใจลำบาก การเข้ารับการรักษาในห้องไอซียู และการได้รับการบำบัดด้วยออกซิเจน เป็นตัวแปรทำนาย 3 อันดับแรกที่ส่งผลต่อโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 มากที่สุด นอกจากนี้ เหตุผลทางชีวการแพทย์ยังแสดงให้เห็นว่า ผู้ป่วยที่จำเป็นต้องใช้เครื่องช่วยหายใจมักอยู่ในระยะที่โรครุนแรงถึงขั้นวิกฤต โดยเฉพาะในกรณีที่เกิดภาวะ ARDS (Acute Respiratory Distress Syndrome) ซึ่งเป็นภาวะที่มีอัตราการเสียชีวิตสูง ทั้งนี้ การใช้เครื่องช่วยหายใจอาจก่อให้เกิด

ภาวะแทรกซ้อนต่าง ๆ เช่น การติดเชื้อในปอด จากเครื่องช่วยหายใจ (Ventilator-associated pneumonia), การบาดเจ็บจากแรงดันอากาศ (Barotrauma) และ ภาวะกล้ามเนื้ออ่อนแรงจากการพึ่งพาเครื่องเป็นเวลานาน ซึ่งล้วนส่งผลให้การพยากรณ์โรคแย่ลง และเพิ่มความเสี่ยงต่อการเสียชีวิตของผู้ป่วย

สรุป

การศึกษาปัจจัยที่สัมพันธ์ต่อการเสียชีวิตของผู้ป่วยโควิด-19 กรณีศึกษาโรงพยาบาลระยอง มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่สัมพันธ์กับการเสียชีวิตของผู้ป่วยโควิด-19 จากข้อมูลผู้ป่วยโควิด-19 จำนวน 1,006 ราย ที่เข้ารับการรักษา ณ โรงพยาบาลระยอง โดยมีตัวแปรอิสระทั้งหมด 46 ตัวแปร นำไปสร้างตัวแบบในการทำนายโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 ด้วยวิธีการวิเคราะห์การถดถอยลอจิสติก วิธีต้นไม้ตัดสินใจ และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น ผลสรุปดังนี้

ประสิทธิภาพของตัวแบบวิธีการวิเคราะห์การถดถอยลอจิสติกในการทำนายชุดข้อมูลทดสอบพบว่า มีค่าความแม่นยำเท่ากับ 0.89 ค่าความเที่ยงเท่ากับ 0.89 ค่าเรียกคืนเท่ากับ 0.92 และค่าความถ่วงดุลเท่ากับ 0.90 โดยตัวแปรที่ส่งผลต่อโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 จากการทำนายด้วยตัวแบบวิธีการวิเคราะห์การถดถอยลอจิสติก ได้แก่ อายุ สัญชาติ การใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ โรคไต และการใช้ยาเมทิลเพรดนิโซโลน ตามลำดับ

ประสิทธิภาพของตัวแบบวิธีต้นไม้ตัดสินใจในการทำนายชุดข้อมูลทดสอบพบว่า มีค่าความแม่นยำเท่ากับ 0.89 ค่าความเที่ยงเท่ากับ 0.88 ค่าเรียกคืนเท่ากับ 0.94 และ

ค่าความถ่วงดุลเท่ากับ 0.91 โดยในส่วนของตัวแปรอิสระ ได้แก่ จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา อายุ การใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ อัตราการหายใจ และปริมาณออกซิเจนในเลือดเป็นตัวแปรอิสระที่มีความสำคัญในการทำนายโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 ด้วยตัวแบบวิธีต้นไม้ตัดสินใจ

ประสิทธิภาพของตัวแบบวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น ในการทำนายชุดข้อมูลทดสอบพบว่า มีค่าความแม่นยำเท่ากับ 0.95 ค่าความเที่ยงเท่ากับ 0.91 ค่าเรียกคืนเท่ากับ 1.00 และค่าความถ่วงดุลเท่ากับ 0.95 โดยตัวแปรอิสระที่มีความสำคัญสูงสุด 5 อันดับแรก ได้แก่ การใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ มีค่าความสำคัญมากที่สุด รองลงมาคือ โรคไต อายุ จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษา และการใช้ยาไฮโดรคอร์ติโซนในการรักษา ตามลำดับ

จากประสิทธิภาพของตัวแบบทั้ง 3 วิธีในการทำนายชุดข้อมูลทดสอบพบว่า วิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น มีประสิทธิภาพสูงสุด รองลงมาคือวิธีต้นไม้ตัดสินใจ และวิธีการวิเคราะห์การถดถอยลอจิสติก ตามลำดับ โดยตัวแปรอิสระที่มีความสำคัญสูงสุด 5 อันดับแรกจากตัวแบบทั้ง 3 วิธีพบว่า อายุ และการใช้เครื่องช่วยหายใจแบบใส่ท่อช่วยหายใจ เป็นตัวแปรอิสระที่มีความสำคัญสูงสุดในการทำนายโอกาสการเสียชีวิตของผู้ป่วยโควิด-19 ด้วยตัวแบบทั้ง 3 วิธี โรคไต เป็นตัวแปรอิสระที่มีความสำคัญในการทำนายด้วยตัวแบบวิธีการวิเคราะห์การถดถอยลอจิสติก และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น จำนวนวันนับตั้งแต่วันที่ตรวจพบเชื้อจนถึงวันจำหน่ายการรักษาเป็น

ตัวแปรอิสระที่มีความสำคัญในการทำนายด้วยตัวแบบวิธีต้นไม้ตัดสินใจ และวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น สัญชาติและการใช้ยาเมทิลเพรดนิโซโลน เป็นตัวแปรอิสระที่มีความสำคัญในการทำนายด้วยตัวแบบวิธีการวิเคราะห์การถดถอยลอจิสติก อัตราการหายใจ และปริมาณออกซิเจนในเลือด เป็นตัวแปรอิสระที่มีความสำคัญในการทำนายด้วยตัวแบบวิธีต้นไม้ตัดสินใจ และการใช้ยาไฮโดรคอร์ติโซนในการรักษาเป็นตัวแปรอิสระที่มีความสำคัญในการทำนายด้วยตัวแบบวิธีโครงข่ายประสาทเทียมแบบหลายลำดับชั้น

ข้อจำกัดของการศึกษา

แม้ว่าผลการศึกษจะสามารถชี้ให้เห็นปัจจัยที่มีความสัมพันธ์ต่อการเสียชีวิตของผู้ป่วย COVID-19 ได้อย่างชัดเจน แต่อย่างไรก็ตาม การวิจัยครั้งนี้ยังมีข้อจำกัดบางประการที่ควรพิจารณา ได้แก่ ความเอนเอียง (Bias) ที่อาจเกิดจากการประมวลผลข้อมูลด้วยเทคนิค SMOTE ซึ่งมีจุดประสงค์เพื่อแก้ปัญหาคลาสไม่สมดุล แต่อาจส่งผลให้เกิดข้อมูลเทียมที่ไม่สะท้อนความเป็นจริงอย่างแท้จริง นอกจากนี้ ข้อมูลที่ใช้ในการศึกษาได้มาจากโรงพยาบาลเพียงแห่งเดียว จึงอาจจำกัดความสามารถในการ นำผลการศึกษาไปใช้กับโรงพยาบาลหรือบริบททางคลินิกอื่น ๆ ที่มีลักษณะประชากรหรือแนวทางการรักษาที่แตกต่างกัน

ข้อเสนอแนะในการวิจัยครั้งต่อไป

แนวทางการศึกษาวิจัยในอนาคต จะพิจารณาตัวแปรอิสระอื่น ๆ ที่เกี่ยวข้องเพิ่มเติมเช่น ผลการตรวจทางห้องปฏิบัติการ อาการแสดงของผู้ป่วย เป็นต้น เพื่อให้ตัวแบบมีประสิทธิภาพเพิ่มมากขึ้น

เอกสารอ้างอิง

1. กรมควบคุมโรค. สถานการณ์โรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) มาตรการสาธารณสุขและปัญหาอุปสรรคการป้องกันควบคุมโรคในผู้เดินทาง; [อินเทอร์เน็ต]. 2564 [เข้าถึง วันที่ 18 สิงหาคม 2564]. เข้าถึงได้จาก : https://ddc.moph.go.th/uploads/files/2017420210_820025238.pdf.
2. คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล. ความรู้พื้นฐาน COVID-19 ตอนที่ 1; [อินเทอร์เน็ต]. 2563 [เข้าถึง วันที่ 20 กันยายน 2564]. เข้าถึงได้จาก : <https://www.rama.mahidol.ac.th/ceb/News/2020/Covid-19>.
3. กรมควบคุมโรค. รายงานสถานการณ์โรคติดเชื้อไวรัสโคโรนา 2019 ฉบับที่ 58 วันที่ 1 มีนาคม 2563; [อินเทอร์เน็ต]. 2563 [เข้าถึง วันที่ 18 สิงหาคม 2564]. เข้าถึงได้จาก : <https://ddc.moph.go.th/viralpneumonia/file/situation/situation-no58-010363.pdf>.
4. กรมควบคุมโรค. สถานการณ์โรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) มาตรการสาธารณสุขและปัญหาอุปสรรคการป้องกันควบคุมโรคในผู้เดินทาง วันที่ 14 มกราคม 2564; [อินเทอร์เน็ต]. 2564 [เข้าถึง วันที่ 31 สิงหาคม 2564]. เข้าถึงได้จาก : [https://ddc.moph.go.th/uploads/ckeditor2/files/1%20วิเคราะห์สถานการณ์%20covid%20\(15%20มค64cp\).pdf](https://ddc.moph.go.th/uploads/ckeditor2/files/1%20วิเคราะห์สถานการณ์%20covid%20(15%20มค64cp).pdf).
5. คณะกรรมการโรคติดต่อจังหวัดระยอง. แฉลงการณ์จังหวัดระยอง ฉบับที่ 1 เรื่อง สถานการณ์ โรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) ในพื้นที่จังหวัดระยอง วันที่ 19 มีนาคม 2563.

6. สาธิต ปิตุเตชะ. ผู้ป่วยโควิด-19 เพศชาย อายุ 45 ปี เสียชีวิต 1 ราย ที่ระยอง; [อินเทอร์เน็ต]. 2563 [เข้าถึง วันที่ 31 สิงหาคม 2564]. เข้าถึงได้จาก: https://www.facebook.com/permalink.php?story_fbid=3648069868612657&id=119568378129508.
7. โรงพยาบาลวิชัยเวช อินเตอร์เนชั่นแนล หนองแขม. โควิดสายพันธุ์โอไมครอนคืออะไร น่ากลัวแค่ไหน; [อินเทอร์เน็ต]. 2564 [เข้าถึง วันที่ 31 สิงหาคม 2564]. เข้าถึงได้จาก : <http://vichaivej-nongkhaem.com/healthinfo/โควิดโอไมครอน/>.
8. โรงพยาบาลพริ้นซ์ สุวรรณภูมิ. เรื่องควรรู้เกี่ยวกับเชื้อไวรัสโควิด-19 สายพันธุ์โอไมครอน (Omicron); [อินเทอร์เน็ต]. 2564 [เข้าถึง วันที่ 31 สิงหาคม 2564]. เข้าถึงได้จาก: <https://www.princsuvarnabhumi.com/conten-nt-omicron/>.
9. กรมควบคุมโรค. รายงานสถานการณ์โรคติดเชื้อไวรัสโคโรนา 2019 ฉบับที่ 724 วันที่ 27 ธันวาคม 2564; [อินเทอร์เน็ต]. 2564 [เข้าถึง วันที่ 31 ธันวาคม 2564]. เข้าถึงได้จาก: <https://ddc.moph.go.th/viralpneumonia/file/situation/situation-no724-271264.pdf>.
10. กรมการแพทย์. แนวทางเวชปฏิบัติ การวินิจฉัย ดูแลรักษา และป้องกันการติดเชื้อในโรงพยาบาล กรณีโรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) สำหรับแพทย์และบุคลากรสาธารณสุข ฉบับปรับปรุง วันที่ 22 เมษายน 2565; [อินเทอร์เน็ต]. 2564 [เข้าถึง วันที่ 30 เมษายน 2565]. เข้าถึงได้จาก : https://covid19.dms.go.th/backend/Content/Content_File/Covid_Health/Attach/25650422162203PM_CPG_COVID-19_n_v.22_20220422.pdf.
11. Das AK, Mishra S, and Gopalan SS. Predicting CoVID-19 community mortality risk using machine learning and development of an online prognostic tool. *PeerJ* 2020; 8: e10083.
12. Sánchez-Montañés M, Rodríguez-Belenguer P, Serrano-López AJ, Soria-Olivas E, Alakhdar-Mohmara Y. Machine learning for mortality analysis in patients with COVID-19. *Int J Environ Res Public Health* 2020; 17(22): 8386.
13. Ferreira AT, Fernandes C, Vieira J, Portela F. Pervasive intelligent models to predict the outcome of COVID-19 patients. *Future Internet* 2021; 13(4): 102.
14. Muhammad LJ, Algehyne EA, Usman SS, Ahmad A, Chakraborty C, Mohammed IA. Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset. *SN Comput Sci* 2021; 2(1): 1-13.
15. Li X, Ge P, Zhu J, Li H, Graham J, Singer A, Richman PS, Duong TQ. Deep learning prediction of likelihood of ICU admission and mortality in COVID-19 patients using clinical variables. *PeerJ*. 2020; 8: e10337.
16. Agbelusi O, Olayemi OC. Prediction of mortality rate of COVID-19 patients using machine learning techniques in nigeria. *Int J Comput Sci Softw Eng* 2020; 9(5): 30-34.
17. Khan IU, Aslam N, Aljabri M, Aljameel, SS, Kamaleldin MMA, Alshamrani FM, Chrouf SMB. Computational intelligence-based model for mortality rate prediction in COVID-19 patients. *Int J Environ Res Public Health* 2021; 18(12): 6429.
18. Pourhomayoun M, Shakibi M. Predicting mortality risk in patients with COVID-19 using machine learning to help medical

- decision-making. *Smart Health* 2021; 20: 100178.
19. Moulaei K, Shanbehzadeh M, Mohammadi-Taghiabad Z, Kazemi-Arpanahi H. Comparing machine learning algorithms for predicting COVID-19 mortality. *BMC Med Inform Decis Mak* 2022; 22(1): 1-12.
20. Stevens J. *Applied multivariate statistics for the social sciences*. 3rd ed. Mahwah (NJ): Lawrence Erlbaum Associates; 1996.