

บทที่ 3

วิธีการดำเนินการวิจัย

นำเสนอรายละเอียดระเบียบวิธีวิจัยและขั้นตอนการทดลองของงานวิจัยอย่างเป็นระบบ อธิบายกระบวนการออกแบบและการพัฒนาแบบจำลองที่นำเสนอ รวมถึงวิธีการประเมินประสิทธิภาพของแบบจำลองในการตรวจจับโรคปอดโดยใช้สัญญาณเสียงจากปอดผ่านโครงข่ายประสาทเทียมลึกแบบหลายงาน

3.1 ชุดข้อมูลและการเตรียมสัญญาณเสียงปอด

3.1.1 ชุดข้อมูลเสียงปอด

งานวิจัยนี้ใช้ชุดข้อมูลเสียงปอดจากฐานข้อมูลที่เผยแพร่สาธารณะ (Fraivan, M., Fraivan, L., Khassawneh, B., and Ibnian, A. A., 2021) ฐานข้อมูลที่ใช้สำหรับงานจำแนกเสียงปอด ซึ่งเป็นชุดข้อมูลที่มีการเก็บข้อมูลจากโรงพยาบาล King Abdullah University Hospital (KAUH) อย่างไรก็ตาม งานวิจัยนี้เลือกใช้ข้อมูล KAUH ในการทดลองเพื่อประเมินประสิทธิภาพของวิธีการที่นำเสนอ เพราะมีความเหมาะสมต่อการทดลองแบบการตรวจสอบไขว้แบบห้าส่วน โดยแยกรายบุคคล (Subject-independent fivefold cross-validation) เพราะในแต่ละกลุ่มโรคมียังมีจำนวนอาสาสมัครมากกว่า 4 ราย (ปกติ, โรคหอบหืด, โรคปอดอักเสบ, โรคปอดอุดกั้นเรื้อรัง) จึงเหมาะสมต่อการแบ่งส่วนข้อมูลโดยหลีกเลี่ยงการปะปนของข้อมูลบุคคลเดียวกันระหว่างชุดฝึกและชุดทดสอบ

การทดลองดำเนินการบนชุดข้อมูลสาธารณะ โดยมีการจัดกลุ่มผู้เข้าร่วมตามสถานะสุขภาพ 4 กลุ่ม ได้แก่ กลุ่มภาวะปกติ 35 ราย โรคหอบหืด 32 ราย โรคปอดอักเสบ 5 ราย และโรคปอดอุดกั้นเรื้อรัง 9 ราย สัญญาณเสียงปอดถูกบันทึกที่ความถี่การสุ่มตัวอย่าง 4 กิโลเฮิร์ตซ์ รายละเอียดเชิงประชากรศาสตร์ (ช่วงอายุ และสัดส่วนเพศ) รวมทั้งจำนวนสัญญาณหลังการแบ่งช่วงสัญญาณ ดังตารางที่ 3.1

ตารางที่ 3.1 ข้อมูลประชากรของกลุ่มตัวอย่างจากฐานข้อมูล KAUH

กลุ่มตัวอย่าง	จำนวน อาสาสมัคร (คน)	ช่วงอายุ (ปี)	เพศ (หญิง, ชาย)	จำนวนสัญญาณเสียง แบ่งส่วน (*จำนวนสัญญาณ)
ภาวะปกติ	35	18–81	หญิง 11, ชาย 24	415 (105)
โรคหอบหืด	32	12–72	หญิง 17, ชาย 15	357 (96)
โรคปอดอักเสบ	5	36–70	หญิง 2, ชาย 3	66 (15)
โรคปอดอุดกั้นเรื้อรัง	9	42–76	หญิง 1, ชาย 8	102 (27)

หมายเหตุ *ตัวเลขในวงเล็บแสดงจำนวนสัญญาณเสียงปอดต้นฉบับ (Original lung sounds)

การเก็บข้อมูลดำเนินการที่ King Abdullah University Hospital ประเทศจอร์แดน โดยใช้หูฟังแพทย์อิเล็กทรอนิกส์ 3M™ Littmann® Electronic Stethoscope model 3200 เป็นอุปกรณ์หลักในการบันทึกเสียงปอด อุปกรณ์ดังกล่าวสามารถบันทึกเสียงจากร่างกายผู้ป่วยและส่งข้อมูลไปยังคอมพิวเตอร์ผ่าน Bluetooth adaptor จากนั้นผู้วิจัยใช้ซอฟต์แวร์ 3M™ Littmann® StethAssist Visualization Software ในการดึงข้อมูลเสียงจากหน่วยความจำของหูฟังแพทย์และส่งออกไฟล์เสียงในรูปแบบ .wav

ในการบันทึกเสียง แพทย์ผู้เชี่ยวชาญเป็นผู้วางหูฟังแพทย์บนตำแหน่งต่าง ๆ ของผนังทรวงอก โดยตำแหน่งที่ใช้บันทึกเสียงไม่ได้ถูกกำหนดให้เหมือนกันทุกคนแบบตายตัว แต่ขึ้นอยู่กับการประเมินและการตรวจวินิจฉัยของแพทย์ในแต่ละกรณี ตำแหน่งการเก็บเสียงถูกระบุเป็นเขตของทรวงอก ได้แก่ ด้านหน้าและด้านหลังของทรวงอก ด้านซ้ายและด้านขวา รวมถึงบริเวณส่วนบนส่วนกลาง และส่วนล่างของทรวงอก การระบุตำแหน่งดังกล่าวมีความสำคัญ เนื่องจากเสียงปอดที่บันทึกได้อาจแตกต่างกันตามตำแหน่งของปอดและลักษณะของพยาธิสภาพในผู้ป่วยแต่ละราย

เสียงที่บันทึกจากผู้เข้าร่วมแต่ละรายมีความยาวประมาณ 5–30 วินาที ซึ่งเป็นระยะเวลาที่เพียงพอสำหรับครอบคลุมอย่างน้อยหนึ่งรอบการหายใจ อย่างไรก็ตาม บทความต้นฉบับระบุว่าไม่ได้มีการกำหนดระยะเวลาขั้นต่ำในการบันทึกเสียงอย่างเคร่งครัด เนื่องจากการบันทึกขึ้นอยู่กับ การตรวจของแพทย์และข้อจำกัดของอุปกรณ์ โดยระยะเวลาสูงสุดของไฟล์เสียงถูกจำกัดจากความสามารถของหูฟังแพทย์อิเล็กทรอนิกส์ที่ใช้ในการศึกษา

ไฟล์เสียงในชุดข้อมูลถูกส่งออกเป็นไฟล์ .wav ภายใต้ตัวกรองความถี่ 3 รูปแบบ ได้แก่ Bell mode, Diaphragm mode และ Extended mode โดยแต่ละตัวกรองมีคุณสมบัติใน

การเน้นช่วงความถี่ของเสียงแตกต่างกัน Bell mode เน้นเสียงความถี่ต่ำ Diaphragm mode เน้นเสียงในช่วงความถี่ที่เหมาะสมกับการฟังเสียงปอดมากขึ้น และ Extended mode ครอบคลุมช่วงความถี่ที่กว้างและเน้นเสียงในช่วงกลาง การมีไฟล์เสียงจากตัวกรองทั้งสามรูปแบบช่วยให้สามารถนำข้อมูลไปวิเคราะห์เปรียบเทียบผลของช่วงความถี่ต่อประสิทธิภาพของแบบจำลอง หรือเลือกใช้ตัวกรองที่เหมาะสมกับวัตถุประสงค์ของการวิเคราะห์ได้

ชุดข้อมูลยังประกอบด้วยไฟล์ data annotation.xlsx ซึ่งระบุข้อมูลประกอบของแต่ละตัวอย่าง ได้แก่ อายุ เพศ การวินิจฉัยโรค ชนิดของเสียงปอด และตำแหน่งที่บันทึกเสียงบนทรวงอก ข้อมูลเหล่านี้ถูกใช้เป็นข้อมูลกำกับในการวิเคราะห์และสร้างแบบจำลองจำแนกประเภทเสียงปอดหรือโรคระบบทางเดินหายใจ ทั้งนี้ ข้อมูลที่เผยแพร่ไม่มีข้อมูลที่สามารถระบุตัวตนของผู้เข้าร่วมได้โดยตรง และการเก็บข้อมูลได้รับการอนุมัติจากคณะกรรมการจริยธรรมการวิจัยของ King Abdullah University Hospital และ Jordan University of Science and Technology

3.1.2 การเตรียมสัญญาณและการแบ่งช่วงสัญญาณ

เพื่อเตรียมข้อมูลสำหรับการเรียนรู้ของแบบจำลอง สัญญาณเสียงปอดในแต่ละไฟล์ถูกแบ่งออกเป็นช่วงสัญญาณด้วยหน้าต่างเวลาแบบไม่ทับซ้อน (Non-overlapping window) โดยกำหนดความยาวหน้าต่างเท่ากับ 4 วินาที ค่าดังกล่าวได้มาจากการทดลองของงานวิจัยนี้ ซึ่งมีความสอดคล้องกับแนวทางในงานวิจัยก่อนหน้า ที่ใช้การแบ่งสัญญาณระดับหลายวินาที เพื่อครอบคลุมลักษณะเชิงคลินิกของเสียงปอด (Abdulhameed, M. F., Al-Zubaidi, A. S., Mohammed, M. A., and Al-Hayani, A. A., 2020; Rocha, B. M., Filos, D., Mendes, L., Vogiatzis, I., Perantoni, E., Kaimakamis, E., Natsiavas, P., Jaume-i-Capó, A., and Paiva, R. P., 2020)

ผลจากการแบ่งช่วงสัญญาณด้วยหน้าต่างความยาว 4 วินาที การเลือกความยาวหน้าต่าง 4 วินาที เป็นค่าประนีประนอมระหว่างการครอบคลุมลักษณะเชิงคลินิกของเสียงปอดในระดับรอบการหายใจ และการควบคุมจำนวนเฟรมให้เหมาะสมต่อการฝึกแบบจำลอง (Chambres et al., 2018; Rangayyan, 2015) ทำให้ได้จำนวนเฟรมแตกต่างกันในแต่ละกลุ่มโรค ดังที่สรุปไว้ในตารางที่ 3.1 โดยกลุ่มภาวะปอดมีจำนวนเฟรมมากที่สุด จำนวน 415 เฟรม ข้อมูลที่ใช้ในการศึกษาได้มาจากแฟ้มข้อมูลสัญญาณเสียงปอดต้นฉบับจำนวน 105 แฟ้ม รองลงมา คือ กลุ่มโรคหอบหืดจำนวน 357 เฟรม จาก 96 แฟ้ม กลุ่มโรคปอดอุดกั้นเรื้อรัง จำนวน 102 เฟรม จาก 27 แฟ้ม และกลุ่มโรคปอดอักเสบ จำนวน 66 เฟรม จาก 15 แฟ้ม ตามลำดับ ความแตกต่างของจำนวนเฟรม

ระหว่างกลุ่มโรคแสดงถึงลักษณะความไม่สมดุลของข้อมูล ซึ่งเป็นปัจจัยสำคัญที่ต้องนำมาพิจารณาในการออกแบบการฝึก และการประเมินประสิทธิภาพของแบบจำลอง

3.1.3 การจัดการข้อมูลไม่สมดุล

ชุดข้อมูลเสียงปอดที่ใช้ในงานวิจัยนี้มีลักษณะเป็นข้อมูลไม่สมดุล (Imbalanced dataset) เนื่องจากจำนวนข้อมูลในแต่ละกลุ่มโรคมีความแตกต่างกัน ทั้งในระดับจำนวนอาสาสมัคร จำนวนแฟ้มข้อมูลเสียง และจำนวนช่วงสัญญาณที่ได้หลังจากการแบ่งข้อมูล โดยงานวิจัยนี้พิจารณาการจำแนกโรคปอดออกเป็น 4 กลุ่ม ได้แก่ ภาวะปกติ โรคหอบหืด โรคปอดอักเสบ และโรคปอดอุดกั้นเรื้อรัง ซึ่งจากฐานข้อมูล KAUH พบว่าจำนวนอาสาสมัครของแต่ละกลุ่มไม่เท่ากัน โดยกลุ่มภาวะปกติและโรคหอบหืดมีจำนวนตัวอย่างมากกว่ากลุ่มโรคปอดอักเสบและโรคปอดอุดกั้นเรื้อรังอย่างชัดเจน ฐานข้อมูลต้นฉบับประกอบด้วยอาสาสมัครกลุ่มปกติ 35 ราย โรคหอบหืด 32 ราย โรคปอดอักเสบ 5 ราย และโรคปอดอุดกั้นเรื้อรัง 9 ราย ซึ่งสะท้อนให้เห็นว่าชุดข้อมูลมีการกระจายตัวของจำนวนตัวอย่างไม่เท่ากันตั้งแต่ระดับแหล่งข้อมูลต้นฉบับ

ภายหลังจากการเตรียมสัญญาณและการแบ่งช่วงสัญญาณเสียง ข้อมูลที่ใช้ในการศึกษายังคงมีจำนวนตัวอย่างในแต่ละกลุ่มแตกต่างกัน โดยกลุ่มโรคหอบหืดมีจำนวน 357 แฟ้ม จาก 96 แฟ้ม กลุ่มโรคปอดอุดกั้นเรื้อรังมีจำนวน 102 แฟ้ม จาก 27 แฟ้ม และกลุ่มโรคปอดอักเสบมีจำนวน 66 แฟ้ม จาก 15 แฟ้ม ความแตกต่างของจำนวนแฟ้มระหว่างกลุ่มโรคดังกล่าวถือเป็นปัจจัยสำคัญที่อาจส่งผลกระทบต่อกระบวนการเรียนรู้ของแบบจำลอง เนื่องจากแบบจำลองอาจมีแนวโน้มเรียนรู้ลักษณะของกลุ่มที่มีจำนวนข้อมูลมากกว่าได้ดีกว่ากลุ่มที่มีจำนวนข้อมูลน้อยกว่า

ในงานวิจัยนี้ไม่ได้ใช้วิธีการปรับสมดุลข้อมูลโดยตรง เช่น การเพิ่มจำนวนข้อมูลด้วยการสุ่มซ้ำ (Oversampling) การลดจำนวนข้อมูลจากกลุ่มที่มีมากกว่า (Undersampling) หรือการสร้างข้อมูลสังเคราะห์ เนื่องจากข้อมูลเสียงปอดเป็นข้อมูลชีวการแพทย์ที่มีลักษณะเฉพาะของแต่ละบุคคลและแต่ละโรค การเพิ่มหรือลดจำนวนข้อมูลโดยตรงอาจทำให้รูปแบบของสัญญาณผิดไปจากลักษณะข้อมูลจริง หรืออาจทำให้เกิดอคติจากข้อมูลที่ถูกรวบรวมขึ้นซ้ำ ดังนั้น งานวิจัยนี้จึงเลือกจัดการผลกระทบของข้อมูลไม่สมดุลผ่านการออกแบบกระบวนการทดลอง การแบ่งข้อมูล และการเลือกตัวชี้วัดประเมินผลให้เหมาะสมกับลักษณะของชุดข้อมูล

แนวทางแรก คือ การเลือกใช้ข้อมูลจากฐานข้อมูล KAUH เฉพาะกลุ่มโรคที่มีจำนวนอาสาสมัครเพียงพอต่อการประเมินผลแบบการตรวจสอบไขว้ห้าส่วนโดยแยกตามบุคคล หรือ Subject-independent 5-fold cross-validation ได้แก่ ภาวะปกติ โรคหอบหืด โรคปอดอักเสบ และโรคปอดอุดกั้นเรื้อรัง โดยข้อมูลของอาสาสมัครรายเดียวกันจะถูกจัดให้อยู่ในส่วนเดียวกันเสมอ

ไม่ให้เฟรมเสียงจากบุคคลเดียวกันปรากฏอยู่ทั้งในชุดฝึกและชุดทดสอบ วิธีดังกล่าวช่วยลดปัญหาการรั่วไหลของข้อมูล และทำให้ผลการประเมินสะท้อนความสามารถของแบบจำลองในการจำแนกข้อมูลจากบุคคลใหม่ได้เหมาะสมมากขึ้น

งานวิจัยนี้ใช้การแบ่งข้อมูลแบบการตรวจสอบไขว้แบบห้าส่วน โดยแยกตามบุคคล เพื่อหลีกเลี่ยงการรั่วไหลของข้อมูลระหว่างชุดฝึกและชุดทดสอบ โดยกำหนดให้ข้อมูลเสียงปอดทั้งหมดของอาสาสมัครรายเดียวกันต้องอยู่ในส่วนเดียวกันเสมอ ไม่ว่าจะเป็นเฟรมเสียงหรือช่วงสัญญาณย่อยใด ๆ แนวทางดังกล่าวมีความสำคัญอย่างยิ่งในงานวิเคราะห์เสียงทางการแพทย์ ช่วยให้การประเมินผลแสดงความสามารถของแบบจำลองในการทำให้เป็นสากลต่อผู้ป่วยรายใหม่ได้อย่าง การแบ่งข้อมูลดำเนินการโดยจัดอาสาสมัครออกเป็น 5 ส่วน และหมุนเวียนให้แต่ละส่วนทำหน้าที่เป็นชุดทดสอบ 1 ครั้ง ในขณะที่อีก 4 ส่วน ใช้เป็นชุดฝึกและชุดตรวจสอบในแต่ละรอบการทดลอง การออกแบบในลักษณะนี้ช่วยลดอคติจากการเลือกชุดทดสอบ และเพิ่มความน่าเชื่อถือของผล

Subjects → [Fold 1] [Fold 2] [Fold 3] [Fold 4] [Fold 5]

Iteration 1: Test = Fold 1, Train = Fold 2–5

Iteration 2: Test = Fold 2, Train = Fold 1,3–5

...

Iteration 5: Test = Fold 5, Train = Fold 1-4

รูปที่ 3.1 การแบ่งข้อมูลแบบการตรวจสอบไขว้แบบห้าส่วน

แนวทางที่สอง คือ การใช้แบบจำลองออโตเอนโคเดอร์แบบหลายงาน (Multi-Task Autoencoder: MTAE) เพื่อเรียนรู้ตัวแทนเชิงแฝงของสัญญาณเสียงปอด โดยแบบจำลองเรียนรู้งานการสร้างข้อมูลกลับคืนร่วมกับการจำแนกประเภท ทำให้ตัวแทนข้อมูลที่ได้ยังคงรักษาสารสนเทศสำคัญของสัญญาณเสียงต้นฉบับไว้ ขณะเดียวกันก็เพิ่มความสามารถในการแยกแยะกลุ่มโรค นอกจากนี้ งานวิจัยยังนำตัวแทนเชิงแฝงที่ได้จาก MTAE ไปใช้ร่วมกับซัพพอร์ตเวกเตอร์แมชชีน เพื่อช่วยกำหนดขอบเขตการตัดสินใจระหว่างกลุ่มโรคให้ชัดเจนยิ่งขึ้น ซึ่งสอดคล้องกับแนวทางหลักของวิทยานิพนธ์ที่มุ่งพัฒนาแบบจำลองแบบผสมระหว่างการเรียนรู้เชิงลึกและการเรียนรู้ของเครื่อง

แนวทางที่สาม คือ การประเมินประสิทธิภาพของแบบจำลองโดยไม่พิจารณาเฉพาะค่าความถูกต้องรวม (Accuracy) เพียงอย่างเดียว เนื่องจากในกรณีที่ข้อมูลแต่ละกลุ่มมีจำนวนไม่เท่ากัน ค่าความถูกต้องอาจได้รับอิทธิพลจากกลุ่มข้อมูลที่มีจำนวนมากกว่า และอาจไม่สะท้อนความสามารถของแบบจำลองต่อกลุ่มโรคที่มีข้อมูลน้อยได้อย่างชัดเจน ดังนั้น งานวิจัยนี้จึงรายงานตัวชี้วัดหลายด้าน ได้แก่ ความแม่นยำ (Precision) ความไวหรือรีคอลล์ (Sensitivity/Recall) และค่าคะแนน F1 (F1-score) รวมถึงการวิเคราะห์เมตริกซ์สับสน เพื่อใช้พิจารณาความสามารถของแบบจำลองในแต่ละกลุ่มโรคอย่างครอบคลุม โดยเฉพาะค่าคะแนน F1 ซึ่งสะท้อนความสมดุลระหว่างความแม่นยำ และความไวหรือรีคอลล์ ได้ดีกว่าการใช้ความถูกต้องเพียงค่าเดียวในกรณีที่ข้อมูลมีความไม่สมดุล

ดังนั้น การจัดการข้อมูลไม่สมดุลในงานวิจัยนี้เป็นการจัดการในเชิงระเบียบวิธีการทดลองมากกว่าการปรับจำนวนข้อมูลโดยตรง กล่าวคือ งานวิจัยนี้ควบคุมผลกระทบของข้อมูลไม่สมดุลผ่านการเลือกชุดข้อมูลที่เหมาะสม การแบ่งข้อมูลแบบแยกตามบุคคล การใช้แบบจำลองที่สามารถเรียนรู้ตัวแทนเชิงแฝงของสัญญาณเสียงปอดได้อย่างมีประสิทธิภาพ และการใช้ตัวชี้วัดประเมินผลที่ครอบคลุมหลายมิติ แนวทางดังกล่าวช่วยให้การประเมินประสิทธิภาพของแบบจำลองมีความน่าเชื่อถือมากขึ้น และลดความเอนเอียงของผลลัพธ์ที่อาจเกิดจากจำนวนข้อมูลที่แตกต่างกันระหว่างกลุ่มโรค

3.2 การออกแบบวิธีการที่นำเสนอ

3.2.1 ออโตเอนโคเดอร์แบบหลายงาน

เมื่อผ่านขั้นตอนการเตรียมสัญญาณ สัญญาณจะถูกนำเข้าสู่กระบวนการสกัดคุณลักษณะ คือ สัมประสิทธิ์เซปสตรัมความถี่เมล ซึ่งเป็นคุณลักษณะที่ออกแบบมาเพื่ออธิบายลักษณะสเปกตรัมระยะสั้น (Short-term spectrum) ของสัญญาณเสียง และแปลงสเกลความถี่ให้สอดคล้องกับการรับรู้ของการได้ยินของมนุษย์ การสร้างคุณลักษณะประกอบด้วย (1) การคำนวณสเปกตรัมกำลังด้วยฟูริเยร์ (Power spectrum calculation using Fourier transform) (2) การทำแผนที่ความถี่เข้าสู่สเกลเมลผ่านชุดตัวกรองเมล (Mel filter bank) (3) การแปลงลอการิทึมเพื่อลดช่วงไดนามิกของพลังงาน และ (4) การแปลงโคไซน์แบบไม่ต่อเนื่องเพื่อให้ได้ค่าสัมประสิทธิ์เซปสตรัมที่มีการรวมตัวของข้อมูลและลดมิติได้อย่างมีประสิทธิภาพ

กำหนดให้การแทนข้อมูลด้วยสัมประสิทธิ์เซปสตรัมความถี่เมล เป็นคุณลักษณะหลักที่ใช้ป้อนเข้าสู่แบบจำลอง มีการกำหนดพารามิเตอร์ ได้แก่ ความยาวเฟรม (Frame length) ทำการ

ทดสอบหลายระดับ การใช้หน้าต่างแฮมมิง (Hamming window) การซ้อนทับของเฟรม 50 เปอร์เซ็นต์ และการคำนวณค่าสัมประสิทธิ์เซปสตรัมความถี่หลัก รวมถึงการเพิ่มอนุพันธ์อันดับหนึ่งและอันดับสองเพื่อแสดงการเปลี่ยนแปลงเชิงเวลา

ตารางที่ 3.2 พารามิเตอร์การสกัดคุณลักษณะด้วยสัมประสิทธิ์เซปสตรัมความถี่

พารามิเตอร์	ค่าที่กำหนด
ความยาวช่วงสัญญาณที่ใช้แบ่ง (Segment duration)	4 วินาที (ไม่ซ้อนทับ)
ความถี่สุ่มตัวอย่าง (Sampling frequency)	4 กิโลเฮิร์ตซ์
ความยาวเฟรม	20, 40, 128, 256, 512, 1000 มิลลิวินาที
หน้าต่างสัญญาณ	หน้าต่างแฮมมิง
การซ้อนทับเฟรม	50% overlap
จำนวนสัมประสิทธิ์เซปสตรัมความถี่	13 มิติ

แนวความคิดการเรียนรู้หลายงานในลักษณะการแชร์ตัวแทนเชิงแฝง สอดคล้องกับงานวิจัยดั้งเดิมด้านการเรียนรู้แบบหลายงาน ซึ่งชี้ให้เห็นว่าการเรียนรู้หลายงานสามารถทำหน้าที่เป็นกลไกการควบคุมความซับซ้อนของแบบจำลอง ช่วยลดการเรียนรู้เกิน และเพิ่มความสามารถในการทำแบบจำลองสามารถใช้งานได้กับข้อมูลที่ไม่เคยพบมาก่อน (Caruana, 1997)

ในบริบทของข้อมูลเสียงพูด ซึ่งมีความแปรปรวนสูงจากปัจจัยผู้ป่วย สภาพแวดล้อม และวิธีการบันทึก การใช้การเรียนรู้หลายงานจึงมีบทบาทสำคัญในการเพิ่มคุณภาพของตัวแทนเชิงแฝง งานการสร้างข้อมูลกลับคืนช่วยรักษาคุณลักษณะสัญญาณพื้นฐาน ขณะที่งานการจำแนกช่วยเน้นสารสนเทศเชิงคลินิกที่เกี่ยวข้องกับโรค ส่งผลให้ตัวแทนเชิงแฝงที่ได้มีความสมดุลและเหมาะสมต่อการนำไปใช้ในขั้นตอนการจำแนก การออกแบบสถาปัตยกรรมออโตเอนโคเดอร์แบบหลายงานมีพื้นฐานทั้งจากแนวคิดระเบียบวิธีและจากทฤษฎีการเรียนรู้หลายงานที่ได้รับการยอมรับ ช่วยสนับสนุนการใช้แบบจำลองกับการจำแนกเสียงพูดที่มีข้อมูลจำกัดและความแปรปรวนสูง

ตารางที่ 3.3 สถาปัตยกรรมของแบบจำลองออโตเอนโคเดอร์แบบหลายงาน

Module	Layer	Kernel	Layer configuration	Activation
Autoencoder (Encoder1)	Conv2d 1	3x3	in_channels=1,	ReLU
	Conv2d 2	3x3	out_channels=16	ReLU
Autoencoder (Encoder2)	Conv2d 3	3x3	in_channels=16,	ReLU
			out_channels=8	
Autoencoder (Encoder3)			in_channels=8,	
			out_channels=2	
Autoencoder (Decoder1)	ConvTranspose2d 1	3x3	in_channels=2,	ReLU
	ConvTranspose2d 2	3x3	out_channels=8	ReLU
Autoencoder (Decoder2)	ConvTranspose2d 3	3x3	in_channels=8,	ReLU and
			out_channels=16	Sigmoid
Autoencoder (Decoder3)			in_channels=16,	
			out_channels=1	
Classification (Block1)	FC (Linear 1)	-	in_features=[Flatten	ReLU
	FC (Linear 2)	-	latent size],	ReLU
Classification (Block2)	FC (Linear 3)	-	out_features=128	ReLU
	Classifier (Linear 4)	-	in_features=128,	Softmax
Classification (Block3)			out_features=128	
			in_features=128,	
			out_features=128	
			in_features=128,	
			out_features=4	

ในตารางที่ 3.3 ออกแบบให้มีการแบ่งปันตัวแทนเชิงแฝงร่วมกันระหว่างงาน การสร้างข้อมูลกลับคืน และงานการจำแนก การออกแบบในลักษณะนี้ช่วยให้ตัวแทนเชิงแฝงรักษาทั้งสารสนเทศเชิงโครงสร้างและสารสนเทศที่เกี่ยวข้องกับการจำแนกโรคไว้พร้อมกัน

3.2.2 ออโตเอนโคเดอร์แบบหลายงานร่วมกับซัพพอร์ตเวกเตอร์แมชชีน

หลังจากได้ตัวแทนเชิงแฝงที่เรียนรู้จากแบบจำลองออโตเอนโคเดอร์แบบหลายงานในหัวข้อก่อน ตัวแทนดังกล่าวถูกนำมาใช้เป็นอินพุตของกระบวนการจำแนกโรค โดยงานวิจัยนี้เลือกใช้ซัพพอร์ตเวกเตอร์แมชชีน เป็นตัวจำแนกหลักในขั้นตอนสุดท้ายแทนการใช้ชั้นฟังก์ชันซอฟต์แวร์แมชชีนของโครงข่ายประสาทเทียม ทั้งนี้ การเลือกใช้ซัพพอร์ตเวกเตอร์แมชชีนมีเหตุผลเชิงระเบียบวิธีที่สอดคล้องกับลักษณะของข้อมูลเสียงปอด ซึ่งมีมิติของคุณลักษณะที่มีแนวโน้มสูง แต่มีจำนวนตัวอย่างจำกัด และมีความแปรปรวนสูง

หลักการของซัพพอร์ตเวกเตอร์แมชชีน เป็นตัวจำแนกแบบอัตราระยะขอบ (Margin) ที่มุ่งหาขอบเขตการตัดสินใจที่มีระยะขอบมากที่สุดระหว่างกลุ่มต่าง ๆ ซึ่งช่วยลดความเสี่ยงของการเรียนรู้เกิน และเพิ่มความสามารถในการทำให้แบบจำลองทั่วไปได้ดีกับข้อมูลที่ไม่เคยพบมาก่อน โดยเฉพาะในกรณีที่ข้อมูลมีการกระจายไม่สมดุลหรือมีการซ้อนทับกันของกลุ่มบางส่วน จากลักษณะดังกล่าวทำให้ซัพพอร์ตเวกเตอร์แมชชีนมาใช้อย่างแพร่หลายในงานจำแนกสัญญาณชีวการแพทย์ รวมถึงงานวิเคราะห์เสียงปอดและเสียงทางเดินหายใจ

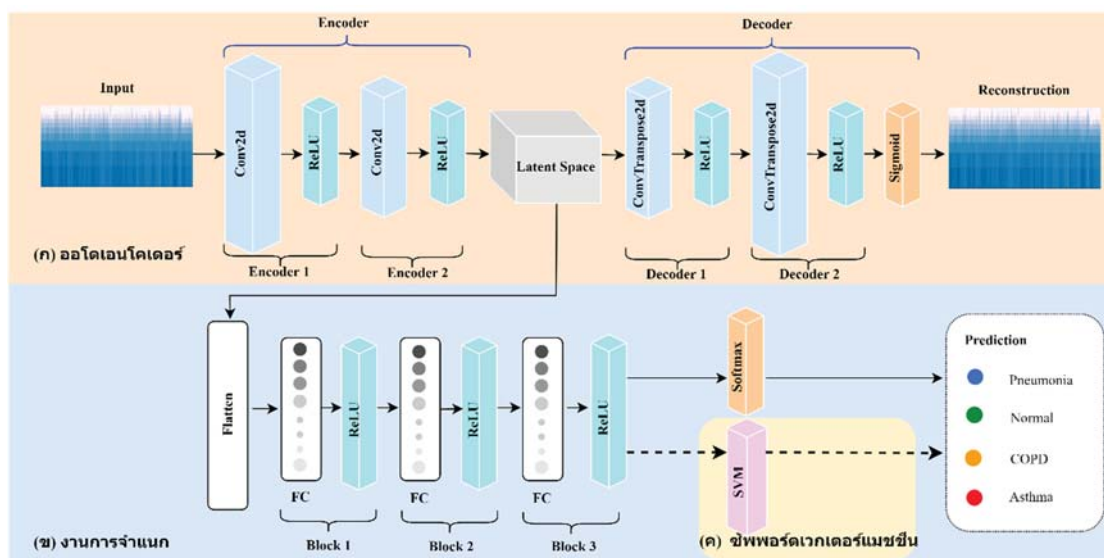
ในงานวิจัยขั้นตอนการจำแนกขั้นสุดท้ายถูกออกแบบให้ใช้ซัพพอร์ตเวกเตอร์แมชชีนร่วมกับตัวแทนเชิงแฝงหรือคุณลักษณะเชิงลึกที่ได้จากแบบจำลองออโตเอนโคเดอร์แบบหลายงาน เพื่อเพิ่มประสิทธิภาพของขอบเขตการตัดสินใจ โดยเฉพาะเมื่อข้อมูลมีความแปรปรวนสูงและมีจำนวนตัวอย่างจำกัด ซึ่งเป็นลักษณะสำคัญของข้อมูลเสียงปอดในบริบททางคลินิก กระบวนการทำงานโดยรวมดำเนินการในรูปแบบ 2 ขั้นตอน คือ (1) ฝึกออโตเอนโคเดอร์แบบหลายงาน ให้เรียนรู้คุณลักษณะที่มีความหมายจากข้อมูลเสียงปอด และ (2) ใช้ออโตเอนโคเดอร์แบบหลายงานที่ฝึกแล้วเป็นตัวสกัดคุณลักษณะ ก่อนนำคุณลักษณะดังกล่าวเข้าสู่ซัพพอร์ตเวกเตอร์แมชชีนเพื่อทำการจำแนกโรคต่อไป

เมื่อทำการฝึกออโตเอนโคเดอร์แบบหลายงาน นำคุณลักษณะจากชั้นเชื่อมโยงสมบูรณ์ขั้นสุดท้ายของออโตเอนโคเดอร์แบบหลายงาน มาใช้เป็นข้อมูลนำเข้าสำหรับซัพพอร์ตเวกเตอร์แมชชีนแทนการใช้ชั้นจำแนกแบบเดิมของโครงข่ายประสาทเทียม จากนั้นในช่วงทดสอบจะใช้ออโตเอนโคเดอร์แบบหลายงานที่ฝึกแล้วสกัดคุณลักษณะจากข้อมูลทดสอบ และให้ซัพพอร์ตเวกเตอร์แมชชีนทำนายกลุ่มโรคจากคุณลักษณะดังกล่าว แนวทางนี้มีนัยเชิงระเบียบวิธีที่สำคัญ คือ แยกบทบาทระหว่างการเรียนรู้ตัวแทน และการสร้างขอบเขตการตัดสินใจ ทำให้การประเมินแสดงความสามารถในการทำให้เป็นสากลของคุณลักษณะที่เรียนรู้ได้ชัดเจน

การตั้งค่าพารามิเตอร์ของซัพพอร์ตเวกเตอร์แมชชีน การจำแนกโดยใช้เคอร์เนลฟังก์ชันเรเดียล ซึ่งพิจารณาพารามิเตอร์สำคัญ 2 ตัว ได้แก่ พารามิเตอร์ควบคุมการแลกเปลี่ยนระหว่างความกว้างของระยะขอบและความผิดพลาดจากการจำแนก และพารามิเตอร์ที่ควบคุมความกว้างของเคอร์เนล คือ พารามิเตอร์แกมมา การเลือกค่าพารามิเตอร์อาศัยการวิเคราะห์บนชุดข้อมูลตรวจสอบ (Validation) ของแบบจำลองที่ฝึกจากชุดฝึก จากทฤษฎีของซัพพอร์ตเวกเตอร์แมชชีนและงานศึกษาพื้นฐานด้านการเรียนรู้ด้วยเคอร์เนล พบว่าค่าพารามิเตอร์ที่ควบคุมความกว้างของเคอร์เนลที่มีค่าสูงจะทำให้ขอบเขตการตัดสินใจมีความซับซ้อนและมีความไวต่อข้อมูลรบกวน ขณะที่ค่าขนาดเล็กจะส่งผลให้ขอบเขตการตัดสินใจมีความเรียบและช่วยเพิ่มความสามารถในการใช้งานทั่วไปของแบบจำลอง โดยเฉพาะในกรณีที่ข้อมูลมีมิติสูงหรือจำนวนตัวอย่างมีจำกัด (Schölkopf, B., and Smola, A. J., 2002; Hastie, T., Tibshirani, R., and Friedman, J., 2009) แนวคิดดังกล่าวสอดคล้องกับหลักการของซัพพอร์ตเวกเตอร์แมชชีน ที่มุ่งเน้นการเพิ่มระยะขอบสูงสุด (maximum margin) เพื่อควบคุมความซับซ้อนของแบบจำลองและลดความเสี่ยงของการเรียนรู้เกิน

นอกจากนี้ งานวิจัยที่ประยุกต์ใช้ซัพพอร์ตเวกเตอร์แมชชีนร่วมกับคุณลักษณะเชิงลึกหรือคุณลักษณะที่ผ่านการเรียนรู้แล้ว ระบุว่าเมื่อคุณลักษณะมีการแยกกลุ่มเชิงโครงสร้างที่ดีอยู่แล้ว การใช้ค่าพารามิเตอร์แกมมา ในช่วงต่ำสามารถช่วยให้ซัพพอร์ตเวกเตอร์แมชชีนทำหน้าที่เป็นกลไกการตัดสินใจที่มีเสถียรภาพ โดยไม่จำเป็นต้องปรับพารามิเตอร์อย่างละเอียด งานของ Niu, X., and Suen, C. Y. (2012) แสดงให้เห็นว่าการใช้ซัพพอร์ตเวกเตอร์แมชชีนบนคุณลักษณะที่สกัดจากโครงข่ายเชิงลึกสามารถให้ผลการจำแนกที่มีความเสถียรโดยอาศัยการกำหนดค่าพารามิเตอร์ในช่วงมาตรฐาน ขณะที่ Huang, G., Song, S., Gupta, J. N. D., and Wu, C. (2014) ชี้ให้เห็นว่าตัวจำแนกแบบมาร์จิน มีประสิทธิภาพสูงเมื่อทำงานบนตัวแทนข้อมูลที่มีโครงสร้างชัดเจนจากกระบวนการเรียนรู้ล่วงหน้า

สถาปัตยกรรมและลำดับขั้นตอนการทำงานของแบบจำลองที่นำเสนอ ซึ่งถูกออกแบบขึ้นเพื่อแก้ไขข้อจำกัดของการเรียนรู้แบบงานเดียวในการวิเคราะห์เสียงปอด โดยผสานแนวคิดการเรียนรู้หลายงานเข้ากับตัวจำแนกแบบระยะขอบ เพื่อยกระดับทั้งความแม่นยำและความเสถียรของการจำแนกโรกระบบทางเดินหายใจ ลำดับขั้นตอนตั้งแต่การรับสัญญาณ การสกัดคุณลักษณะ การเรียนรู้ตัวแทนเชิงแฝง ไปจนถึงการจำแนกโรค ดังรูปที่ 3.2 แบบจำลองออโตเอนโคเดอร์แบบหลายงานร่วมกับซัพพอร์ตเวกเตอร์แมชชีน ซึ่งประกอบด้วย 3 ส่วนหลัก ได้แก่



รูปที่ 3.2 แบบจำลองออโตเอนโคเดอร์แบบหลายงานร่วมกับซัพพอร์ตเวกเตอร์แมชชีน

ในส่วนที่ (ก) ออโตเอนโคเดอร์ ทำหน้าที่เรียนรู้ตัวแทนเชิงแฝงจากข้อมูลอินพุต โดยส่วนตัวเข้ารหัส ประกอบด้วยบล็อกคอนโวลูชันสองชั้น (Encoder 1 และ Encoder 2) ซึ่งใช้เลเยอร์ Convolution 2D ร่วมกับฟังก์ชันกระตุ้นฟังก์ชันเรกติไฟต์ลิเนียร์ยูนิต เพื่อสกัดคุณลักษณะเชิงลึกและลดมิติของข้อมูล จากนั้นข้อมูลจะถูกแปลงให้อยู่ในรูปของปริภูมิเชิงแฝง ซึ่งเป็นตัวแทนเชิงแฝงที่สรุปสารสนเทศสำคัญของสัญญาณเสียงปอด ส่วนตัวถอดรหัสประกอบด้วยเลเยอร์คอนโวลูชันทรานสโพสแบบสองมิติ (Convolution Transpose 2D) และฟังก์ชันกระตุ้นฟังก์ชันเรกติไฟต์ลิเนียร์ยูนิต เพื่อสร้างสัญญาณกลับคืน ให้ใกล้เคียงกับอินพุตเดิม โดยมีวัตถุประสงค์เพื่อบังคับให้ตัวแทนเชิงแฝงรักษาโครงสร้างของสัญญาณต้นฉบับไว้

ในส่วนที่ (ข) งานการจำแนก ตัวแทนเชิงแฝงจากตัวแทนเชิงแฝง จะถูกนำมาทำการปรับรูปข้อมูลเป็นเวกเตอร์หนึ่งมิติ (Flatten) และป้อนเข้าสู่ชั้นเชื่อมโยงสมบูรณ์ จำนวนหลายบล็อก โดยแต่ละบล็อกใช้ฟังก์ชันกระตุ้นฟังก์ชันเรกติไฟต์ลิเนียร์ยูนิต เพื่อเพิ่มความไม่เชิงเส้นในการเรียนรู้คุณลักษณะ จากนั้นมีการใช้ฟังก์ชันซอฟต์แวร์แม็กซ์ เพื่อให้ได้ค่าความน่าจะเป็นของแต่ละคลาสในกรณีการจำแนกหลายกลุ่ม ซึ่งสะท้อนผลการทำนายเบื้องต้นของโครงข่าย

ในส่วนที่ (ค) ส่วนตัวจำแนกขั้นสุดท้ายด้วยซัพพอร์ตเวกเตอร์แมชชีน ตัวแทนเชิงแฝงหรือคุณลักษณะเชิงลึกที่ได้จากการจำแนกจะถูกส่งต่อไปยังตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน ซึ่งทำหน้าที่เป็นตัวจำแนกขั้นสุดท้าย โดยใช้หลักการระยะขอบสูงสุด เพื่อค้นหาไฮเปอร์เพลน

ที่เหมาะสมในการแยกกลุ่มข้อมูล ส่งผลให้การตัดสินใจมีความคมชัดและเสถียรมากขึ้น โดยเฉพาะในกรณีที่กลุ่มโรคมียุคมีลักษณะใกล้เคียงกัน ผลสุดท้าย คือ การจำแนกโรคระบบทางเดินหายใจเป็น 4 กลุ่ม โดยสรุป สถาปัตยกรรมดังกล่าวสะท้อนแนวคิดการผสมผสานระหว่างการเรียนรู้หลายงาน เพื่อยกระดับคุณภาพของตัวแทนเชิงแฝง และการใช้ตัวจำแนกซัพพอร์ตเวกเตอร์แมชชีนเพื่อเพิ่มความเสถียรของการตัดสินใจ

3.3 ทรัพยากรที่ใช้ในการวิจัยวิทยานิพนธ์

เครื่องมือที่ใช้ในการพัฒนางานวิจัยนี้ ใช้หน่วยประมวลผลกลาง : AMD Ryzen Threadripper 3970x หน่วยประมวลผลกราฟิก : NVIDIA GeForce RTX 3090, หน่วยความจำหลัก : 128 GB หน่วยความจำสำรอง SSD/ HDD : 2TB/6TB ระบบปฏิบัติการ : Ubuntu 20.04 ภาษาที่ใช้ในการพัฒนา : Python 3 และเฟรมเวิร์ก : PyTorch ในการฝึกแบบจำลอง มีการกำหนดค่าพารามิเตอร์การเรียนรู้ที่สำคัญ เช่น จำนวนรอบการฝึก ขนาดกลุ่มข้อมูลย่อย ตัวปรับค่าพารามิเตอร์ และฟังก์ชันกระตุ้น เพื่อให้ผลการทดลองมีความเสถียรและสามารถทำซ้ำได้ รายละเอียดของพารามิเตอร์ที่ใช้ในการฝึกแบบจำลอง ดังแสดงในตารางที่ 3.4

ตารางที่ 3.4 พารามิเตอร์สำคัญที่ใช้ในการฝึกแบบจำลอง

พารามิเตอร์	ค่าที่กำหนด
ขนาดแบตช์ (Batch size)	32
จำนวนยุคการฝึก (Epochs)	350
ตัวปรับค่าเหมาะสม (Optimizer)	อדם (Adam)
ค่าความสูญเสีย (Loss)	ครอสเอนโทรปี
ฟังก์ชันกระตุ้นของชั้นเอาต์พุต (Output activation)	ซอฟต์แมกซ์
ฟังก์ชันกระตุ้นของชั้นซ่อน (Hidden activation)	ฟังก์ชันเรกติไฟต์ไลเนียร์ยูนิต

ตารางที่ 3.4 แสดงพารามิเตอร์สำคัญที่ใช้ในการฝึกโครงข่าย ซึ่งทำหน้าที่เรียนรู้ตัวแทนเชิงแฝงของสัญญาณเสียงปอดสำหรับการนำไปใช้ในขั้นตอนการจำแนก โดยกำหนดขนาดแบตช์ 32 เพื่อรักษาสมดุลระหว่างความเสถียรของการเรียนรู้และประสิทธิภาพการคำนวณ จำนวนยุคการฝึก 350 รอบ เพื่อให้สามารถเรียนรู้โครงสร้างเชิงแฝงของข้อมูลได้อย่างเพียงพอ

ในขั้นตอนการฝึกอัตโนมัติเอนโคเดอร์ ฟังก์ชันความสูญเสียแบบครอสเอนโทรปี ถูกใช้สำหรับงานการจำแนกที่ผสมอยู่ในกรอบการเรียนรู้หลายงาน ขณะที่ฟังก์ชันเรกติไฟต์ไลเนียร์ยูนิตถูกใช้เพื่อเพิ่มความไม่เชิงเส้นของแบบจำลอง ส่วนฟังก์ชันซอฟต์แมกซ์ ถูกใช้ในชั้นเอาต์พุตของงานการจำแนกภายในอัตโนมัติเอนโคเดอร์เท่านั้น โดยผลลัพธ์จากชั้นเชิงแฝงที่ได้หลังการฝึกจะถูกนำไปใช้เป็นเวกเตอร์คุณลักษณะอินพุตสำหรับซัพพอร์ตเวกเตอร์แมชชีน ซึ่งทำหน้าที่เป็นตัวจำแนกขั้นสุดท้ายของงานแยกจากกระบวนการฝึกโครงข่ายอัตโนมัติเอนโคเดอร์

3.4 การประเมินประสิทธิภาพ

3.4.1 เมตริกซ์สับสนและตัวชี้วัดการประเมินผลรายกลุ่ม

เมตริกซ์สับสนเป็นเครื่องมือมาตรฐานที่ใช้ในการประเมินประสิทธิภาพของแบบจำลองสำหรับปัญหาการจำแนกประเภท โดยเฉพาะในกรณีการจำแนกหลายกลุ่ม (Multi-class classification) เมตริกซ์ดังกล่าวแสดงความสัมพันธ์ระหว่างป้ายกำกับจริง (True labels) และผลการทำนายของแบบจำลอง (Predicted labels) ในรูปแบบตาราง ซึ่งช่วยให้สามารถวิเคราะห์รูปแบบความถูกต้องและความสับสนระหว่างกลุ่มได้อย่างเป็นระบบ

True	Class 1	Class 2	Class 3	Class 4
	TP1	FP2	FP3	FP4
	FN1	TP2	FP3	FP4
	FN1	FN2	TP3	FP4
Class 4	FN3	FN2	FN3	TP4
	Class 1	Class 2	Class 3	Class 4
Predictions				

(ก) พื้นฐานสำหรับการจำแนก 4 กลุ่ม

Actual label	Pneumonia	Normal	COPD	Asthma
	42	7	4	13
	12	325	33	45
	3	34	55	10
Asthma	16	83	16	242
	Pneumonia	Normal	COPD	Asthma
Predicted label				

(ข) การจำแนกเสียงปอดออกเป็น 4 กลุ่ม

รูปที่ 3.3 เมตริกซ์สับสนแบบจำแนก 4 กลุ่ม

รูปที่ 3.3 (ก) แสดงโครงสร้างพื้นฐานของเมตริกซ์สับสนสำหรับการจำแนก 4 กลุ่ม โดยค่าในแนวทแยงมุม (Diagonal elements) แทนจำนวนตัวอย่างที่แบบจำลองจำแนกได้ถูกต้องในแต่ละกลุ่ม ซึ่งเรียกว่าผลจริงบวก (True Positive: TP) ของกลุ่มนั้น ซึ่งเป็นข้อมูลจริงที่มีผลเฉลยเป็น

ค่าจริง และแบบจำลองทำนายผลว่าอยู่ในเป็นกลุ่มนั้น ซึ่งถือว่าแบบจำลองทำนายถูกต้องตรงกับข้อมูลจริง ขณะที่ค่านอกแนวทแยงมุมแสดงกรณีการจำแนกผิดพลาด ซึ่งสามารถแบ่งออกเป็นค่าผลบวกเทียม (False Positive: FP) ข้อมูลจริงที่มีผลเฉลยเป็นกลุ่มอื่น แต่แบบจำลองทำนายผลว่าอยู่ในกลุ่มค่าจริงบวก ซึ่งถือว่าแบบจำลองทำนายผิดกลุ่มข้อมูล และผลลบเทียม (False Negative: FN) ข้อมูลจริงที่มีผลเฉลยเป็นคลาสบวก แต่แบบจำลองทำนายผลว่าอยู่ในกลุ่มอื่น ซึ่งถือว่าแบบจำลองทำนายผิดจากข้อมูลจริง ทั้งนี้ การกระจายของค่าเหล่านี้สะท้อนระดับความสัมพันธ์ระหว่างกลุ่มต่าง ๆ อย่างชัดเจน สำหรับงานวิจัยนี้ จำแนกเสียงปอดออกเป็น 4 กลุ่ม ได้แก่ ภาวะปกติ โรคปอดอักเสบ โรคปอดอุดกั้นเรื้อรัง และโรคหอบหืด โดยรูปที่ 3.3 (ข) แสดงตัวอย่างเมตริกซ์สับสนที่ได้จากผลการทดลอง จากเมตริกซ์สับสนสามารถคำนวณตัวชี้วัดการประเมินผลได้ดังนี้

ความแม่นยำ สำหรับรายกลุ่ม ดังสมการที่ (3.1)

$$\text{Precision}_{id} = \frac{TP_{id}}{TP_{id} + FP_{id}} \quad (3.1)$$

เมื่อ

Precision_{id}	คือ ค่าความแม่นยำของแบบจำลองสำหรับกลุ่มที่ id
FP_{id}	คือ จำนวนตัวอย่างที่แบบจำลองทำนายว่าเป็นกลุ่มที่ id แต่ในความเป็นจริงเป็นกลุ่มอื่น

ความไว สำหรับรายกลุ่ม ดังสมการที่ (3.2)

$$\text{Recall}_{id} = \frac{TP_{id}}{TP_{id} + FN_{id}} \quad (3.2)$$

เมื่อ

Recall_{id}	คือ ค่าความไวของแบบจำลองสำหรับกลุ่มที่ id
FN_{id}	คือ จำนวนตัวอย่างของกลุ่มที่ id ที่แบบจำลองไม่สามารถตรวจพบได้และถูกจำแนกเป็นกลุ่มอื่น

ค่าคะแนน F1 สำหรับรายกลุ่ม ดังสมการที่ (3.3)

$$F1_{id} = 2 * \frac{Precision_{id} + Recall_{id}}{Precision_{id} \times Recall_{id}} \quad (3.3)$$

เมื่อ

$F1_{id}$ คือ ค่าคะแนน F1 ของแบบจำลองสำหรับกลุ่มที่ id

ค่าคะแนน F1 เป็นตัวชี้วัดเชิงสมดุลระหว่างความแม่นยำและความไว ซึ่งเหมาะสมสำหรับการประเมินประสิทธิภาพในกรณีที่จำนวนตัวอย่างของแต่ละกลุ่มไม่สมดุล

ความถูกต้องโดยรวม (Overall Accuracy: OA) ดังสมการที่ (3.4)

$$OA = \sum_{C_{id=1}}^{C_{all}} \frac{TP_{id}}{N_{all}} \quad (3.4)$$

เมื่อ

C_{id}	คือ ดัชนีของกลุ่ม
C_{all}	คือ จำนวนกลุ่มทั้งหมด
TP_{id}	คือ จำนวนตัวอย่างของกลุ่มที่ id ที่แบบจำลองจำแนกได้ถูกต้อง
N_{all}	คือ จำนวนตัวอย่างทั้งหมด
id	คือ ดัชนีของกลุ่มหรือคลาสที่ใช้ในการจำแนก โดย $id = 1, 2, \dots, C_{all}$

ในงานวิจัยนี้มีการรายงานค่าตัวชี้วัดรายกลุ่ม เพื่อสะท้อนสมรรถนะของแบบจำลองในแต่ละโรคจากเมตริกซ์สับสน โดยสำหรับกลุ่มที่ id ได้คำนวณความไว ($Recall_{id}$) ซึ่งนิยามเป็นอัตราส่วนระหว่างจำนวนตัวอย่างที่ถูกจำแนกถูกต้องของกลุ่มนั้นต่อจำนวนตัวอย่างทั้งหมดของกลุ่ม ดังแสดงในสมการที่ (3.3) และในงานวิจัยนี้กำหนดสัญลักษณ์การรายงานรายโรคเป็น ACC_{Normal} , $ACC_{Pneumonia}$, ACC_{Asthma} และ ACC_{COPD} สำหรับกลุ่มภาวะปกติ กลุ่มโรคปอดอักเสบ กลุ่มโรคปอดอุดกั้นเรื้อรัง และโรคหอบหืด ตามลำดับ เพื่อให้การนำเสนอผลลัพธ์ในแต่ละกลุ่มมีความชัดเจน

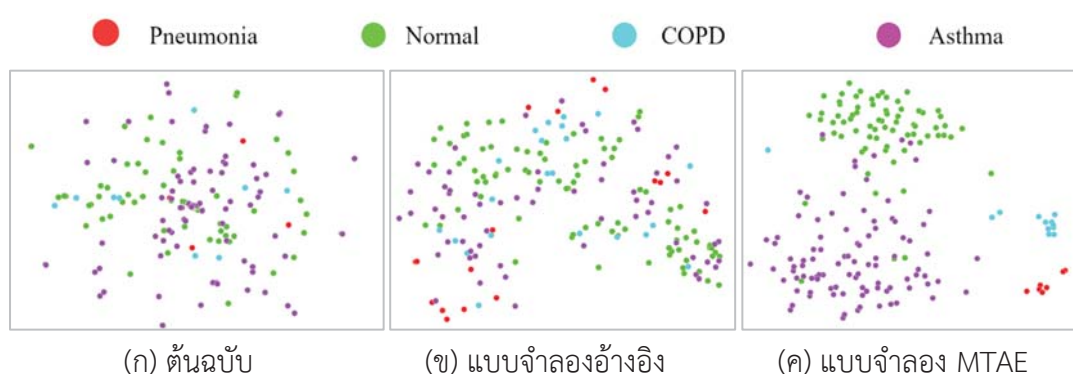
เพื่อให้การประเมินผลมีความครอบคลุมมากยิ่งขึ้น งานวิจัยนี้รายงานทั้งค่าตัวชี้วัดรายกลุ่มและค่าเฉลี่ยรวม โดยให้ความสำคัญกับค่าคะแนน F1 เนื่องจากเป็นตัวชี้วัดที่สะท้อนสมดุล

ระหว่างความแม่นยำ และความไวได้ดีกว่าค่าความถูกต้องเพียงอย่างเดียว โดยเฉพาะในกรณีที่จำนวนตัวอย่างของแต่ละกลุ่มไม่สมดุล นอกจากนี้ เมตริกซ์สับสนยังช่วยสนับสนุนการตีความเชิงคลินิก โดยทำให้เห็นรูปแบบความสัมพันธ์ระหว่างกลุ่มโรค และเป็นข้อมูลสำคัญสำหรับการปรับปรุงแบบจำลอง (Powers, 2011; Sokolova and Lapalme, 2009)

3.4.2 การวิเคราะห์เชิงภาพด้วยเทคนิค t-SNE

นอกเหนือจากการประเมินประสิทธิภาพของแบบจำลองด้วยตัวชี้วัดเชิงปริมาณและเมตริกซ์สับสนแล้ว งานวิจัยนี้ยังใช้การวิเคราะห์เชิงภาพด้วยเทคนิค t-SNE เพื่อแสดงลักษณะการกระจายตัวของข้อมูลและคุณลักษณะเชิงลึกหลังผ่านกระบวนการสกัดคุณลักษณะหรือการเรียนรู้ตัวแทนเชิงแฝง เทคนิค t-SNE เป็นวิธีการลดมิติข้อมูล ซึ่งสามารถแปลงข้อมูลจากปริภูมิมิติสูงสู่ปริภูมิสองมิติได้ โดยยังคงโครงสร้างของข้อมูลไว้ ทำให้สามารถสังเกตแนวโน้มการรวมกลุ่มและการแยกกลุ่มของข้อมูลได้อย่างชัดเจน (van der Maaten, L., and Hinton, G. (2008))

ในงานวิจัยนี้ t-SNE ถูกนำมาใช้เพื่อสนับสนุนการอธิบายความสามารถของแบบจำลองในการจำแนกประเภทของโรค โดยตัวอย่างผลการแสดงภาพ แสดงไว้ในรูปที่ 3.3 ซึ่งเป็นผลลัพธ์จากตัวอย่าง 1 ส่วน ของกระบวนการตรวจสอบความถูกต้องแบบไขว้ทำส่วน ทำการเปรียบเทียบการกระจายตัวของข้อมูลจาก 3 กรณี ได้แก่ ข้อมูลต้นฉบับ (Raw features) คุณลักษณะที่ได้จากแบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชันซึ่งใช้เป็นแบบจำลองอ้างอิง และตัวแทนเชิงแฝงที่ได้จากแบบจำลองออโตเอนโคเดอร์แบบหลายงาน



รูปที่ 3.4 ตัวอย่างการแสดงผลภาพด้วยเทคนิค t-SNE

จากรูปที่ 3.4 สามารถตีความสัญลักษณ์กลุ่ม ได้ดังนี้ สีแดง หมายถึง โรคปอดอักเสบ (Pneumonia), สีเขียว หมายถึง ปกติ (Normal), สีฟ้า หมายถึง โรคปอดอุดกั้นเรื้อรัง (COPD) และ สีบานเย็น หมายถึง โรคหอบหืด (Asthma) ทั้งนี้ ผลการแสดงผลภาพด้วยเทคนิค t-SNE ถูกใช้เพื่อสนับสนุนข้อสังเกตเชิงคุณภาพควบคู่กับผลการประเมินเชิงปริมาณ และไม่ได้ถูกนำมาใช้เป็นเกณฑ์หลักในการตัดสินประสิทธิภาพของแบบจำลอง

สังเกตได้ว่า การกระจายตัวของข้อมูลในกรณีอื่นพุตต้นฉบับมีการซ้อนทับกันของกลุ่มค่อนข้างมาก แสดงให้เห็นถึงความยากในการแยกประเภทของโรคโดยอาศัยคุณลักษณะดั้งเดิมเพียงอย่างเดียว ขณะที่กรณีของแบบจำลองอ้างอิง เริ่มแสดงแนวโน้มการรวมกลุ่มของข้อมูลในบางคลาส แต่ยังคงมีความสับสนระหว่างกลุ่มที่มีลักษณะใกล้เคียงกัน ในทางตรงกันข้าม ตัวแทนเชิงแฝงที่ได้จากแบบจำลองออโตเอนโคเดอร์แบบหลายงาน แสดงให้เห็นถึงการแยกกลุ่มของข้อมูลในแต่ละโรคได้ชัดเจนมากขึ้น ซึ่งสะท้อนถึงความสามารถของการเรียนรู้หลายงานในการสร้างตัวแทนข้อมูลที่มีโครงสร้างและความจำแนกได้ดีกว่า