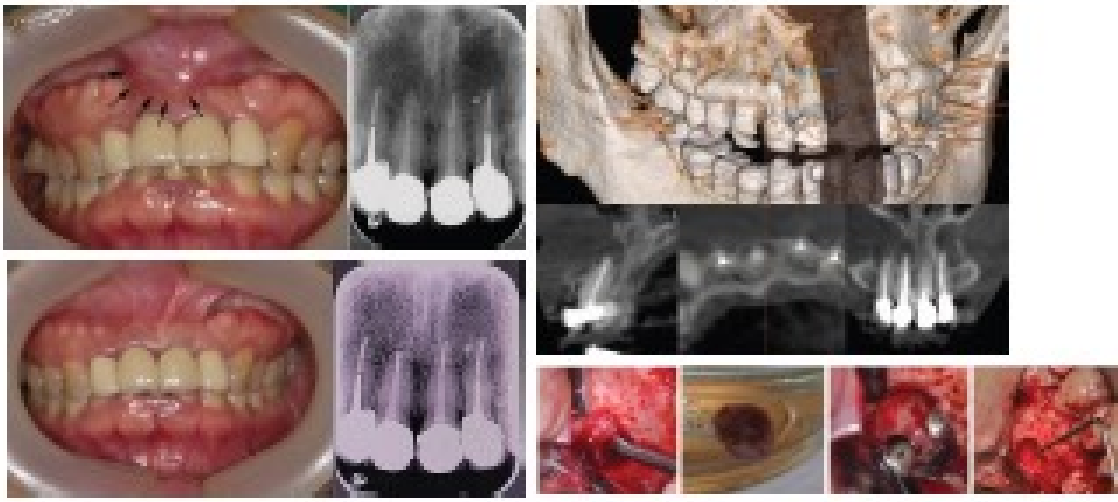
Volume 2 (2) July-December 2023

Thai Endodontic Journal (ENDOSARN)

The official journal of the Thai Endodontic Association

TEJ



Interesting articles

- *การเรียนรู้ของเครื่องและการประยุกต์ใช้ทางวิทยาเอ็นโดดอนต์ (Machine Learning and its Application in Endodontics)*
- *Apical Debris Extrusion in Endodontics: Unraveling Influential Factors*
- *การทำศัลยกรรมปลายรากฟันร่วมกับการจัดการปุ่มกระดูกส่วนงอกของขากรรไกรที่มีขนาดใหญ่: รายงานผู้ป่วย (Endodontic Microsurgery on Tooth 11 Concurrent with a Large Exostosis Removal: A Case Report)*
- *สถิติกับงานวิทยาเอ็นโดดอนต์ ตอนที่ 2: ความเชื่อถือได้ (Statistics in Endodontics Part 2: Reliability)*

Text and Journal Publication Co., LTD.

ISSN 1685-3709

การเรียนรู้ของเครื่องและการประยุกต์ใช้ทางวิทยาเอ็นโดดอนต์ (Machine Learning and its Application in Endodontics)

ปกิต ตุ้งสวัสดิ์ (Pakit Tungsawat)
College of Dental Medicine, Rangsit University

บทคัดย่อ

การเรียนรู้ของเครื่อง (machine learning) เป็นศาสตร์สาขาหนึ่งทางปัญญาประดิษฐ์ที่เริ่มเข้ามามีบทบาทในทางทันตกรรมจากความก้าวหน้าทางวิทยาการคอมพิวเตอร์และวิทยาการข้อมูล เริ่มมีการนำมาประยุกต์ใช้ในทางวิทยาเอ็นโดดอนต์ เช่น การระบุรอยโรครอบปลายรากในภาพรังสี การจำแนกลักษณะคลองราก การตรวจสอบรากฟันแตก เป็นต้น ส่วนใหญ่จะใช้การเรียนรู้ของเครื่องแบบ supervised machine learning โดยมนุษย์จัดเตรียมชุดข้อมูลนำเข้าพร้อมป้ายกำกับให้คอมพิวเตอร์เรียนรู้โดยใช้สมการทางสถิติศาสตร์และสร้างเป็นโมเดลที่ซับซ้อน จากนั้นประเมินประสิทธิภาพของโมเดลที่สร้างขึ้นเมื่อมีประสิทธิภาพเพียงพอจึงนำไปใช้งานจริง ข้อมูลส่วนใหญ่ที่ใช้ในงานวิจัยทางวิทยาเอ็นโดดอนต์จะเป็นข้อมูลรูปภาพและภาพรังสีที่จะใช้วิธีการเรียนรู้แบบ convolution neural network (CNN) ในการสกัดข้อมูลสำคัญเพื่อสร้างโมเดล ผลการศึกษาส่วนใหญ่พบว่าการเรียนรู้เครื่องมีประสิทธิภาพ เพิ่มความแม่นยำในการปฏิบัติงานและลดเวลาการตัดสินของมนุษย์ มีแนวโน้มที่สามารถนำมาใช้เสริมการปฏิบัติงานให้มีประสิทธิภาพมากขึ้นได้

คำสำคัญ: การเรียนรู้ของเครื่อง, การเรียนรู้เชิงลึก, ประสิทธิภาพโมเดล, วิทยาเอ็นโดดอนต์, การรักษาคลองราก

Abstract

With developments in computer science and data science, the field of artificial intelligence known as machine learning entered the dental field. It has been used in endodontics, for example, to locate periapical lesions in the radiographs, Identification of the root canal morphology, detection of root fracture etc. The majority of studies employ supervised machine learning, wherein humans provide a labeled input data set for the computer to train using statistical equations, producing intricate mathematical models. Once the created model has proven to be sufficiently efficient, it is then put into use, and its performance is reviewed. Convolutional neural network (CNN) learning techniques are utilized to extract crucial data for model creation from imaging and radiography data, which make up most of the data used in endodontic research. Most research has concluded that machine learning is effective. Enhance decision-making speed while increasing operational precision. It might serve as a supplementary mechanism to improve operational efficiency.

Keywords: deep learning, endodontics, machine learning, model evaluation, root canal treatment

Correspondence: Pakit Tungsawat
College of Dental Medicine, Rangsit University, 52/347 Muang-Ake, Phaholyothin Road, Lak-Hok, Muang, Pathum Thani 12000 Thailand.
Email address: pakit.tu@rsu.ac.th
Received: 13 March 2023 Revised: 15 August 2023 Accepted: 12 December 2023

บทนำ

ความก้าวหน้าทางวิทยาการ เทคโนโลยีและอุปกรณ์คอมพิวเตอร์ในปัจจุบันทำให้ปัญญาประดิษฐ์ (Artificial intelligence; AI) เริ่มเข้ามามีบทบาทในชีวิตประจำวันมากขึ้น (1) เช่น ระบบการจดจำใบหน้าของโทรศัพท์เคลื่อนที่ ระบบขับเคลื่อนรถยนต์อัตโนมัติ ระบบแนะนำสินค้าและบริการ เป็นต้น งานทางทันตกรรมเริ่มมีการศึกษาวิจัยเพื่อประยุกต์ใช้ AI ตั้งแต่ช่วงปี ค.ศ.2010 (2) และมีแนวโน้มที่เพิ่มมากขึ้นมีการประยุกต์ใช้เพื่อวินิจฉัยจำแนกรอยโรค (3, 4) ประเมินความเสี่ยง (5) วางแผนการรักษา (6, 7) การพยากรณ์โรค (8) เป็นต้น สาขาวิทยาเอ็นโดดอนต์เป็นสาขาหนึ่งที่เริ่มนำ AI มาประยุกต์ใช้ในงานทางคลินิกเพิ่มมากขึ้น องค์ความรู้พื้นฐานทางศาสตร์ด้านวิทยาการข้อมูลและ AI จึงเป็นสิ่งจำเป็นเพื่อความเข้าใจและประเมินประสิทธิภาพของระบบ AI ที่จะนำมาใช้งานในอนาคต บทความนี้จึงมีวัตถุประสงค์เพื่อรวบรวมองค์ความรู้พื้นฐานเกี่ยวกับวิทยาการข้อมูลและ AI ที่จำเป็นโดยเฉพาะในเรื่องการเรียนรู้ของเครื่อง (Machine learning; ML) เพื่อความเข้าใจหลักการทำงานพื้นฐาน การประเมินประสิทธิภาพของโมเดลที่จะนำมาใช้งาน การประยุกต์ใช้ การพัฒนาต่อยอด และขอบเขตงานวิจัยในสาขาวิทยาเอ็นโดดอนต์ในปัจจุบัน

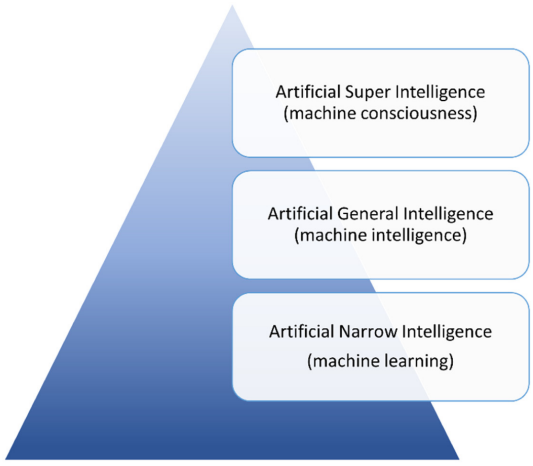
ปัญญาประดิษฐ์ (Artificial intelligence)

คำว่าปัญญาประดิษฐ์เริ่มต้นใช้ตั้งแต่ปี ค.ศ. 1956 โดย John McCarthy นักวิทยาศาสตร์คอมพิวเตอร์ชาวอเมริกันที่ต่อยอดมาจากแนวคิดของ Alan Turing ผู้ซึ่งประดิษฐ์คอมพิวเตอร์ขึ้นในช่วงสงครามโลกครั้งที่สองเพื่อถอดรหัสจากข้อความที่สื่อสารกันของฝ่ายเยอรมันได้สำเร็จโดยที่มนุษย์ไม่สามารถถอดรหัสได้ โดยในปี ค.ศ. 1950 Alan Turing (9) ได้เสนอแนวคิดว่าคอมพิวเตอร์เมื่อพัฒนาให้มีประสิทธิภาพมากขึ้นจะมีความสามารถทางปัญญาเสมือนมนุษย์ โดยเสนอทฤษฎีทดสอบว่าคอมพิวเตอร์มีปัญญาเท่ากับมนุษย์หรือไม่โดยใช้การทดสอบที่เรียกว่า การทดสอบทัวริง (10) (Turing test) มีวิธีการทดสอบโดยจัดให้ผู้ทดสอบพิมพ์สนทนาโต้ตอบกับคอมพิวเตอร์เปรียบเทียบกับการสนทนามนุษย์ที่อยู่อีกห้องหนึ่งถ้าผู้ทดสอบไม่สามารถระบุได้ว่ามนุษย์และคอมพิวเตอร์อยู่ในห้องใด แสดงว่าคอมพิวเตอร์มีความสามารถทางปัญญาเท่ากับมนุษย์ โดยทั่วไป AI สามารถแบ่งได้เป็น 3 ประเภท (11) ประกอบด้วย Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI), และ Artificial Super Intelligence (ASI) (รูปที่ 1)

Artificial Narrow Intelligence (ANI) หรือ weak AI คือ AI ที่มีความสามารถในการทำงานจำกัดหรือทำได้เพียงอย่างเดียว เป็น AI ชนิดที่พบได้ทั่วไปในปัจจุบัน เช่น ระบบจดจำใบหน้า ระบบแนะนำภาพยนตร์หรือเพลง ระบบช่วยพิมพ์คำให้สมบูรณ์โดยอัตโนมัติ ระบบตรวจสอบแกรมมา ระบบกรองอีเมลขยะ ระบบเหล่านี้ใช้ AI ปฏิบัติงานได้เฉพาะอย่างโดยมีพื้นฐานการทำงานจากเงื่อนไขเบื้องต้นที่กำหนดไว้ AI ชนิดนี้สร้างโดยการฝึก (training) ด้วยข้อมูลนำเข้าที่ป้อนให้จึงไม่สามารถทำงานอื่นที่อยู่นอกเหนือขอบเขตได้

Artificial General Intelligence (AGI) หรือ strong AI เป็น AI ชนิดที่มีความสามารถหลากหลาย สามารถคิดได้ด้วยตัวเอง ไม่ต้องการข้อมูลนำเข้าจากมนุษย์ สามารถเรียนรู้ได้ด้วยตนเอง มีตรรกะและอารมณ์ ในปัจจุบันAGI นี้ยังคงเป็นเพียงทฤษฎียังไม่มีโมเดลที่สร้างขึ้นได้จริงจึงยังต้องการการพัฒนาต่อไป

Artificial Super Intelligence (ASI) จัดเป็น strong AI อีกประเภทหนึ่งที่มีความสามารถ ความฉลาด ตรรกะ ความคิดสร้างสรรค์เหนือมนุษย์ ซึ่ง AI ประเภทนี้ยังคงเป็นเพียงทฤษฎีที่ปรากฏอยู่ในนวนิยายแนววิทยาศาสตร์

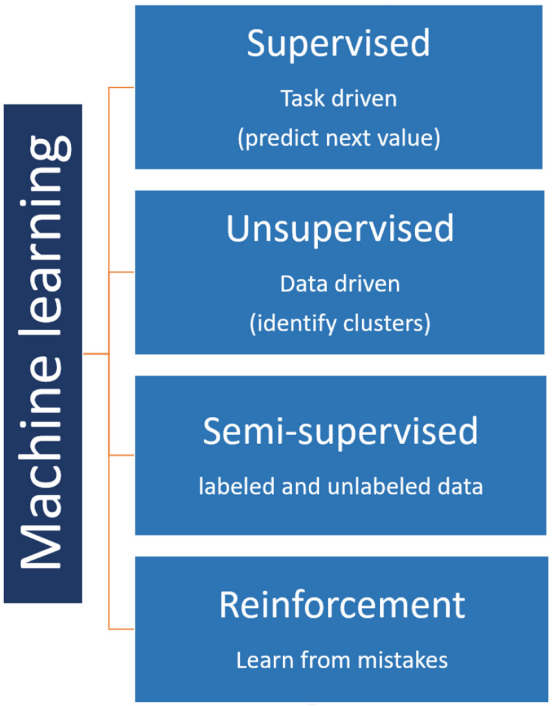


รูปที่ 1 Stages of artificial intelligence

การตัดสินใจของคอมพิวเตอร์ในช่วงเริ่มแรกจะเกิดจากการเขียนโปรแกรมป้อนกฎการตัดสินใจสำหรับสถานการณ์ที่คาดการณ์ว่าจะเกิดขึ้นและคอมพิวเตอร์จะตัดสินใจและตอบสนองต่อสถานการณ์ตามที่โปรแกรมระบุไว้ ตัวอย่างเช่น โปรแกรมในเครื่องแลกเหรียญที่ถูกเขียนเอาไว้อย่างชัดเจนว่าเมื่อเครื่องรับธนบัตรเข้ามาจะแลกเป็นเหรียญชนิดที่กำหนดออกไปจำนวนเท่ากัน หรือโปรแกรมในเครื่องซื้อตั๋วอัตโนมัติ เป็นต้น แต่ความฉลาดของคอมพิวเตอร์ที่มีพื้นฐานจากกฎการตัดสินใจที่ถูกป้อนเข้าโดยมนุษย์นี้มีข้อจำกัด เนื่องจากสถานการณ์ต้องเป็นไปตามที่กฎเขียนระบุไว้ คอมพิวเตอร์จึงสามารถตัดสินใจได้มิฉะนั้นจะเกิดความผิดพลาดขึ้นในระบบ และข้อมูลหรือสถานการณ์บางอย่างไม่สามารถเขียนออกมาเป็นกฎในการตัดสินใจได้ เช่น การเขียนกฎเพื่อจำแนกรอยโรคในภาพรังสีรอบปลายรากเพื่อจำแนกว่ามีรอยโรคหรือไม่ สถานการณ์ตัวอย่างนี้สามารถแก้ปัญหาโดยให้คอมพิวเตอร์เรียนรู้และตัดสินใจด้วยตัวเองโดยมนุษย์จะเป็นผู้ป้อนข้อมูลนำเข้าเป็นภาพรังสีรอบปลายรากพร้อมคำเฉลยว่ามีรอยโรคหรือไม่จำนวนหนึ่งให้คอมพิวเตอร์ได้เรียนรู้และสร้างแบบแผนการตัดสินใจขึ้นเองโดยมนุษย์ไม่ต้องเขียนกฎการตัดสินใจใด ๆ ให้คอมพิวเตอร์ การเรียนรู้ลักษณะนี้จึงถูกเรียกว่า การเรียนรู้ของเครื่อง (machine learning)

การเรียนรู้ของเครื่อง (machine learning; ML)

ML เป็นปัญญาประดิษฐ์ประเภทหนึ่งที่ทำนายหรือตัดสินใจโดยใช้อัลกอริทึมที่สร้างจากชุดข้อมูลตัวอย่างที่มนุษย์ป้อนเข้าแล้วเรียนรู้โดยใช้วิธีทางสถิติศาสตร์เพื่อสร้างเป็นโมเดลสำหรับทำนายหรือตัดสินใจได้เองโดยมนุษย์ไม่ต้องเขียนโปรแกรมสร้างกฎพื้นฐานสำหรับการตัดสินใจไว้อย่างชัดแจ้ง ML แบ่งออกเป็น 4 ประเภทตามลักษณะการเรียนรู้ (12) คือ supervised learning, unsupervised learning, semi-supervised และ reinforcement learning (รูปที่ 2)

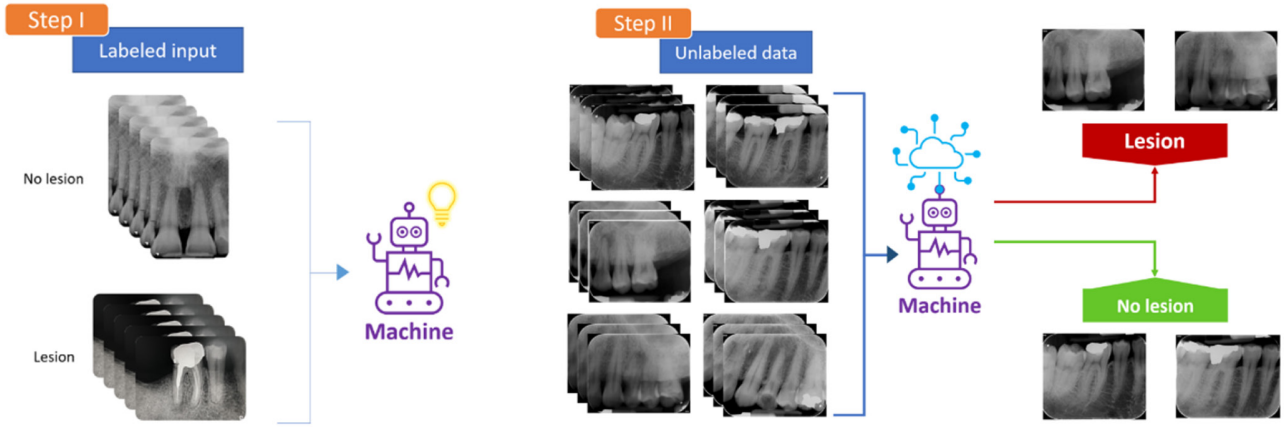


รูปที่ 2 ประเภทของการเรียนรู้ของเครื่อง (Types of machine learning)

Supervised learning เป็นวิธีการเรียนรู้โดยมนุษย์เป็นผู้ป้อนเข้าชุดข้อมูลตัวอย่างที่มีป้ายกำกับให้คอมพิวเตอร์เรียนรู้สร้างเป็นโมเดลการตัดสินใจ ตัวอย่างชุดข้อมูลเช่น รูปสุนัขและรูปแมวพร้อมระบุว่ารูปใดเป็นสัตว์ชนิดใด หรือรูปฟันที่ระบุว่าฟันซี่ใดมีรอยผุและฟันซี่ใดไม่มีรอยผุ (รูปที่ 3)

Unsupervised learning เป็นวิธีการเรียนรู้โดยมนุษย์ป้อนข้อมูลนำเข้าที่ไม่มีป้ายกำกับให้คอมพิวเตอร์เรียนรู้เอง การเรียนรู้ลักษณะนี้จะใช้สำหรับแก้ไขปัญหาในรูปแบบของการจัดกลุ่ม (clustering) หรือการหาความสัมพันธ์ (association)

Semi-supervised learning เป็นวิธีการเรียนรู้ที่ใช้ทั้งแบบ supervised และ unsupervised ร่วมกันคือข้อมูลนำเข้าที่ป้อนให้คอมพิวเตอร์จะมีทั้งกลุ่มที่มีป้ายกำกับและกลุ่มที่ไม่มีป้ายกำกับเนื่องจากไม่มีข้อมูลที่ระบุป้ายกำกับจำนวนมากพอที่จะให้คอมพิวเตอร์เรียนรู้แบ supervised วิธีนี้จะช่วยประหยัดต้นทุนและเวลาในการทำงาน



รูปที่ 3 วิธีการเรียนรู้ของเครื่องแบบ supervised learning *ขั้นตอน 1* จะป้อนข้อมูลที่ระบุป้ายกำกับแต่ละภาพให้เครื่องเรียนรู้ *ขั้นตอนที่ 2* เมื่อเครื่องเรียนรู้สำเร็จจะป้อนข้อมูลใหม่ที่ไม่ระบุป้ายกำกับเพื่อให้เครื่องระบุประเภทของข้อมูลแต่ละภาพ

Reinforcement learning เป็นการเรียนรู้ของคอมพิวเตอร์มีลักษณะคล้ายการลองผิดลองถูกของมนุษย์ โดยคอมพิวเตอร์จะทดลองตัดสินใจแก้ปัญหาและตรวจสอบผลลัพธ์จากนั้นนำมาปรับปรุงวิธีการตัดสินใจแก้ไขปัญหาให้มีประสิทธิภาพมากขึ้นไปเรื่อย ๆ ตัวอย่างเช่น โปรแกรม AlphaGO ที่พัฒนาโดยบริษัท Google DeepMind ซึ่งเป็นโปรแกรมคอมพิวเตอร์ที่เล่นเกมกระดานหมากล้อมชนะแชมป์โลกที่เป็นมนุษย์ได้

วิธีการเรียนรู้ของเครื่องแบบ supervised learning

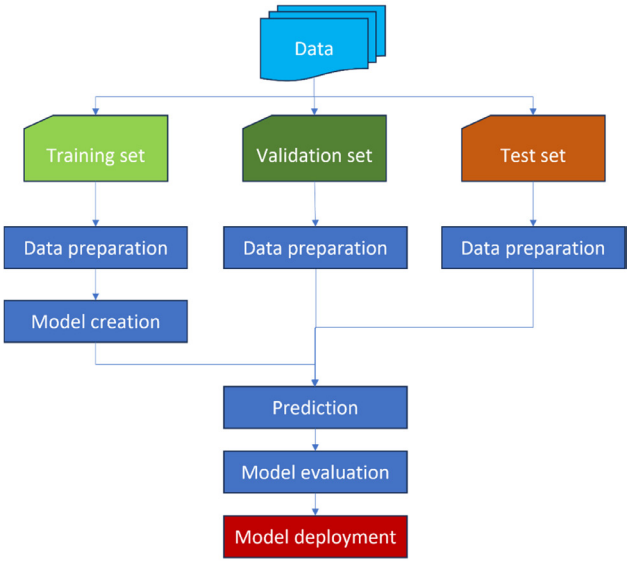
งานวิจัยด้านปัญญาประดิษฐ์ในทางทันตกรรมและวิทยาเอ็นโดดอนต์ในปัจจุบันใช้โมเดลการเรียนรู้ของเครื่องแบบ supervised learning ดังนั้นขอบเขตของบทความนี้จะครอบคลุมถึงวิธีการเรียนรู้ของเครื่องแบบ supervised learning เท่านั้น

การเรียนรู้ของเครื่องแบบ supervised learning เริ่มต้นจากมนุษย์เป็นผู้รวบรวมและจัดเตรียมชุดข้อมูลตัวอย่างที่ระบุป้ายกำกับให้มีความหลากหลายและมีจำนวนมากเพียงพอ จากนั้นนำเข้าสู่โปรแกรมเพื่อให้คอมพิวเตอร์เรียนรู้โดยแบ่งข้อมูลออกเป็นสามส่วนคือ ชุดฝึก (training set) ชุดตรวจสอบ (validation set) และชุดทดสอบ (test set) ในสัดส่วนที่เหมาะสม เช่น 70:15:15 หรือ 80:10:10 จากนั้นให้เครื่องเรียนรู้เพื่อสร้างโมเดลการตัดสินใจจากชุดฝึก จากชุดฝึกและทดสอบหาโมเดลที่ดีที่สุดโดยใช้ข้อมูลชุดสุดท้ายทดสอบว่าเครื่องสามารถเรียนรู้ได้สำเร็จหรือไม่โดยใช้ชุดทดสอบเป็นชุดข้อมูลให้เครื่องทดลองแก้ปัญหาและประเมินผลการเรียนรู้ของเครื่องจากตัวชี้วัดที่กำหนด (13-15) ขั้นตอนการเรียนรู้ดังรูปที่ 4

ลักษณะของปัญหาสำหรับโมเดลการตัดสินใจสามารถแบ่งออกเป็น 2 แบบตามพื้นฐานคณิตศาสตร์และสถิติที่ใช้แก้ปัญหา คือ การถดถอย (regression) และการจัดหมวดหมู่ (classification)

ปัญหาแบบ regression จะมีคำตอบเป็นเลขจำนวนจริงที่คำนวณได้จากสมการทำนาย เช่น การทำนายความเสี่ยงของโรคปริทันต์ (5) วิธีทางสถิติหลักที่ใช้ในสร้างโมเดล คือ สมการ linear regression

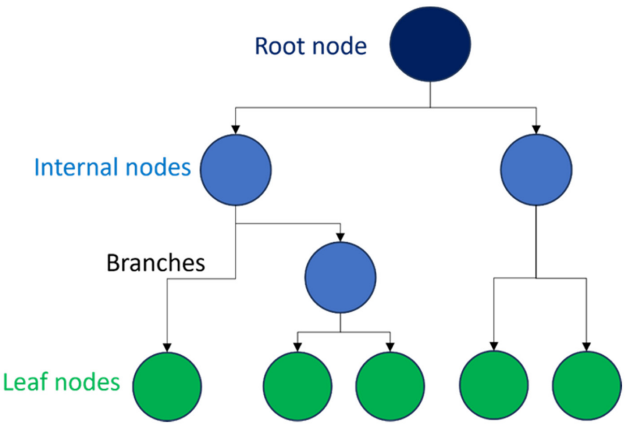
ปัญหาแบบ classification จะเป็นปัญหาที่ต้องการระบุประเภทหรือกลุ่มของข้อมูลว่าจัดอยู่ในหมวดหมู่ใด เช่น การตัดสินว่าฟันผุหรือไม่ผุ การระบุว่าเป็นฟันชนิดใด การประเมินระดับความรุนแรงของโรคปริทันต์ เป็นต้น วิธีการหลักทางสถิติที่ใช้สร้างโมเดล คือ สมการ logistic regression



รูปที่ 4 ขั้นตอนการสร้างและประเมินประสิทธิภาพโมเดลแบบ supervised learning.

อัลกอริทึมการเรียนรู้ของเครื่องแบบ supervised learning

1. **Decision tree** เป็นอัลกอริทึมที่ใช้โครงสร้างแบบ flow chart เหมือนต้นไม้แต่กลับส่วนรากขึ้นบนและส่วนใบอยู่ด้านล่าง มีส่วนประกอบคือ root, internal nodes, branches, และ leaves (รูปที่ 5) การตัดสินใจของอัลกอริทึมนี้จะพิจารณาตามแผนผังโดยเริ่มจากจุดบนสุดคือ root node เมื่อพบข้อมูลว่าเป็นไปตามเงื่อนไขอัลกอริทึมก็จะตัดสินใจไปตามทิศทางของแผนผัง จนกระทั่งถึงส่วนสุดท้ายที่ตัดสินใจคือ leaf nodes

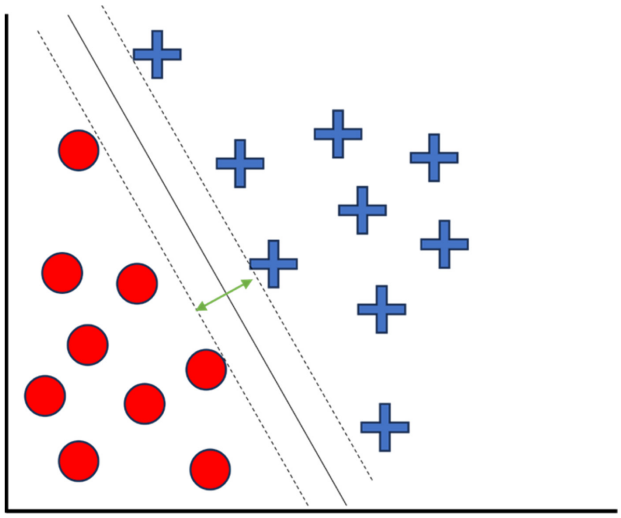


รูปที่ 5 Decision tree algorithm

2. **Random Forest (RF)** เป็นอัลกอริทึมที่สร้างจาก decision tree หลายชุดมารวมกันโดยสร้างแบบสุ่มให้แต่ละชุดมีความแตกต่างเพื่อให้มีความหลากหลายและมีความเป็นอิสระต่อกันซึ่งช่วยลดปัญหาการเกิด overfitting กับข้อมูลชุดที่นำมาฝึก การใช้ decision tree จำนวนมากร่วมกันวิเคราะห์ข้อมูลทำให้ประสิทธิภาพการทำงานและการคาดการณ์มีความแม่นยำมากขึ้น

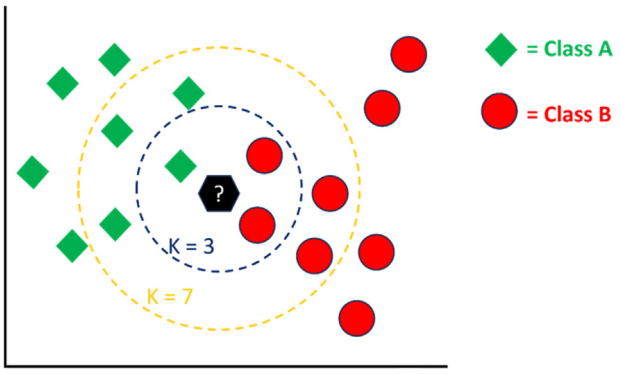
3. **Gradient boosting machine (GBM)** เป็นอัลกอริทึมแบบ ensemble คือโมเดลที่เกิดจากโมเดลมากกว่าหนึ่งโมเดลที่นำมาประกอบกันเป็นโมเดลที่มีความซับซ้อน โดย GBM เป็นโมเดลที่สร้างจาก decision tree จำนวนหลายชุดมาต่อกันโดยการฝึก decision tree แต่ละชุดจะเรียนรู้จากค่าผิดพลาดของ decision tree ชุดก่อนหน้าทำให้การคาดการณ์มีความแม่นยำมากขึ้นตามจำนวนการเรียนรู้ของ decision tree ที่มากขึ้นต่อเนื่องกัน

4. **Support vector machine (SVM)** เป็นอัลกอริทึมที่ตัดสินใจแบ่งกลุ่มข้อมูลของชุดข้อมูลที่ปรากฏบนกราฟ (รูปที่ 6) โดยการสร้างเส้นแบ่งข้อมูลตรงกลางหนึ่งเส้นและเส้นขอบเขตอีกสองเส้นที่แต่ละฝั่งของเส้นกลางจากนั้น SVM จะปรับหาตำแหน่งและมุมของเส้นกลางที่ทำให้ระยะห่างระหว่างเส้นขอบทั้งสองกว้างมากที่สุดเพื่อใช้ในการตัดสินใจแบ่งชุดข้อมูลออกจากกัน



รูปที่ 6 Support vector machine

5. K-nearest neighbor (KNN) เป็นอัลกอริทึมการสร้างโมเดลทำนายที่ตั้งชื่อตามวิธีการทำงานของอัลกอริทึม คือ การจำแนกกลุ่มผลลัพธ์ตามความใกล้เคียงของข้อมูลกับข้อมูลจุดอื่นบนกราฟ ตัวอย่างเช่น ชุดข้อมูลชุดหนึ่งจำแนกออกเป็น 2 กลุ่ม (รูปที่ 7) คือ จุดข้อมูลที่มีป้ายกำกับเป็นสีเขียวและสีแดง เมื่อมีข้อมูลจุดใหม่ (สีดำ) ที่ต้องการจำแนกว่าอยู่กลุ่มใด อัลกอริทึม KNN จะนับข้อมูลจุดที่ใกล้จุดใหม่มากที่สุดเพื่อใช้ในการตัดสินใจ ค่า K หมายถึงจำนวนของจุดข้อมูลใกล้ที่สุดที่ใช้พิจารณา ตัวอย่างเช่น ถ้ากำหนดค่า K เท่ากับ 3 อัลกอริทึมจะใช้จุดข้อมูลตำแหน่งที่อยู่ใกล้กับข้อมูลจุดใหม่มากที่สุด 3 ตำแหน่งและตัดสินใจตามจำนวนจุดข้อมูลที่มีมากกว่า จากรูปจุดที่ใกล้จุดข้อมูลใหม่ 3 จุด ประกอบด้วยจุดเขียว 1 จุดและจุดแดง 2 จุด ดังนั้น KNN จะจำแนกจุดใหม่ให้เป็นส่วนหนึ่งของกลุ่มสีแดง

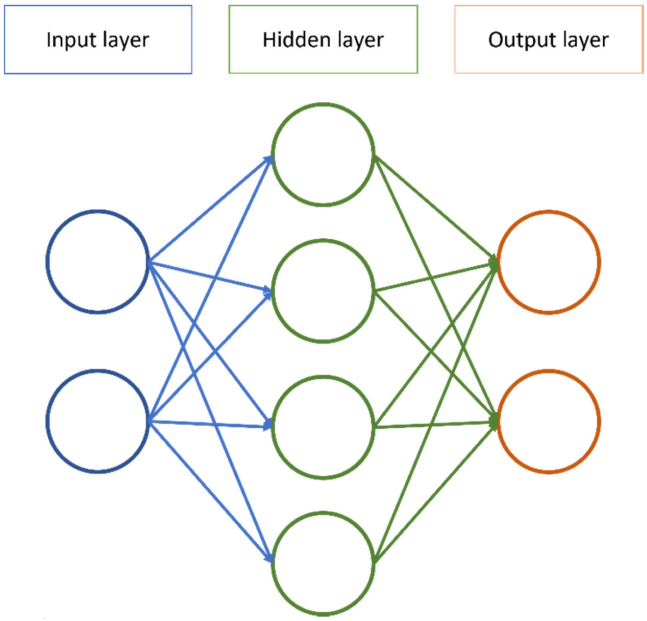


รูปที่ 7 K-nearest neighbor algorithm

6. Naive Bayes เป็นอัลกอริทึมที่ตัดสินใจแบ่งกลุ่มข้อมูลโดยการอิงทฤษฎีความน่าจะเป็นของ Bayes ที่มีแนวคิดว่าความน่าจะเป็นที่เกิดเหตุการณ์หนึ่งก็ต่อเมื่อมีอีกเหตุการหนึ่งได้เกิดไปแล้ว อัลกอริทึมจะอาศัยความน่าจะเป็นของการเกิดแต่ละเหตุการณ์มาใช้ในการตัดสินใจ ส่วนใหญ่ใช้ในการวิเคราะห์ข้อมูลที่เป็นเหตุการณ์ต่อเนื่อง

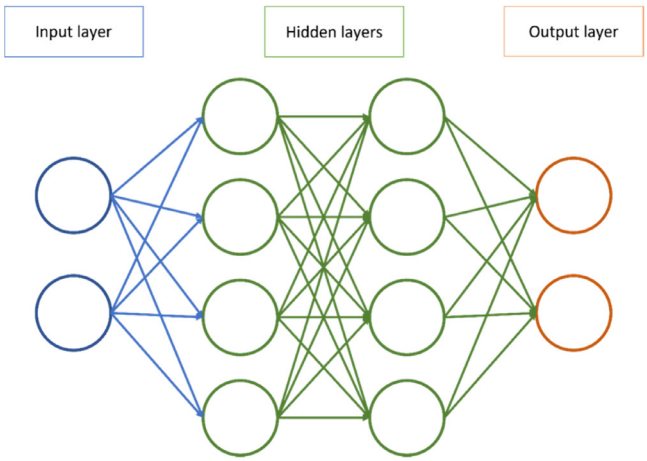
การเรียนรู้เชิงลึก (Deep learning; DL)

DL เป็น machine learning ชนิดหนึ่งที่คอมพิวเตอร์จะนำข้อมูลนำเข้ามาวิเคราะห์ด้วยสมการทางสถิติและสร้างเป็นโมเดล โดยจำนวนสมการจะมากขึ้นตามความซับซ้อนของข้อมูลซึ่งสามารถเขียนออกมาเป็นแผนภาพเรียกว่าเครือข่ายประสาทเทียม (neural network) (รูปที่ 8)



รูปที่ 8 เครือข่ายประสาทเทียม (neural network) ของ machine learning

จากรูปที่ 8 แสดงการทำงานของ machine learning โดย input layer คือชั้นข้อมูลนำเข้าที่ป้อนเข้าสู่คอมพิวเตอร์ hidden layer คือชั้นของสมการทางสถิติที่เกิดการคำนวณและส่งข้อมูลระหว่างกันมีความซับซ้อนและไม่สามารถรู้การทำงานที่แน่ชัดได้เมื่อได้ผลลัพธ์การตัดสินใจจะส่งผลลัพธ์ออกมาปรากฏที่ output layer วงกลมแต่ละวงในชั้น hidden layer เรียกว่า node การส่งต่อข้อมูลที่คำนวณไปมาระหว่าง node อย่างซับซ้อนทำให้ถูกนำไปเปรียบเทียบกับการส่งต่อข้อมูลของระบบประสาท จึงเรียกการทำงานของ machine learning นี้ว่า artificial neural networks (ANN) กรณีข้อมูลมีจำนวนมากและมีความซับซ้อนในการวิเคราะห์การใช้ hidden layer เพียงชั้นเดียวจะทำให้เสียเวลาและทรัพยากรในการคำนวณมากจึงมีการพัฒนาโดยเพิ่มจำนวน hidden layer ขึ้นทำให้การคำนวณมีต้นทุนลดลงจะเรียก ANN ที่มีจำนวน hidden layer มากกว่าหนึ่งชั้นว่า deep learning(12) สามารถเขียนแผนผังได้ดังรูปที่ 9



รูปที่ 9 เครือข่ายประสาทเทียมของการเรียนรู้เชิงลึก (deep neural network) จะมีจำนวน hidden layer มากกว่าหนึ่งชั้น

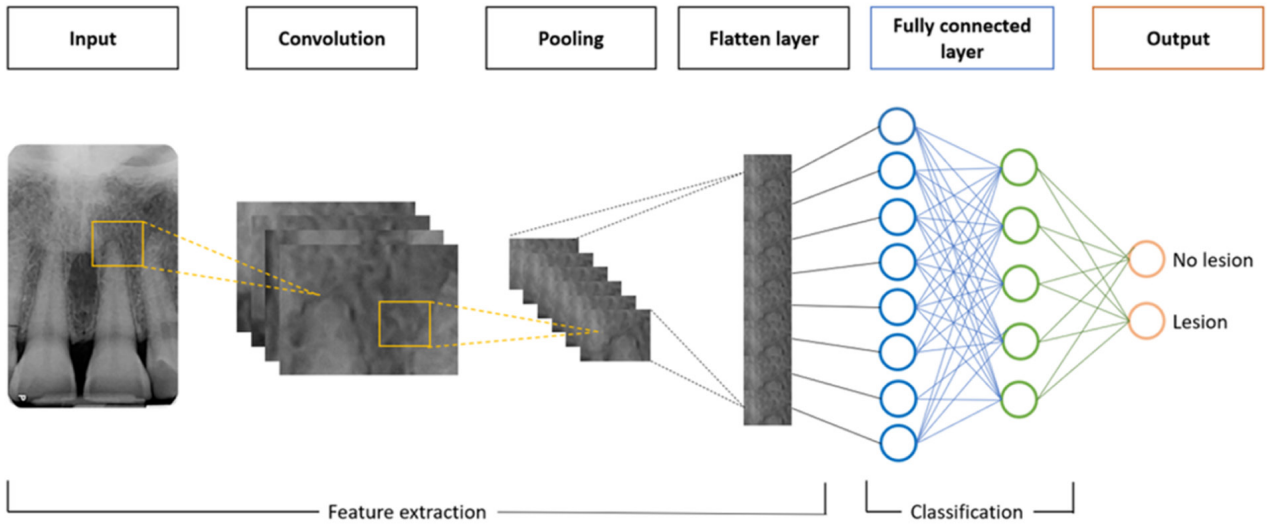
สถาปัตยกรรมการเรียนรู้เชิงลึกที่ใช้ในงานวิทยาเอ็นโดดอนต์

Convolutional neural networks (CNN) เป็น ML ชนิด supervised learning ที่ถูกพัฒนาขึ้นเพื่อแก้ปัญหาเกี่ยวกับรูปภาพ (16) โดยปกติ ANN จะสามารถแก้ปัญหารูปภาพมีไม่มีพื้นหลังได้อย่างมีประสิทธิภาพ เช่น รูปตัวเลขหรือตัวอักษรบนกระดาษสีขาว แต่จะไม่มีประสิทธิภาพเมื่อรูปภาพมีรายละเอียดพื้นหลัง เช่น รูปแผ่นป้ายทะเบียนรถยนต์ที่วิ่งอยู่บนถนน ทำให้ต้องสกัดข้อมูล (feature extraction) สำคัญออกมาจากรูปภาพที่มีพื้นหลังก่อนที่จะนำข้อมูลเข้าสู่ neural networks เรียกขั้นตอนนี้ว่า convolution ดังนั้นความหมายของ CNN คือ neural networks ที่รับข้อมูลนำเข้าที่ผ่านกระบวนการสกัดข้อมูลสำคัญมาแล้ว (รูปที่ 10)

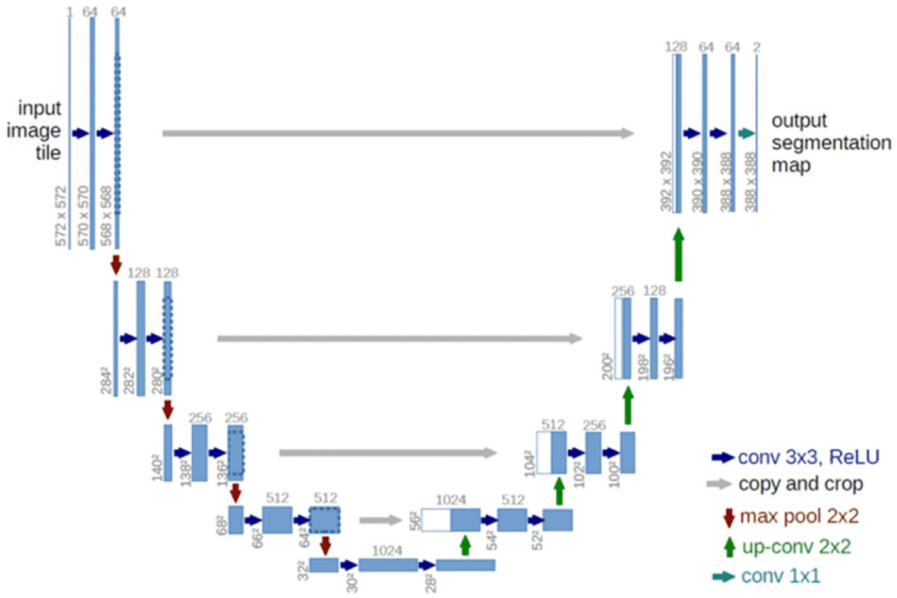
U-net (17) เป็นอัลกอริทึมที่พัฒนามาจาก convolutional neural network แบบดั้งเดิมที่คิดค้นขึ้นในปีค.ศ.2015 นำมาใช้กับรูปทางการแพทย์เพื่อวิเคราะห์โรคหรือระบุความผิดปกติในภาพถ่าย รูปแบบสถาปัตยกรรมจะประกอบด้วย CNN 2 ส่วน คือ รูปภาพแบบย่อและขยายในแต่ละชั้นเมื่อเขียนเป็นแผนภาพมีลักษณะคล้ายรูปตัว U (รูปที่ 11) เป็นอัลกอริทึมที่สามารถระบุพื้นที่และแยกเส้นขอบของรูปภาพได้ดีในระดับพิกเซลจึงเหมาะสำหรับงาน image segmentation หรือ image classification

Probabilistic neural network (PNN) เป็นเครือข่ายประสาทเทียมที่ใช้สำหรับปัญหาการจำแนกประเภทและการจดจำรูปแบบ อัลกอริทึมจะใช้ probability distribution function โดยสร้างจากเครือข่าย Bayesian มีข้อดีคือมีความเร็วและความแม่นยำมากกว่าเครือข่าย perceptron หลายชั้น และไม่อ่อนไหวต่อค่าผิดปกติ

Deep Residual Network (ResNet) (18) เป็นโมเดลที่ใช้สำหรับงานประมวลรูปภาพ ถูกพัฒนาขึ้นเพื่อแก้ปัญหาเรื่องของ vanishing gradients จากโมเดลที่มีเครือข่ายลึกหลายชั้นทำให้ข้อมูลจากชั้นบนจะเลือนหายไป

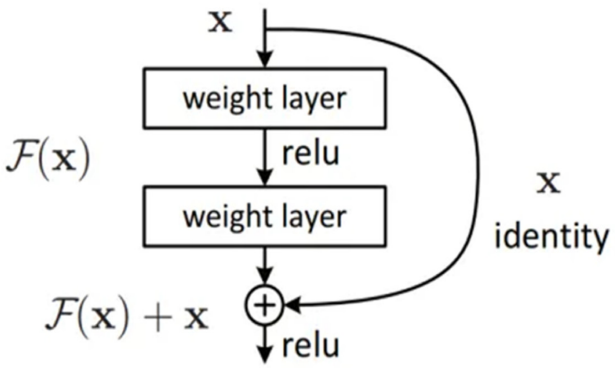


รูปที่ 10 Convolutional Neural Networks



รูปที่ 11 U-net architecture มีลักษณะคล้ายตัว U (17)

เมื่อมีการส่งผ่านไปยังชั้นลึก จึงมีแนวคิดให้ส่งผ่านข้อมูลจากชั้นแรก (ค่า x) ลงไปบวกกับผลลัพธ์ที่ได้จากชั้นลึก (รูปที่ 12) ซึ่งการทำเช่นนี้ทำให้โมเดลที่ได้มีประสิทธิภาพดีขึ้นเมื่อโมเดลมีชั้นประสาทเทียมลึกหลายชั้น



รูปที่ 12 Residual leaning แสดงการส่งผ่านค่า x จากชั้นบนลงไปบวกกับผลลัพธ์ในชั้นลึก (18)

Xception net ถูกนำเสนอโดย Chollet (19) ในปี ค.ศ.2017 เป็นโมเดลแบบ depthwise separable convolutions ที่พัฒนาปรับปรุงจาก CNN แบบดั้งเดิม โดยมีชั้นเครือข่ายประสาทเทียมแบบ convolutional จำนวน 36 ชั้นเพื่อลดการทำงานและประมวลผลแต่ยังคงมีประสิทธิภาพเท่ากับ ResNet เป็นโมเดลที่มีประสิทธิภาพและใช้เวลาในการคำนวณต่ำกว่าโมเดล CNN แบบดั้งเดิม

การประเมินประสิทธิภาพของโมเดล

เมื่อสร้างโมเดลจากชุดฝึกสำเร็จจะประเมินประสิทธิภาพของโมเดลด้วยข้อมูลชุดทดสอบโดยตัวบ่งชี้ประสิทธิภาพของโมเดลจะแบ่งตามประเภทของโมเดลที่สร้างขึ้น

การประเมินประสิทธิภาพของโมเดลแบบ linear regression (20, 21) จะประเมินประสิทธิภาพของโมเดลด้วย R-squared score (R^2), mean squared error (MSE), mean absolute error (MAE), หรือ mean absolute percentage error (MAPE) โดยมีสมการคำนวณดังตารางที่ 1

R-squared score เป็นค่าที่บอกความแตกต่างระหว่างค่าจริงกับค่าพยากรณ์ โดยจะมีค่าน้อยที่สุดเป็นอนันต์และค่ามากที่สุดเท่ากับ 1 ยิ่งค่าเข้าใกล้ 1 มากเท่าใดแสดงว่ามีประสิทธิภาพการพยากรณ์ใกล้เคียงค่าจริงมากเท่านั้น โมเดลที่ดีควรมีค่า R^2 มากกว่า 0 และเข้าใกล้ 1

Mean squared error คือค่าเฉลี่ยความแตกต่างของค่าจริงกับค่าพยากรณ์ยกกำลังสอง เป็นค่าที่บอกช่วงความคลาดเคลื่อนของค่าทำนายจากค่าจริง เช่น โมเดลพยากรณ์อุณหภูมิได้ค่าพยากรณ์เท่ากับ 37 องศาเซลเซียส มีค่า MSE เท่ากับ 9 แสดงว่าค่าจริงมีโอกาสอยู่ในช่วง 37 ± 3 องศาเซลเซียส ($\sqrt{9}$ = 3) สาเหตุที่ต้องถอดรากที่สองเนื่องจากเป็นค่าที่คำนวณได้จากสมการเป็นค่ากำลังสอง

Mean absolute error คือค่าเฉลี่ยความแตกต่างของค่าจริงกับค่าพยากรณ์สัมบูรณ์ ที่บอกช่วงความคลาดเคลื่อนของค่าทำนายจากค่าจริง ลักษณะการใช้งานเหมือนกับค่า MSE

Mean absolute percentage error คือ MAE ที่แปลงให้อยู่ในรูปร้อยละเพื่อความสะดวกในการสื่อสาร เช่น โมเดลที่สร้างขึ้นมีค่า MAPE 2.1% หมายความว่าค่าพยากรณ์ที่ได้มีโอกาสคลาดเคลื่อน ±2.1%

การประเมินประสิทธิภาพของโมเดล(22) แบบ logistic regression จะประเมินประสิทธิภาพของโมเดลด้วย confusion matrix และคำนวณค่าตัวบ่งชี้ประสิทธิภาพประกอบด้วย accuracy, precision, recall, negative predictive value, sensitivity, specificity, F1-score โดยสมการคำนวณดังตารางที่ 2

Accuracy คือ ความแม่นยำของการพยากรณ์บอกร้อยละที่โมเดลพยากรณ์ได้ถูกต้องจากจำนวนครั้งที่พยากรณ์ทั้งหมด การใช้ค่า accuracy ประเมินประสิทธิภาพของโมเดลมีข้อควรระวังกรณีกลุ่มค่าบวกและกลุ่มค่าลบมีจำนวนต่างกันมากหรือเรียกว่า imbalance class เนื่องจากน้ำหนักของความแม่นยำในการพยากรณ์เกือบทั้งหมดมากจากกลุ่มชุดข้อมูลที่มีจำนวนมากกว่าทำให้แม้โมเดลจะพยากรณ์กลุ่มที่มีจำนวนน้อยผิดทั้งหมดค่าความแม่นยำของการพยากรณ์ยังคงมีค่าสูง ซึ่งอาจนำไปสู่การแปลความที่ผิดพลาดได้ ดังนั้นจึงควรพิจารณาค่าบ่งชี้อื่นประกอบการประเมินประสิทธิภาพโมเดลด้วย

Precision หรือ Positive predictive value คือสัดส่วนที่บอกถึงประสิทธิภาพของการพยากรณ์ค่าบวกว่าพยากรณ์ค่าบวกถูกต้องร้อยละเท่าใดจากการพยากรณ์ค่าบวกทั้งหมด ค่า precision ในอุดมคติคือ 1 หมายความว่าโมเดลพยากรณ์ค่าบวกได้ถูกต้องทั้งหมดโดยไม่มีค่าที่เป็นค่าบวกลวง (false positive)

Recall หรือ Sensitivity หรือ True positive rate คือ ความน่าจะเป็นที่โมเดลจะพยากรณ์ค่าบวกได้ถูกต้องจากชุดข้อมูลจริงที่มีค่าบวกทั้งหมด ค่า recall ในอุดมคติคือ 1 หมายความว่าโมเดลพยากรณ์ค่าบวกได้ถูกต้องทั้งหมดโดยไม่มีค่าที่เป็นค่าลบลวง (false negative)

ตารางที่ 1 สมการคำนวณของตัวชี้วัดเพื่อประเมินประสิทธิภาพโมเดลแบบ regression

ตัวชี้วัด	สมการ
R-squared score (R^2)	$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}$
Mean squared error (MSE)	$MSE = \frac{1}{N}\sum_{i=1}^{n}(y_i - \hat{y})^2$
Mean absolute error (MAE)	$MAE = \frac{1}{N}\sum_{i=1}^{n}\lvert y_i - \hat{y}\rvert$
Mean absolute percentage error (MAPE)	$MAPE = \left[\frac{1}{N}\sum_{i=1}^{n}\frac{\lvert y_i - \hat{y}\rvert}{\lvert y_i\rvert}\right] \times 100$

*หมายเหตุ: y_i คือ ค่าที่แท้จริงของข้อมูลแต่ละตัว, $\hat{y}$ คือ ค่าพยากรณ์ของข้อมูลแต่ละตัว, N คือ จำนวนข้อมูลทั้งหมด, n คือจำนวนข้อมูลแต่ละกลุ่มตัวอย่าง

F1 score คือ ค่าที่ใช้ในการประเมินประสิทธิภาพของโมเดลโดยรวมทั้งค่า precision และ recall เข้าไว้ด้วยกัน โดยใช้การคำนวณค่าเฉลี่ยค่าเฉลี่ยฮาร์มอนิกของค่าทั้งสอง โดยค่าอุดมคติของ F1-score คือ 1 ซึ่งหมายความว่าทั้งค่า precision และ recall ต้องมีค่าเท่ากับ 1 ทั้งคู่ การที่ค่า F1-score จะมีค่าสูงได้นั้นค่าที่เป็นองค์ประกอบคือ precision และ recall ต้องมีค่าสูงทั้งคู่ จึงเหมาะสมที่จะใช้ในการประเมินประสิทธิภาพของโมเดลมากกว่าค่า accuracy

Dice Score หรือ Sørensen–Dice coefficient เป็นการคำนวนสถิติที่ใช้วัดค่าความเหมือนกันของข้อมูลตัวอย่างสองชุด คือ ค่าทำนายและค่าจริง ใช้ในงานการเรียนรู้เชิงลึกแบบแยกส่วน (deep learning-based segmentation) โดยเฉพาะในกรณีมีปัญหาชั้นข้อมูลอสมดุลขั้นสูง (High-class Imbalance Problem) มีค่าอยู่ระหว่าง 0 และ 1 โดยค่ายิ่งเข้าใกล้ 1 ยิ่งดี

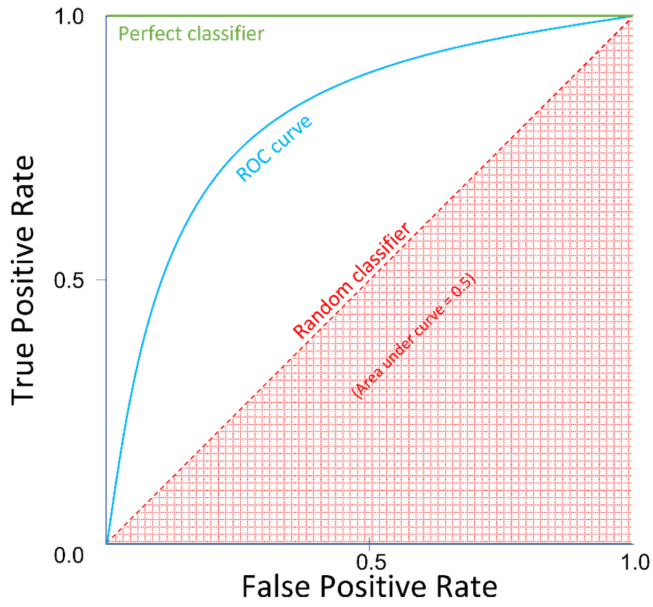
Negative predictive value คือ สัดส่วนที่บอกถึงประสิทธิภาพของการพยากรณ์ค่าลบ บอกว่าพยากรณ์ค่าลบถูกต้องร้อยละเท่าใดจากการพยากรณ์ค่าลบทั้งหมด

Specificity หรือ True negative rate คือ ความน่าจะเป็นที่โมเดลจะพยากรณ์ค่าลบได้ถูกต้องจากชุดข้อมูลจริงที่มีค่าลบทั้งหมด

Area Under the Curve of the Receiver Operating Characteristic Curve (AUC-ROC) เป็นกราฟความน่าจะเป็นที่วาดอยู่บนแกน True positive rate (sensitivity) กับ False positive rate (คือ 1-specificity) (รูปที่ 13) แสดงประสิทธิภาพของ classification model ที่แต่ละ classification threshold โดยพื้นที่ใต้กราฟ (AUC) จะเป็นค่าที่บ่งบอกความสามารถในการจำแนกข้อมูล โมเดลที่มีค่าพื้นที่ใต้กราฟมากแสดงว่ามีประสิทธิภาพสูงในการจำแนกข้อมูล ถ้า AUC = 1 หมายถึงโมเดลสามารถจำแนกข้อมูลแบบ positive และ negative ได้ถูกต้องทั้งหมด (perfect classifier) ถ้า AUC = 0 หมายถึงโมเดลจำแนกข้อมูลผิดทั้งหมด คือ ข้อมูลที่เป็น positive จะระบุว่าเป็น negative ขณะที่ข้อมูล negative จะระบุว่าเป็น positive ถ้า AUC = 0.5 แสดงว่า model มีประสิทธิภาพในการแยกข้อมูลระหว่าง positive กับ negative ได้เท่ากับการสุ่ม (random classifier) ดังนั้นโมเดลที่ใช้งานได้ควรมีค่า AUC ไม่น้อยกว่า 0.5 (23)

ตารางที่ 2 Confusion matrix แสดงสมการคำนวณตัวบ่งชี้ประเมินประสิทธิภาพโมเดลแบบ classification

		Predicted class: Positive	Predicted class: Negative	
Actual class	Positive	True positive (TP)	False negative (FN) -Type II error-	**Sensitivity** $\frac{TP}{TP+FN}$
	Negative	False positive (FP) -Type I error-	True negative (TN)	**Specificity** $\frac{TN}{TN+FP}$
		Precision $\frac{TP}{TP+FP}$	**Negative predictive value** $\frac{TN}{TN+FN}$	**Accuracy** $\frac{TP+TN}{TP+TN+FP+FN}$
		F1 score $2\left[\frac{precision \cdot recall}{precision + recall}\right]$	**Sørensen–Dice coefficient** $\frac{2TP}{2TP+FP+FN}$	



รูปที่ 13 Receiver operating characteristic (ROC) curve

การประยุกต์ใช้การเรียนรู้ของเครื่องสำหรับงานสาขาวิทยาเอ็นโดดอนต์ในปัจจุบัน

การศึกษาวิจัย ML เพื่อใช้ในงานวิทยาเอ็นโดดอนต์ถูกนำไปใช้หลากหลายด้าน ทั้งการพยากรณ์โรค จำแนกลักษณะกายวิภาคของฟันและอวัยวะที่เกี่ยวข้อง แปลผลภาพรังสี ประเมินความยาวทำงาน ประเมินความยากในการรักษา ลักษณะการใช้งานส่วนใหญ่จะเกี่ยวข้องกับรูปภาพ ภาพรังสีทั้งแบบสองมิติและสามมิติ จึงเป็นปัญหาแบบ classification โดยมีการวิจัยในหัวข้อต่าง ๆ ดังนี้ (ตารางที่ 3)

การประเมินระดับความยากของเคสและการตัดสินใจส่งต่อ

การพัฒนาโมเดล ML ใช้สำหรับประเมินระดับความยากของเคสและการตัดสินใจส่งต่อ (24) โดยใช้วิธีการเรียนรู้แบบ deep neural network และ support vector machine ชุดข้อมูลนำเข้าสำหรับการเรียนรู้ของโมเดลเป็นข้อมูลจากแบบประเมินความยากของเคสของ American Association of Endodontists (AAE) จากผู้ป่วยจำนวน 500 รายและถูกเตรียมป้ายกำกับข้อมูลโดยผู้เชี่ยวชาญสาขาวิทยาเอ็นโดดอนต์จำนวน 2 คน ผลลัพธ์การตัดสินใจของโมเดลจะแบ่งความยากออกเป็น 3 ระดับ คือ น้อย ปานกลาง และมาก การประเมินประสิทธิภาพหลังจากโมเดลเรียนรู้สำเร็จพบว่ามีค่า sensitivity 94% และช่วยเพิ่มความรวดเร็วในการตัดสินใจส่งต่อ การใช้ ML ในการประเมินความยากและตัดสินใจส่งต่อจะช่วยทันตแพทย์ทั่วไปหรือทันตแพทย์ที่มีประสบการณ์น้อยสามารถตัดสินใจส่งต่อเคสที่ยากได้เหมือนกับการตัดสินใจของทันตแพทย์ผู้เชี่ยวชาญ ช่วยลดเวลาในการประเมินและเป็นภววิสัย (objective) ทำให้ตัดสินใจมีความเที่ยงตรงและแม่นยำ

การพยากรณ์โรคของการรักษาจุลศัลยกรรมทางวิทยาเอ็นโดดอนต์ (endodontic microsurgery)

โมเดลสำหรับการพยากรณ์โรคของการรักษา endodontic microsurgery (25) ที่สร้างด้วยโมเดลทางคณิตศาสตร์ 2 แบบคือ gradient boosting machine (GBM) และ random forest (RF) โดยใช้ข้อมูลนำเข้าจากแบบประเมินความยากของเคสและภาพรังสี CBCT จำนวนตัวอย่างทั้งหมด 234 ซี่ พบว่าประสิทธิภาพของโมเดลในการทำนายโรคของ GBM มีค่า accuracy = 0.80, sensitivity = 0.92, specificity = 0.71, PPV = 0.71, NPV = 0.92, F1 = 0.80, และ AUC = 0.88 ขณะที่โมเดล RF มีค่า accuracy = 0.80, sensitivity = 0.85, specificity = 0.76, PPV = 0.73, NPV = 0.87, F1 = 0.79, และ AUC = 0.83 โมเดลที่พัฒนาด้วยตัวแปร 8 ตัวประกอบด้วย อายุ เพศ ชนิดฟัน ชนิดการทำลายกระดูก ขนาดรอยโรค ความยาววัสดุอุดคลองราก ความแน่นวัสดุอุดคลองรากและความยาวเดือยฟัน มีความสามารถในการพยากรณ์โรคของการทำ endodontic microsurgery ได้ และโมเดล GMB มีประสิทธิภาพดีกว่า RF เล็กน้อย ทั้งนี้ในการศึกษาดังกล่าวใช้ภาพรังสีติดตามผลการรักษาภายหลังการผ่าตัด 1 ปีในการตัดสินว่าการรักษาประสบความสำเร็จหรือไม่โดยแบ่งเป็นสองกลุ่ม คือ กลุ่ม healed จะไม่ปรากฏเงาโปร่งรังสีรอบปลายราก และกลุ่ม not healed จะปรากฏเงาโปร่งรังสีรอบปลายรากไม่ว่าจะเป็นกรณีการหายไม่สมบูรณ์แม้ว่ารอยโรคเล็กลงหรือไม่มีการหายโดยรอยโรคมีขนาดเท่าเดิมหรือใหญ่ขึ้นซึ่งผู้วิจัยได้อ้างอิงตามการศึกษาของ Rud และคณะ (26) และ Molven และคณะ (27)

การจำแนกลักษณะกายวิภาคระบบคลองราก

มีการพัฒนาโมเดลเพื่อระบุคลองรากรูปตัวซี (C-shaped canal) จากภาพรังสีส่วนตัดอาศัยคอมพิวเตอร์ลำรังสีรูปกรวย (CBCT) บนสถาปัตยกรรม U-net, residual U-net และ Xception U-net(28) จากชุดข้อมูลภาพรังสี CBCT จำนวน 135 ภาพ พบว่าโมเดลที่สร้างขึ้นสามารถช่วยระบุคลองรากรูปตัวซีในภาพรังสี CBCT ได้และประสิทธิภาพของโมเดลที่สร้างจาก U-Net, residual U-Net, และ Xception U-Net มีค่า Dice coefficients เท่ากับ 0.660, 0.736, และ 0.768 ตามลำดับ แต่การศึกษานี้มีข้อจำกัดเนื่องจากจำนวนชุดข้อมูลตัวอย่างน้อยทำให้ประสิทธิภาพของโมเดลยังไม่สูงจึงยังต้องการชุดข้อมูลภาพรังสี CBCT ของคลองรากรูปตัวซีเพิ่มขึ้นและการปรับภาพ เช่น การเพิ่ม contrast อาจทำให้โมเดลสามารถระบุขอบเขตในรูปได้ดีขึ้นอาจทำให้ประสิทธิภาพของโมเดลเพิ่มมากขึ้น นอกจากนี้มีการพัฒนาโมเดลเพื่อระบุรากด้านลิ้นไกลกลางของฟันกรามล่างซี่ที่หนึ่งจากภาพรังสี panoramic (29) จากชุดข้อมูลจำนวน 760 ภาพ ที่สร้างโดยใช้ AlexNet และ GoogleNet พบว่าโมเดลมีประสิทธิภาพสูงในการระบุว่ารากไกลกลางของฟันกรามล่างซี่ที่หนึ่งมีรากเดียวหรือมากกว่าหนึ่งราก โดยประสิทธิภาพโมเดลมีค่า accuracy 87.4% และ 85.3% ตามลำดับ

การระบุรอยโรครอบปลายรากในภาพรังสี

โมเดลส่วนใหญ่ที่ถูกสร้างขึ้นเพื่อใช้งานทางวิทยาเอ็นโดดอนต์เป็นโมเดลเพื่อระบุรอยโรครอบปลายรากในภาพรังสีทั้งภาพรังสีรอบปลายราก ภาพรังสีพาโนรามิก (panoramic) และภาพรังสี CBCT

โมเดลที่สร้างจากชุดข้อมูลนำเข้าที่เป็นภาพรังสีรอบปลายรากโดยใช้ CNN (30) ถูกสร้างจากภาพรังสีรอบปลายรากจำนวน 4,129 ภาพ พบว่าโมเดลสามารถช่วยเพิ่มความแม่นยำและความเที่ยงในการแปลผลรอยโรครอบปลายรากในภาพรังสีโดยมีค่า F1 score = 0.828 ขณะที่อีกการศึกษาหนึ่งใช้ชุดข้อมูลภาพรังสีรอบปลายรากจำนวน 280 ภาพ (31) สร้างบนสถาปัตยกรรม CNN เช่นเดียวกันมีค่า sensitivity, specificity, และ ROC-AUC เท่ากับ 0.79, 0.88, และ 0.86 ตามลำดับ ดังนั้นโมเดลที่พัฒนาขึ้นเพื่อระบุรอยโรครอบปลายรากจากภาพรังสีรอบปลายรากมีประสิทธิภาพ และช่วยเพิ่มความแม่นยำในการทำงาน

โมเดลที่สร้างจากชุดข้อมูลนำเข้าที่เป็นภาพรังสี panoramic มีการใช้สถาปัตยกรรมแบบ CNN (32) ในการสร้างโมเดลจากชุดตัวอย่างจำนวน 2,001 ภาพ ประสิทธิภาพโมเดลพบว่ามีค่า AUC 0.85, sensitivity 0.65, specificity 0.87 ดังนั้น CNN มีประสิทธิภาพในการระบุรอยโรครอบปลายรากบนภาพรังสี panoramic ขณะที่โมเดลซึ่งพัฒนาโดยใช้ deep learning algorithm (33) จากภาพรังสีจำนวน 3,240 ภาพ โมเดลมีประสิทธิภาพโดยค่า F1 score = 0.58, PPV = 0.67, และ Recall = 0.51 นอกจากนี้มีการพัฒนาโมเดลโดยใช้ ML ร่วมกับวิธีการประมวลผลภาพแบบ wavelet transformation (34) ประสิทธิภาพของโมเดลมีค่า sensitivity 70% และ specificity 89%

โมเดลที่สร้างจากชุดข้อมูลนำเข้าที่เป็นภาพรังสี CBCT โดยใช้ CNN บนสถาปัตยกรรม U-net ใช้ชุดข้อมูลนำเข้าจำนวน 153 ภาพ (35) ประสิทธิภาพของโมเดลพบว่ามีค่า recall, precision, and F-measure เท่ากับ 0.89, 0.95 และ 0.93 ตามลำดับ ขณะที่การศึกษาที่ใช้ deep learning algorithm ที่มีชุดข้อมูลจำนวน 20 ภาพ (36) พบว่าโมเดลที่สร้างขึ้นมีค่า accuracy 0.93, specificity 0.88, PPV 0.87, NPV 0.93, และ Dice index 0.67 จะเห็นได้ว่าโมเดลที่ถูกสร้างขึ้นโดยมีชุดข้อมูลเพื่อใช้ฝึกจำนวนมากกว่าทำให้โมเดลมีประสิทธิภาพสูงกว่า นอกจากนี้มีความพยายามในการปรับปรุงประสิทธิภาพของโมเดลโดยการใช้ ML ร่วมกับ graph-based random walks segmentation (37) ที่มีพื้นฐานมาจาก semiautomatic multilabel random walks algorithm เพื่อปรับปรุงประสิทธิภาพในการสกัดข้อมูลจากจุดภาพ (pixel) ของภาพรังสีโดยวิธีการทางคณิตศาสตร์ พบว่าโมเดลมีประสิทธิภาพโดยมีค่า accuracy 94.1%

โมเดลที่พัฒนาขึ้นเพื่อระบุรอยโรครอบปลายรากในภาพรังสีทั้ง 3 แบบมีประสิทธิภาพสูง ถ้าจำนวนชุดข้อมูลนำเข้าที่ใช้สร้างโมเดลมีจำนวนมากขึ้นจะทำให้โมเดล

มีประสิทธิภาพเพิ่มขึ้น นอกจากนี้การเพิ่มประสิทธิภาพโดยใช้กระบวนการทางคณิตศาสตร์ เช่น wavelet transformation หรือ semiautomatic multilabel random walks algorithm ทำให้โมเดลมีประสิทธิภาพเพิ่มขึ้น ชุดข้อมูลภาพรังสี CBCT มีประสิทธิภาพสูงในการใช้ระบุรอยโรครอบปลายรากเนื่องจากเป็นภาพสามมิติไม่มีการซ้อนทับกับอวัยวะข้างเคียง ขณะที่ภาพรังสีแบบ panoramic คุณภาพจะขึ้นกับมุมในการถ่าย มีการบิดเบี้ยว หดขยายของภาพ มีการซ้อนทับกันของอวัยวะรอบข้าง ภาพบางตำแหน่งไม่ชัดเจนทำให้ประสิทธิภาพของโมเดลอาจด้อยกว่าภาพรังสี CBCT และภาพรังสีรอบปลายราก

การระบุรอยแตกของรากในภาพรังสี

การระบุรอยร้าวหรือรอยแตกแนวดิ่งของราก (vertical root fracture) ในภาพรังสีรอบปลายราก มีการศึกษาโมเดลที่พัฒนาโดยใช้สถาปัตยกรรม probabilistic neural network (PNN) (38) ใช้ภาพตัวอย่าง 200 ภาพในการฝึก แบ่งเป็นฟันปกติ 50 ภาพและฟันที่มีรอยแตก 150 ภาพ ประเมินประสิทธิภาพของโมเดลมีค่า sensitivity 98%, specificity 90.5%, และ accuracy 95.7% และโมเดลที่ถูกสร้างขึ้นเพื่อระบุรอยแตกแนวดิ่งของรากในภาพรังสี panoramic จากชุดข้อมูลจำนวน 300 ภาพ (39) พบว่าประสิทธิภาพของโมเดลมีค่า recall 0.75 และ F measure 0.83 มีความพยายามในการสร้างโมเดลจากชุดข้อมูลภาพรังสีส่วนตัดอาศัยคอมพิวเตอร์ขนาดเล็ก (micro CT) จำนวน 36 ภาพ (40) โดยแบ่งเป็นฟันที่มีรอยร้าว 22 ซี่และฟันปกติ 14 ซี่ บนสถาปัตยกรรม U-net ร่วมกับการเตรียมข้อมูลด้วย 3D Isotropic Wavelets โดยให้โมเดลทำนายฟันที่มีรอยร้าวเป็นเชิงปริมาณโดยใช้สัดส่วนของ Voxel ต่อปริมาตรฟันทั้งหมด พบว่าค่าสัดส่วนที่สูงจะเป็นฟันที่มีรอยร้าวซึ่งโมเดลดังกล่าวมีประสิทธิภาพสูงในการระบุฟันที่มีรอยร้าวได้ ขณะที่การใช้ high-resolution CBCT (hr-CBCT) ร่วมกับการตั้งค่าแบบ endodontic mode จะมีการปรับภาพโดยใช้ smoothing filters จะทำให้ประสิทธิภาพในการระบุรอยร้าวลดลง นอกจากนี้ยังมีการศึกษาเปรียบเทียบโมเดลที่สร้างจากชุดข้อมูลภาพรังสีรอบปลายรากเปรียบเทียบกับโมเดลที่สร้างจากชุดข้อมูล CBCT โดยใช้ PNN (41) พบว่าโมเดลที่สร้างโดยใช้ภาพรังสี CBCT ในการฝึกมีประสิทธิภาพสูงกว่าโดยมีค่า accuracy, sensitivity, และ specificity เท่ากับ 96.6, 93.3, และ 100% ตามลำดับ ขณะที่โมเดลที่ใช้ภาพรังสีรอบปลายรากในการฝึกจะมีค่าเท่ากับ 70.00, 97.78, และ 67.7% ตามลำดับ

ชุดข้อมูลภาพรังสี CBCT ช่วยให้โมเดลระบุรอยแตกแนวดิ่งของรากได้มีประสิทธิภาพมากกว่าภาพรังสีรอบปลายรากที่มีข้อจำกัดจากเทคนิคการถ่ายที่เป็นสองมิติทำให้ภาพของอวัยวะทับซ้อนกัน ภาพรังสีที่ได้ขึ้นกับมุมของลำรังสีขณะถ่ายและความกว้างของรอยแตก ขณะที่ภาพรังสี CBCT เป็นสามมิติสามารถระบุรอยแตกได้ชัดเจนกว่า นอกจากนี้ความกว้างของรอยแตกมีผลต่อประสิทธิภาพของโมเดลคือเมื่อรอยแตกกว้างมากขึ้นประสิทธิภาพของโมเดลจะเพิ่มขึ้น มีข้อสังเกตว่าการศึกษาส่วนใหญ่ในการวิจัยในห้องปฏิบัติการมีการควบคุมชุดข้อมูลภาพรังสี เช่น เป็นฟันรากเดียว ไม่มีวัสดุอุดคลองราก ไม่มีเดือยฟัน และไม่มีอวัยวะรอบรากฟันจึงทำให้ผลการศึกษาอาจไม่ตรงกับลักษณะการใช้งานจริงทางคลินิก

การระบุตำแหน่งรอยคอดรูเปิดปลายราก (minor apical constriction)

การระบุตำแหน่งรอยคอดรูเปิดปลายรากโดยใช้โมเดลที่สร้างบน ANN ฝึกด้วยชุดข้อมูลภาพรังสีรอบปลายรากจำนวน 50 ภาพร่วมการเตรียมข้อมูลโดยวิธี Otsu method และ wave let theory (42, 43) พบว่าโมเดลที่สร้างขึ้นสามารถระบุตำแหน่ง minor constriction ได้ถูกต้องถึง 96% เมื่อเปรียบเทียบกับผู้เชี่ยวชาญสาขาวิทยาเอ็นโดดอนต์ที่สามารถระบุได้ถูกต้องเพียง 76% และมีการใช้โมเดล ML เพื่อพัฒนาการระบุตำแหน่งปลายรากของเครื่องระบุปลายรากอิเลกทรอนิกส์ (electronic apex locator) โดยใช้วิธี multifrequency impedance ร่วมกับ DL (44) พบว่ามี accuracy 95% มากกว่าวิธี dual-frequency impedance ratio ที่มีความแม่นยำประมาณ 85%

ตารางที่ 3 การศึกษาวิจัยการเรียนรู้ของเครื่องเพื่อใช้ในงานวิทยาเอ็นโดดอนต์

Author, year	Sample size	Features	Types of DL	Results
Referral decisions				
Mallishery *et al*, 2020 (24)	500	AAE Case Difficulty Assessment Form	SVM and DNN	SVM: Accuracy = 94.80% Sensitivity = 94.96% Specificity = 94.63% Precision = 94.96% DNN: Accuracy = 93.40% Sensitivity = 93.02% Specificity = 93.80% Precision = 94.12%
Prediction				
Qu *et al*, 2022 (25)	341	CBCT	Gradient boosting machine (GBM) and random forest (RF)	GBM: Accuracy = 0.80 Sensitivity = 0.92 Specificity = 0.71 PPV = 0.71 NPV = 0.92 F1 = 0.80 AUC = 0.88 RF: Accuracy = 0.80 Sensitivity = 0.85 Specificity = 0.76 PPV = 0.73 NPV = 0.87 F1= 0.79 AUC = 0.83
Gao *et al*, 2021 (45)	300	13 parameters	Back propagation (BP) artificial neural network	Accuracy = 95.60%
Root canal morphology identification				
Sherwood *et al*, 2021 (28)	135	CBCT	U-net, residual U-net, and Xception U-net	Dice coefficients U-Net = 0.660 Residual U-Net = 0.736 Xception U-Net = 0.768

ตารางที่ 3 การศึกษาวิจัยการเรียนรู้ของเครื่องเพื่อใช้ในงานวิทยาเอ็นโดดอนต์ (ต่อ)

Author, year	Sample size	Features	Types of DL	Results
Hiraiwa *et al*, 2019 (29)	760	Panoramic radiographs	AlexNet and GoogleNet	AlexNet: Accuracy = 87.4%, Sensitivity = 77.3% Specificity = 97.1% PPV = 96.3% NPV = 81.8% GoogleNet: Accuracy = 85.3% Sensitivity = 74.2% Specificity = 95.9% PPV = 94.7% NPV = 80.0%
Periapical lesion detection				
Li *et al*, 2022 (30)	4129	Periapical radiographs	Modified ResNet-18 backbone	F1-score = 0.828
Pauwels *et al*, 2021 (31)	5600	Periapical radiographs	CNN	Sensitivity = 0.79 Specificity = 0.88 ROC-AUC = 0.86
Ekert *et al*, 2019 (32)	2001	panoramic radiographs	CNN	AUC = 0.85 Sensitivity = 0.65 Specificity = 0.87 PPV = 0.49 NPV = 0.93
Endres *et al*, 2020 (33)	3240	panoramic radiographs	CNN	precision = 0.60 F1 score = 0.58 PPV = 0.67
Birdal *et al*, 2016 (34)	9	Panoramic radiographs	Wavelet transformation and polynomial regression	Specificity = 89% Sensitivity = 70%
Orhan *et al*, 2020 (35)	3900	CBCT	U-Net	Reliability = 92.8%

ตารางที่ 3 การศึกษาวิจัยการเรียนรู้ของเครื่องเพื่อใช้ในงานวิทยาเอ็นโดดอนต์ (ต่อ)

Author, year	Sample size	Features	Types of DL	Results
Setzer *et al*, 2020 (36)	61	CBCT	U-Net	Accuracy = 0.93 Specificity = 0.88 PPV = 0.87 NPV = 0.93 DICE index = 0.67
Okada *et al*, 2015 (37)	28	CBCT	Linear discriminant analysis and AdaBoost	Accuracy = 94.1%
Root fracture detection				
Kositbowornchai *et al*, 2013 (38)	200	digital radiography	PNN	sensitivity = 98% Specificity = 90.5% Accuracy = 95.7%
Fukuda *et al*, 2020 (39)	330	panoramic radiographs	DetectNet	Recall = 0.75 Precision = 0.93 F measure = 0.83
Vicory *et al*, 2021 (40)	22	high resolution CBCT and microCT	U-Net	Probability maps give an opportunity to the physician to evaluate the severity of the crack
Johari *et al*, 2017 (41)	240	Periapical and CBCT radiographs	PNN	Periapical: Accuracy = 70.00% Sensitivity = 97.78% Specificity = 67.70% CBCT: Accuracy = 96.6% Sensitivity = 93.3% Specificity = 100%
Working length determination				
Saghiri *et al*, 2012 (42)	36	Periapical radiographs	ANN	Accuracy = 93%
Saghiri *et al*, 2012 (43)	50	Periapical radiographs	ANN	Accuracy = 96%
Qiao *et al*, 2020 (44)	21	Multifrequency Impedance Measurement	ANN	Accuracy = 95%

บทวิจารณ์

การนำปัญญาประดิษฐ์มาประยุกต์ใช้ในงานวิทยาเอ็นโดดอนต์ การศึกษาส่วนใหญ่จะใช้ในการจำแนกลักษณะในภาพรังสี เช่น การระบุรอยโรครอบปลายราก การตรวจสอบรอยแตกของราก การจำแนกรูปร่างของคลองรากรูปตัวซี หรือการระบุรากพิเศษ ขณะที่การประยุกต์ในในการพยากรณ์โรค การตัดสินใจส่งต่อ ยังมีการศึกษาจำนวนน้อย ทั้งนี้อาจเป็นเพราะการสร้างโมเดลต้องใช้ข้อมูลนำเข้าเป็นจำนวนมาก ภาพรังสีเป็นข้อมูลสำคัญในงานวิทยาเอ็นโดดอนต์ที่มีการเก็บบันทึกในการรักษาผู้ป่วยทุกรายจึงเป็นข้อมูลที่มีจำนวนมากเพียงพอและพร้อมที่จะนำมาใช้ในการสร้างโมเดล ขณะที่ข้อมูลชนิดอื่น เช่น แบบประเมินความยาก ข้อมูลพื้นฐานของผู้ป่วย ต้องเก็บข้อมูลเพิ่มเติมจำนวนมากเพื่อให้เพียงพอสำหรับการสร้างโมเดลที่มีประสิทธิภาพจึงเป็นอุปสรรคต่อการศึกษา ดังนั้นในการศึกษาวิจัยเกี่ยวกับ ML ควรมีการวางแผนการเก็บข้อมูลอย่างเป็นระบบ การร่วมมือกันเก็บข้อมูลของศูนย์วิจัยหลายศูนย์ ใช้ระบบเครือข่ายคอมพิวเตอร์ออนไลน์เพื่อช่วยในการเก็บข้อมูลให้ได้จำนวนมากพอซึ่งเป็นปัจจัยหนึ่งที่จะนำไปสู่การสร้างโมเดลที่มีประสิทธิภาพ

โมเดลที่สร้างขึ้นเพื่อใช้ตัดสินใจระบุรอยโรคบนภาพรังสีหรือจำแนกรูปร่างคลองราก มีประสิทธิภาพความแม่นยำสูงแม้จะใช้สถาปัตยกรรมของ ML ต่างชนิดกันและจำนวนข้อมูลนำเข้าที่แตกต่างกัน เมื่อเปรียบเทียบชนิดของภาพรังสีจะพบว่าส่งผลต่อประสิทธิภาพของโมเดลโดยภาพรังสีแบบ CBCT จะให้โมเดลที่มีประสิทธิภาพมากกว่าภาพรังสีพานอรามิกและภาพรังสีรอบปลายราก (30, 32, 33, 35-37, 41)

การศึกษาเกี่ยวกับการใช้ ML ช่วยตัดสินใจส่งต่อกรณีการรักษาที่ยากพบเพียงหนึ่งการศึกษา โมเดลถูกสร้างโดยใช้ข้อมูลจากแบบประเมินความยากของ AAE พบว่ามีประโยชน์หลายประการคือ ช่วยให้ทันตแพทย์ทั่วไปหรือที่มีประสบการณ์น้อยสามารถประเมินระดับความยากได้เหมือนความเห็นของทันตแพทย์ผู้เชี่ยวชาญทำให้มีความสม่ำเสมอของการประเมินระดับความยากของเคส การใช้งานโมเดล ML เป็นการประมวลผลแบบอัตโนมัติทำให้ง่ายในการตัดสินใจในการประเมิน สามารถประเมินได้แม้มีข้อมูลในการกรอกไม่ครบถ้วน และโมเดล ML สามารถฝึกให้เรียนรู้ปรับเปลี่ยนไปตามมาตรฐานระดับความชำนาญของทันตแพทย์ทั่วไปที่แตกต่างกันในแต่ละประเทศได้ (24)

จากการทบทวนงานวิจัยพบว่าโมเดลที่สร้างจาก ML สำหรับการตัดสินใจงานต่าง ๆ ทางวิทยาเอ็นโดดอนต์มีประสิทธิภาพสูง สามารถช่วยลดความลำเอียงและระยะเวลาในการตัดสินใจ เนื่องจาก ML สามารถเรียนรู้ข้อมูลในภาพรังสีได้มากกว่าการรับรู้ของมนุษย์ เช่น ความเข้ม รูปร่าง ความสว่าง รายละเอียดเล็กน้อยที่ปรากฏในภาพรังสี (46, 47)

โมเดล ML ที่สร้างขึ้นแตกต่างกันในแต่ละการศึกษาทำให้ค่าประสิทธิภาพโมเดลมีความแตกต่างกัน ดังนั้นควรมีการศึกษาเพิ่มเติมเพื่อตรวจสอบความตรงและความเที่ยงของแต่ละสถาปัตยกรรมที่ใช้ในการสร้างโมเดลด้วยจำนวนข้อมูลตัวอย่างที่มากเพียงพอก่อนที่จะนำโมเดลที่สร้างได้ไปใช้งานจริง (48)

บทสรุป

ปัจจุบันเริ่มมีการศึกษาวิจัยเพื่อประยุกต์ใช้ ML กับงานทางทันตกรรมมากขึ้น สำหรับสาขาวิทยาเอ็นโดดอนต์งานที่พบว่ามีการศึกษานำมาใช้จำนวนมากจะเกี่ยวข้องกับการตรวจสอบรอยโรครอบปลายรากจากภาพรังสีทั้งแบบสองมิติและสามมิติ งานอื่นที่มีการนำมาประยุกต์ใช้ เช่น การระบุรอยแตกของราก การประเมินความยาวทำงาน การจำแนกประเภทของคลองราก การประเมินความยากในการทำงาน ซึ่งล้วนแล้วแต่เป็นการประเมินโดยใช้ภาพรังสี ดังนั้น ML ที่ใช้ส่วนใหญ่จึงเป็น CNN เป็นพื้นฐานซึ่งเป็น DL มีความสามารถในการสกัดข้อมูลสำคัญจากรูปภาพก่อนที่จะส่งเข้าสู่เครื่องข่ายประสาทเทียมเพื่อสร้างเป็นโมเดลช่วยตัดสินใจร่วมกับการเตรียมข้อมูลด้วยวิธีการทางคณิตศาสตร์เพื่อสกัดข้อมูลสำคัญจากรูปภาพก่อนส่งเข้าเครือข่ายประสาทเทียม จากงานวิจัยเกือบทั้งหมดพบว่า ML ช่วยเพิ่มประสิทธิภาพ ความแม่นยำในการตัดสินใจอยู่ในระดับที่น่าพอใจ รวมทั้งมีความรวดเร็วเหมาะที่จะพัฒนาเพื่อใช้ช่วยประกอบการทำงานทางด้านวิทยาเอ็นโดดอนต์ต่อไปในอนาคต

กิตติกรรมประกาศ

ไม่มี

แหล่งทุนสนับสนุน

ไม่มี

ผลประโยชน์ขัดกัน

ไม่มี

เอกสารอ้างอิง

1. Park WJ, Park JB. History and application of artificial neural networks in dentistry. Eur J Dent. 2018;12(4):594-601.
2. Fatima A, Shafi I, Afzal H, Díez IT, Lourdes DRM, Breñosa J, et al. Advancements in dentistry with artificial intelligence: current clinical applications and future perspectives. Healthcare (Basel). 2022;10(11).
3. Papantonopoulos G, Takahashi K, Bountis T, Loos BG. Artificial neural networks for the diagnosis of aggressive periodontitis trained by immunologic parameters. PLoS One. 2014;9(3):e89757.
4. Geetha V, Aprameya KS, Hinduja DM. Dental caries diagnosis in digital radiographs using back-propagation neural network. Health Inf Sci Syst. 2020;8(1):8.
5. Shankarapillai RM, L.K.; Nair, M.A.; George, R. Periodontitis risk assessment using two artificial neural network algorithms—A comparative study. Int J Dental Clin. 2012;4:17-21.
6. Cui Q, Chen Q, Liu P, Liu D, Wen Z. Clinical decision support model for tooth extraction therapy derived from electronic dental records. J Prosthet Dent. 2021;126(1):83-90.
7. Choi HI, Jung SK, Baek SH, Lim WH, Ahn SJ, Yang IH, et al. Artificial intelligent model with neural network machine learning for the diagnosis of orthognathic surgery. J Craniofac Surg. 2019;30(7):1986-9.
8. Javed S, Zakirulla M, Baig RU, Asif SM, Meer AB. Development of artificial neural network model for prediction of post-streptococcus mutans in dental caries. Comput Methods Programs Biomed. 2020;186:105198.
9. Turing AM. Computing machinery and intelligence. Mind. 1950;59(236):433-60.
10. Oppy GaD, David. The Turing Test. In: Zalta EN, editor. Winter 2021 ed: Metaphysics Research Lab, Stanford University; 2021.
11. Bostrom N. Superintelligence: Paths, dangers, strategies. New York, NY, US: Oxford University Press; 2014. xvi, 328-xvi, p.
12. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to machine learning, neural networks, and deep learning. Transl Vis Sci Technol. 2020;9(2):14.
13. Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial intelligence and machine learning in pathology: The present landscape of supervised methods. Acad Pathol. 2019;6:2374289519873088.
14. Ren R, Luo H, Su C, Yao Y, Liao W. Machine learning in dental, oral and craniofacial imaging: a review of recent progress. PeerJ. 2021;9:e11451.
15. Moore MM, Slonimsky E, Long AD, Sze RW, Iyer RS. Machine learning concepts, concerns and opportunities for a pediatric radiologist. Pediatr Radiol. 2019;49(4):509-16.
16. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: a review. J Med Syst. 2018;42(11):226.

17. Ronneberger O, Fischer P, Brox T, editors. U-Net: Convolutional networks for biomedical image segmentation. medical image computing and computer-assisted intervention – MICCAI 2015; Cham: Springer International Publishing.
18. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. 2016 IEEE conference on computer vision and pattern recognition (CVPR); 2016 27-30 June 2016.
19. Chollet F, editor Xception: Deep learning with depthwise separable convolutions. 2017 IEEE conference on computer vision and pattern recognition (CVPR); 2017 21-26 July 2017.
20. Schneider A, Hommel G, Blettner M. Linear regression analysis: part 14 of a series on evaluation of scientific publications. Dtsch Arztebl Int. 2010;107(44):776-82.
21. Gareth James DWTHRT. An introduction to statistical learning : with applications in R: New York : Springer, [2013] ©2013; 2013.
22. Wu H, Jun Meng F. Review on evaluation criteria of machine learning based on big data. J Phys: Conference Series. 2020;1486(5):052026.
23. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition. 1997;30(7):1145-59.
24. Mallishery S, Chhatpar P, Banga KS, Shah T, Gupta P. The precision of case difficulty and referral decisions: an innovative automated approach. Clin Oral Investig. 2020;24(6):1909-15.
25. Qu Y, Lin Z, Yang Z, Lin H, Huang X, Gu L. Machine learning models for prognosis prediction in endodontic microsurgery. J Dent. 2022;118:103947.
26. Rud J, Andreasen JO, Jensen JE. Radiographic criteria for the assessment of healing after endodontic surgery. Int J Oral Surg. 1972;1(4):195-214.
27. Molven O, Halse A, Grung B. Observer strategy and the radiographic classification of healing after endodontic surgery. Int J Oral Maxillofac Surg. 1987;16(4):432-9.
28. Sherwood AA, Sherwood AI, Setzer FC, K SD, Shamili JV, John C, et al. A deep learning approach to segment and classify C-Shaped canal morphologies in mandibular second molars using cone-beam computed tomography. J Endod. 2021;47(12):1907-16.
29. Hiraiwa T, Ariji Y, Fukuda M, Kise Y, Nakata K, Katsumata A, et al. A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. Dentomaxillofac Radiol. 2019;48(3):20180218.
30. Li S, Liu J, Zhou Z, Zhou Z, Wu X, Li Y, et al. Artificial intelligence for caries and periapical periodontitis detection. J Dent. 2022;122:104107.
31. Pauwels R, Brasil DM, Yamasaki MC, Jacobs R, Bosmans H, Freitas DQ, et al. Artificial intelligence for detection of periapical lesions on intraoral radiographs: comparison between convolutional neural networks and human observers. Oral Surg Oral Med Oral Pathol Oral Radiol. 2021;131(5):610-6.
32. Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, et al. Deep learning for the radiographic detection of apical lesions. J Endod. 2019;45(7):917-22.e5.
33. Endres MG, Hillen F, Salloumis M, Sedaghat AR, Niehues SM, Quatela O, et al. Development of a deep learning algorithm for periapical disease detection in dental radiographs. Diagnostics (Basel). 2020;10(6).

34. Birdal RG, Gumus E, Sertbas A, Birdal IS. Automated lesion detection in panoramic dental radiographs. Oral Radiol. 2016;32(2):111-8.
35. Orhan K, Bayrakdar IS, Ezhov M, Kravtsov A, Özyürek T. Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans. Int Endod J. 2020;53(5):680-9.
36. Setzer FC, Shi KJ, Zhang Z, Yan H, Yoon H, Mupparapu M, et al. Artificial Intelligence for the computer-aided detection of periapical lesions in cone-beam computed tomographic images. J Endod. 2020;46(7):987-93.
37. Okada K, Rysavy S, Flores A, Linguraru MG. Noninvasive differential diagnosis of dental periapical lesions in cone-beam CT scans. Med Phys. 2015;42(4):1653-65.
38. Kositbowornchai S, Plermkamon S, Tangkosol T. Performance of an artificial neural network for vertical root fracture detection: an ex vivo study. Dent Traumatol. 2013;29(2):151-5.
39. Fukuda M, Inamoto K, Shibata N, Ariji Y, Yanashita Y, Kutsuna S, et al. Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography. Oral Radiol. 2020;36(4):337-43.
40. Vicory J, Chandradevan R, Hernandez-Cerdan P, Huang WA, Fox D, Qdais LA, et al. Dental microfracture detection using wavelet features and machine learning. Proc SPIE Int Soc Opt Eng. 2021;11596.
41. Johari M, Esmaeili F, Andalib A, Garjani S, Saberkari H. Detection of vertical root fractures in intact and endodontically treated premolar teeth by designing a probabilistic neural network: an ex vivo study. Dentomaxillofac Radiol. 2017;46(2):20160107.
42. Saghiri MA, Asgar K, Boukani KK, Lotfi M, Aghili H, Delvarani A, et al. A new approach for locating the minor apical foramen using an artificial neural network. Int Endod J. 2012;45(3):257-65.
43. Saghiri MA, Garcia-Godoy F, Gutmann JL, Lotfi M, Asgar K. The reliability of artificial neural network in locating minor apical foramen: a cadaver study. J Endod. 2012;38(8):1130-4.
44. Qiao X, Zhang Z, Chen X. Multifrequency impedance method based on neural network for root canal length measurement. Appl Sci. 2020;10(21):7430.
45. Gao X, Xin X, Li Z, Zhang W. Predicting postoperative pain following root canal treatment by using artificial neural network evaluation. Sci Rep. 2021;11(1):17243.
46. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng. 2017;19:221-48.
47. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500-10.
48. Asiri AF, Altuwalah AS. The role of neural artificial intelligence for diagnosis and treatment planning in endodontics: a qualitative review. Saudi Dent J. 2022;34(4):270-81.