

บทที่ 4

ผลการวิจัยและอภิปรายผล

ในบทนี้นำเสนอผลจากการดำเนินงานวิจัย ทั้งผลการกำหนดพารามิเตอร์ การสกัดคุณลักษณะ การปรับแต่งพารามิเตอร์การฝึกของแบบจำลอง การออกแบบสถาปัตยกรรมของแบบจำลองการเรียนรู้หลายงานแบบอัตโนมัติแบบหลายงาน และการประยุกต์ใช้ตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน เพื่อเพิ่มความสามารถในการจำแนกของโรคที่เกี่ยวข้องกับระบบทางเดินหายใจจากเสียงปอด ตลอดจนการเปรียบเทียบประสิทธิภาพกับแบบจำลองอ้างอิง

4.1 ผลการดำเนินงานวิจัยในส่วนของการหาความยาวของเสียง

การกำหนดความยาวเฟรม สำหรับการสกัดคุณลักษณะของสัมประสิทธิ์เซปสตรัมความถี่เมลเป็นปัจจัยสำคัญต่อคุณภาพของคุณลักษณะเชิงเวลาและความถี่ ที่ป้อนเข้าแบบจำลองโดยเฉพาะสำหรับสัญญาณเสียงปอดซึ่งมีลักษณะไม่คงที่และมีองค์ประกอบความถี่ที่ซับซ้อน ความยาวเฟรมที่สั้นเกินไปอาจทำให้สัญญาณมีความแปรปรวนสูงและเกิดความไวต่อสัญญาณรบกวน ขณะที่ความยาวเฟรมที่ยาวเกินไปจะลดความสามารถในการเก็บรายละเอียดการเปลี่ยนแปลงเชิงเวลา ส่งผลให้ความสามารถในการจำแนกลดลง

เป็นสิ่งที่สำคัญที่จะระบุความยาวของเฟรมให้เหมาะสมของสัมประสิทธิ์เซปสตรัมความถี่เมล โดยทำการตั้งค่าเบื้องต้น ประกอบด้วยตัวเข้ารหัส 3 ชั้น และตัวถอดรหัส 3 ชั้น และชั้นเชื่อมโยงสมบูรณ์ จำนวน 3 ชั้น อัตราการเรียนรู้ (Learning rate: LR) ถูกตั้งค่าเริ่มต้นเป็น 0.00001 ผลการทดลองแสดงให้เห็นผลกระทบของความยาวของเฟรมสัมประสิทธิ์เซปสตรัมความถี่เมลที่แตกต่างกันว่ามีผลต่อประสิทธิภาพของแบบจำลอง ดังตารางที่ 4.1 การศึกษาของสัมประสิทธิ์เซปสตรัมความถี่เมลกับระยะเวลาของเฟรม 20 มิลลิวินาที ในแบบจำลองแสดงให้เห็นถึงประสิทธิภาพสูงสุดในการจำแนกประเภทของโรคโดยใช้สัญญาณเสียงปอด ระยะเวลาเฟรมจึงเป็นสิ่งสำคัญในการศึกษาของสัญญาณเสียงปอด เพราะช่วยในการเก็บรวบรวมข้อมูลที่ถูกต้องของคุณสมบัติเสียง รายละเอียดที่จำเป็นสำหรับการแยกเสียงภาวะปกติกับเสียงที่มีความภาวะผิดปกติ ในทางตรงกันข้ามความยาว

เฟรมที่ยาวกว่ามีแนวโน้มที่จะทำให้คุณลักษณะเฉพาะลดลง ส่งผลให้ความถูกต้องลดลง การทดลองนี้สอดคล้องกับการศึกษาของงานวิจัยที่มีอยู่ก่อนหน้านี้ซึ่งเน้นความสำคัญระยะเฟรมในการศึกษาของสัญญาณเสียงพูด

ตารางที่ 4.1 ผลของความยาวเฟรมของแบบจำลองออโตเอนโคเดอร์แบบหลายงาน

ความยาวเฟรม (มิลลิวินาที)	ความถูกต้อง (เปอร์เซ็นต์)	ความแม่นยำ (เปอร์เซ็นต์)	ความไว (เปอร์เซ็นต์)	ค่าคะแนน F1 (เปอร์เซ็นต์)
20	71.28 ± 3.30	74.40 ± 9.21	52.60 ± 6.11	54.60 ± 5.13
40	61.17 ± 5.10	47.60 ± 16.58	41.80 ± 5.50	40.20 ± 6.76
128	56.60 ± 1.71	36.80 ± 11.34	35.40 ± 2.30	32.60 ± 2.70
256	55.75 ± 4.62	33.00 ± 12.02	33.80 ± 2.28	31.20 ± 2.95
512	50.74 ± 3.21	25.40 ± 2.61	30.80 ± 1.48	27.40 ± 1.52
1000	51.38 ± 4.71	25.40 ± 2.51	31.40 ± 2.51	27.80 ± 2.49

หมายเหตุ $\pm$ ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD)

ผลการทดลองในตารางที่ 4.1 ความยาวเฟรม 20 มิลลิวินาที ให้ประสิทธิภาพสูงสุด เมื่อเทียบกับความยาวเฟรมอื่น ๆ โดยมีค่าความถูกต้องเท่ากับ 71.28 ± 3.30 เปอร์เซ็นต์ และค่าความแม่นยำเท่ากับ 74.40 ± 9.21 เปอร์เซ็นต์ ซึ่งสูงที่สุดในชุดการทดลองนี้ ขณะเดียวกันค่าความไวเท่ากับ 52.60 ± 6.11 เปอร์เซ็นต์ และค่าคะแนน F1 เท่ากับ 54.60 ± 5.13 เปอร์เซ็นต์ แสดงให้เห็นว่าความยาวเฟรม 20 มิลลิวินาที สามารถรักษาสมดุลระหว่างการตรวจจับตัวอย่างภาวะผิดปกติ และความแม่นยำของการทำนายได้ดีที่สุดภายใต้เงื่อนไขการทดลอง

เมื่อเพิ่มความยาวเฟรมเป็น 40 มิลลิวินาที พบว่าประสิทธิภาพลดลงอย่างมีนัยสำคัญ โดยความถูกต้องลดลงเหลือ 61.17 ± 5.10 เปอร์เซ็นต์ และค่าคะแนน F1 ลดลงเหลือ 40.20 ± 6.76 เปอร์เซ็นต์ และเมื่อใช้ความยาวเฟรมที่ยาวมากขึ้น (128 ถึง 1000 มิลลิวินาที) ค่าตัวชี้วัดลดลงต่อเนื่อง เช่น ที่ 512 มิลลิวินาที ค่าความถูกต้องลดลงเหลือ 50.74 ± 3.21 เปอร์เซ็นต์ และค่าคะแนน F1 ลดลงเหลือ 27.40 ± 1.52 เปอร์เซ็นต์ แสดงให้เห็นว่าเฟรมที่ยาวทำให้คุณลักษณะจำเพาะของเสียงพูดถูกเฉลี่ยหรือเจือจาง ส่งผลให้แบบจำลองจำแนกกลุ่มได้ยากขึ้น ดังนั้นเพื่อให้ได้คุณลักษณะ

สัมประสิทธิ์สหสัมพันธ์ความถี่ที่เหมาะสมสำหรับการวินิจฉัยโรคระบบทางเดินหายใจจากเสียงปอด งานวิจัยนี้จึงเลือกใช้ความยาวเฟรม 20 มิลลิวินาที เป็นพารามิเตอร์มาตรฐานสำหรับการทดลอง

4.2 ผลการดำเนินงานวิจัยในส่วนของการเลือกอัตราการเรียนรู้

อัตราการเรียนรู้ เป็นพารามิเตอร์หลักที่ควบคุมขนาดการปรับปรุงค่าน้ำหนักในระหว่างการฝึก หากอัตราการเรียนรู้สูงเกินไปอาจทำให้การปรับค่าน้ำหนักข้ามจุดเหมาะสม และไม่สามารถเข้าหาคำตอบที่เสถียรได้ ขณะที่อัตราการเรียนรู้ต่ำเกินไปจะทำให้การฝึกช้าและอาจติดอยู่ในจุดที่ต่ำกว่าค่าที่เหมาะสมที่สุด (Sub-optimal) การศึกษาการใช้อัตราการเรียนรู้หลายค่า และผลลัพธ์ที่ได้ นำเสนอในตารางที่ 4.2

ตารางที่ 4.2 ผลการทดลองของอัตราการเรียนรู้

อัตราการเรียนรู้	ความถูกต้อง (เปอร์เซ็นต์)	ความแม่นยำ (เปอร์เซ็นต์)	ความไว (เปอร์เซ็นต์)	ค่าคะแนน F1 (เปอร์เซ็นต์)
0.01	44.15 ± 3.63	11.00 ± 0.71	25.00 ± 0.00	15.40 ± 0.89
0.001	75.96 ± 10.33	68.60 ± 20.50	67.20 ± 13.94	67.00 ± 17.59
0.0001	81.70 ± 3.44	80.20 ± 5.93	77.20 ± 3.56	78.00 ± 4.30
0.00001	71.28 ± 3.03	74.40 ± 9.21	55.60 ± 5.50	59.20 ± 7.33
0.000001	46.17 ± 5.30	18.40 ± 7.40	27.60 ± 3.21	21.40 ± 6.11

หมายเหตุ ± ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD)

ผลการทดลองในตารางที่ 4.2 แสดงให้เห็นว่า อัตราการเรียนรู้ เท่ากับ 0.0001 ให้ผลลัพธ์ดีที่สุด โดยมีค่าความถูกต้องเท่ากับ 81.70 ± 3.44 เปอร์เซ็นต์ และค่าความแม่นยำ ค่าความไว และค่าคะแนน F1 เท่ากับ 80.20 ± 5.93 เปอร์เซ็นต์, 77.20 ± 3.56 เปอร์เซ็นต์ และ 78.00 ± 4.30 เปอร์เซ็นต์ ตามลำดับ ซึ่งสะท้อนการเรียนรู้ที่มีเสถียรภาพและให้ผลสมมูลมากที่สุดในทุกตัวชี้วัด เมื่อใช้อัตราการเรียนรู้ เท่ากับ 0.001 แม้ค่าความถูกต้องอยู่ที่ 75.96 ± 10.33 เปอร์เซ็นต์ และค่าคะแนน F1 อยู่ที่ 67.00 ± 17.59 เปอร์เซ็นต์ แต่มีส่วนเบี่ยงเบนมาตรฐานสูงมาก แสดงถึงความไม่เสถียรระหว่างพบกับจำนวนยุคการฝึก ในขณะที่อัตราการเรียนรู้ เท่ากับ 0.01 ให้ผลลัพธ์ต่ำสุดอย่างชัดเจน (ความถูกต้อง 44.15 ± 3.63 เปอร์เซ็นต์, ค่าคะแนน F1 15.40 ± 0.89 เปอร์เซ็นต์) สอดคล้องกับ

ภาวะการปรับค่าเกินจุดเหมาะสม (Over-shooting) และการฝึกที่ไม่เข้าหาอย่างเหมาะสม ส่วนอัตราการเรียนรู้ที่มีค่าต่ำมาก เช่น 0.000001 ทำให้ค่าความถูกต้อง ลดลงเหลือ 46.17 ± 5.30 เปอร์เซ็นต์ และค่าคะแนน F1 21.40 ± 6.11 เปอร์เซ็นต์ สะท้อนว่าแบบจำลองอัปเดตน้ำหนักช้าเกินไปจนไม่สามารถเรียนรู้รูปแบบจำแนกได้อย่างมีประสิทธิภาพ จากผลการวิเคราะห์นี้จึงเลือกใช้ อัตราการเรียนรู้ เท่ากับ 0.0001 สำหรับการฝึกแบบจำลองในการทดลองขั้นตอนต่อไป

4.3 ผลการดำเนินงานวิจัยในส่วนการกำหนดค่าการเข้ารหัสและการถอดรหัส

ผลการดำเนินงานวิจัยในส่วนของการกำหนดสถาปัตยกรรมการเข้ารหัสและการถอดรหัสของแบบจำลองออโตเอนโคเดอร์แบบหลายงาน แสดงให้เห็นว่า จำนวนชั้นของตัวเข้ารหัส และตัวถอดรหัส ส่งผลโดยตรงต่อความสามารถของแบบจำลองในการเรียนรู้ตัวแทนเชิงแฝง และความสามารถในการทำให้คุณลักษณะสามารถใช้ทั่วไปได้กับข้อมูลที่ไม่เคยปรากฏในชุดฝึก หากสถาปัตยกรรมมีความตื้นเกินไป แบบจำลองอาจไม่สามารถจับคุณลักษณะสำคัญของสัญญาณได้อย่างครบถ้วน ในทางกลับกัน หากสถาปัตยกรรมมีความลึกมากเกินไป อาจนำไปสู่ปัญหาการเรียนรู้เกิน และลดประสิทธิภาพในการใช้งานทั่วไปของแบบจำลอง

นอกจากนี้ ประสิทธิภาพของแบบจำลองออโตเอนโคเดอร์แบบหลายงานยังได้รับอิทธิพลอย่างมีนัยสำคัญจากการออกแบบสถาปัตยกรรมในส่วนของหัวงานการจำแนก (Classification head) โดยเฉพาะจำนวนชั้นของชั้นเชื่อมโยงสมบูรณ์ ทำหน้าที่แปลงตัวแทนเชิงแฝงไปสู่การตัดสินใจเชิงกลุ่ม การกำหนดจำนวนชั้น และโครงสร้างของส่วนนี้ที่เหมาะสมมีความสำคัญต่อการรักษาสมดุลระหว่างความสามารถในการแยกกลุ่ม และความเสถียรของแบบจำลอง

เพื่อศึกษาผลกระทบของสถาปัตยกรรมต่อประสิทธิภาพ ทำการทดลองโดยปรับเปลี่ยนจำนวนชั้นและการจัดเรียงของบล็อกการเข้ารหัสและการถอดรหัส รวมถึงจำนวนชั้นของชั้นเชื่อมโยงสมบูรณ์ในส่วนการจำแนก จากนั้นจึงประเมินผลที่ได้โดยอาศัยค่าความถูกต้องของการจำแนกเป็นตัวชี้วัดหลัก ผลการทดลองที่สรุปไว้ในตารางที่ 4.3 แสดงให้เห็นแนวโน้มที่ชัดเจนของผลกระทบของการออกแบบแบบจำลอง

ตารางที่ 4.3 ประสิทธิภาพของการกำหนดชั้นการเข้ารหัสและการถอดรหัส

เข้ารหัส-ถอดรหัส	ความถูกต้อง (เปอร์เซ็นต์)	ความแม่นยำ (เปอร์เซ็นต์)	ความไว (เปอร์เซ็นต์)	ค่าคะแนน F1 (เปอร์เซ็นต์)
1-1	88.51 ± 7.10	91.60 ± 5.13	89.00 ± 5.52	89.80 ± 5.72
2-2	89.47 ± 4.21	92.05 ± 3.15	88.72 ± 6.63	89.96 ± 5.07
3-3	81.70 ± 3.44	80.20 ± 5.93	77.20 ± 3.56	78.00 ± 4.30
4-4	72.87 ± 10.21	69.60 ± 14.84	63.80 ± 12.03	64.60 ± 12.28
5-5	71.17 ± 9.28	66.60 ± 9.66	65.60 ± 13.58	65.40 ± 11.63

หมายเหตุ ± ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD)

ในตารางที่ 4.3 แสดงว่าแบบจำลองที่มีโครงสร้างแบบสมมาตร 2-2 ให้สมรรถนะที่ดีที่สุด โดยมีค่าความถูกต้อง เท่ากับ 89.47 ± 4.21 เปอร์เซ็นต์ และค่าความแม่นยำ ค่าความไว และค่าคะแนน F1 เท่ากับ 92.05 ± 3.15 เปอร์เซ็นต์, 88.72 ± 6.63 เปอร์เซ็นต์ และ 89.96 ± 5.07 เปอร์เซ็นต์ ตามลำดับ ซึ่งเป็นค่าที่สูงที่สุดหรือใกล้เคียงค่าสูงสุดในทุกตัวชี้วัด โครงสร้าง 1-1 ให้ค่าความถูกต้องสูงใกล้เคียง (88.51 ± 7.10 เปอร์เซ็นต์) แต่มีส่วนเบี่ยงเบนมาตรฐานสูงกว่า และค่าเฉลี่ยโดยรวมน้อยกว่า 2-2 แสดงให้เห็นว่า 1-1 ยังไม่เพียงพอสำหรับการเรียนรู้คุณลักษณะที่ซับซ้อนของเสียงปอดในทุกกลุ่ม เมื่อเพิ่มจำนวนชั้นความลึกเป็น 3-3, 4-4 และ 5-5 พบว่าค่าความถูกต้องลดลงอย่างต่อเนื่อง จาก 81.70 เปอร์เซ็นต์ เป็น 72.87 เปอร์เซ็นต์ และ 71.17 เปอร์เซ็นต์ ตามลำดับ พร้อมกับค่าคะแนน F1 ลดลงจาก 78.00 เปอร์เซ็นต์ เป็น 64.60 เปอร์เซ็นต์ และ 65.40 เปอร์เซ็นต์ ตามลำดับ แสดงให้เห็นว่า การเพิ่มความซับซ้อนของแบบจำลองไม่ได้ทำให้ดีขึ้นเสมอไป และมีแนวโน้มเข้าสู่ภาวะการเรียนรู้เกิน หรือเพิ่มภาระการเรียนรู้จนทำให้แบบจำลองไม่เสถียร โครงสร้างการเข้ารหัสและการถอดรหัส แบบสมมาตร 2-2 เป็นสถาปัตยกรรมที่เหมาะสม

4.4 ผลการดำเนินงานวิจัยในส่วนของจำนวนชั้นเชื่อมโยงสมบูรณ์

ชั้นเชื่อมโยงสมบูรณ์ ทำหน้าที่แปลงคุณลักษณะเชิงแฝงจากพื้นที่เชิงแฝง (Latent space) ไปสู่การตัดสินใจจำแนกกลุ่ม จำนวนชั้นเชื่อมโยงสมบูรณ์มากเกินไปอาจเพิ่มพารามิเตอร์จนเกิดการเรียนรู้ที่มากเกินไป ขณะที่น้อยเกินไปอาจไม่เพียงพอในการสร้างขอบเขตการจำแนกที่ซับซ้อน

ผลการทดลองในตารางที่ 4.4 แสดงว่า ชั้นเชื่อมโยงสมบูรณ์ จำนวน 3 ชั้น ให้ผลดีที่สุด โดยมีค่าความถูกต้อง เท่ากับ 89.47 ± 4.21 เปอร์เซ็นต์ และค่าคะแนน F1 เท่ากับ 89.96 ± 5.07 เปอร์เซ็นต์ ซึ่งเป็นค่าที่สูงที่สุดในชุดนี้ แสดงให้เห็นว่า 3 ชั้น สามารถเพิ่มความสามารถในการสร้างขอบเขตการจำแนกโดยไม่ทำให้แบบจำลองซับซ้อนเกินไป เมื่อใช้ชั้นเชื่อมโยงสมบูรณ์ จำนวน 2 ชั้น ค่าความถูกต้องอยู่ที่ 87.55 ± 1.50 เปอร์เซ็นต์ ซึ่งน้อยกว่า 3 ชั้น และเมื่อใช้ชั้นเชื่อมโยงสมบูรณ์ จำนวน 1 ชั้น ค่าความถูกต้องลดลงเป็น 85.84 ± 5.79 เปอร์เซ็นต์ สะท้อนว่าโครงสร้างที่ตื้นเกินไปจำกัดความสามารถในการแยกกลุ่ม ในทางตรงกันข้ามเมื่อเพิ่มชั้นเป็น 4 และ 5 ชั้น ค่าความถูกต้องลดลงเป็น 85.70 ± 4.86 เปอร์เซ็นต์ และ 86.38 ± 2.15 เปอร์เซ็นต์ ตามลำดับ โดยค่าคะแนน F1 อยู่ที่ 83.80 เปอร์เซ็นต์ ทั้งสองกรณี แสดงแนวโน้มว่าการเพิ่มชั้นมากเกินไปก่อให้เกิดประโยชน์เพิ่ม และอาจทำให้แบบจำลองเริ่มเรียนรู้มากเกินไป หรือมีความไม่เสถียร ดังนั้นงานวิจัยจึงเลือกชั้นเชื่อมโยงสมบูรณ์ จำนวน 3 ชั้น เป็นค่าที่เหมาะสมสำหรับแบบจำลอง

ตารางที่ 4.4 ประสิทธิภาพของชั้นเชื่อมโยงสมบูรณ์

ชั้นเชื่อมโยง สมบูรณ์	ความถูกต้อง (เปอร์เซ็นต์)	ความแม่นยำ (เปอร์เซ็นต์)	ความไว (เปอร์เซ็นต์)	ค่าคะแนน F1 (เปอร์เซ็นต์)
1	85.84 ± 5.79	91.20 ± 2.77	84.60 ± 7.50	86.80 ± 5.36
2	87.55 ± 1.50	89.40 ± 2.88	85.80 ± 7.43	86.80 ± 4.44
3	89.47 ± 4.21	92.05 ± 3.15	88.72 ± 6.63	89.96 ± 5.07
4	85.70 ± 4.86	86.80 ± 5.12	84.20 ± 7.89	83.80 ± 9.18
5	86.38 ± 2.15	85.80 ± 3.63	82.20 ± 1.79	83.80 ± 1.48

หมายเหตุ $\pm$ ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD)

4.5 การเปรียบเทียบแบบจำลอง

ผลการทดลองเปรียบเทียบประสิทธิภาพการจำแนกแบบ 4 กลุ่ม ระหว่างแบบจำลองอ้างอิงโครงข่ายประสาทเทียมแบบคอนโวลูชันสองชั้น กับแบบจำลองออโตเอนโคเดอร์แบบหลายงานซึ่งเป็นวิธีที่นำเสนอ พบว่าแบบจำลองที่นำเสนอให้ผลที่ดีกว่าอย่างชัดเจนในตัวชี้วัด ได้แก่ ความถูกต้อง ความแม่นยำ ความไว และค่าคะแนน F1 ดังตารางที่ 4.5

ตารางที่ 4.5 ผลการทดลองแบบจำลอง

แบบจำลอง	ความถูกต้อง (เปอร์เซ็นต์)	ความแม่นยำ (เปอร์เซ็นต์)	ความไว (เปอร์เซ็นต์)	ค่าคะแนน F1 (เปอร์เซ็นต์)
CNN-2L (Baseline)	70.64 ± 8.66	70.60 ± 10.43	65.40 ± 8.44	65.20 ± 8.70
MTAE	89.47 ± 4.21	92.05 ± 3.15	88.72 ± 6.63	89.96 ± 5.07

หมายเหตุ ± ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD)

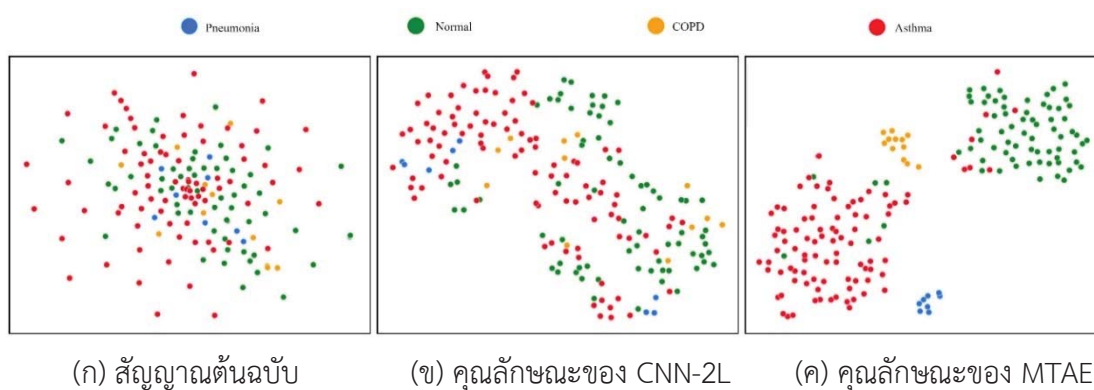
โดยโครงข่ายประสาทเทียมแบบคอนโวลูชันสองชั้น ให้ค่าความถูกต้อง เท่ากับ 70.64 ± 8.66 เปอร์เซ็นต์ และมีค่าความแม่นยำ ค่าความไว และค่าคะแนน F1 เท่ากับ 70.60 ± 10.43 เปอร์เซ็นต์, 65.40 ± 8.44 เปอร์เซ็นต์ และ 65.20 ± 8.70 เปอร์เซ็นต์ ตามลำดับ แสดงว่าแบบจำลองอ้างอิงยังมีข้อจำกัดด้านความสามารถในการจำแนกกลุ่ม โดยเฉพาะเมื่อพิจารณาค่าความไว และค่าคะแนน F1 ที่อยู่ในระดับปานกลาง ซึ่งบ่งชี้ถึงความผิดพลาดในการจำแนกและความสมดุลของการจำแนกที่ยังไม่สูง ในทางตรงกันข้าม แบบจำลองที่นำเสนอ ให้ค่าความถูกต้องสูงถึง 89.47 ± 4.21 เปอร์เซ็นต์ และมีค่าความแม่นยำ ค่าความไว และค่าคะแนน F1 เท่ากับ 92.05 ± 3.15 เปอร์เซ็นต์ 88.72 ± 6.63 เปอร์เซ็นต์ และ 89.96 ± 5.07 เปอร์เซ็นต์ ตามลำดับ แสดงให้เห็นว่าแบบจำลองที่นำเสนอ สามารถเพิ่มทั้งความแม่นยำในการทำนาย และความไวในการตรวจพบกลุ่มได้พร้อมกัน ส่งผลให้ค่าคะแนน F1 สูงขึ้นอย่างมีนัยสำคัญเมื่อเทียบกับแบบจำลองอ้างอิง

นอกจากนี้ค่าส่วนเบี่ยงเบนมาตรฐานของแบบจำลองที่นำเสนอ มีแนวโน้มต่ำกว่าแบบจำลองอ้างอิง ในหลายตัวชี้วัด โดยเฉพาะค่าความถูกต้อง และค่าความแม่นยำ ซึ่งแสดงว่ามีความเสถียรในการทำนายมากกว่าเมื่อประเมินแบบการตรวจสอบไขว้ (Cross-validation) สนับสนุนว่าแนวทางการเรียนรู้ตัวแทนเชิงแฝงแบบหลายงานของแบบจำลองออโตเอนโคเดอร์แบบหลายงาน ช่วยยกระดับประสิทธิภาพการจำแนกโรกระบบทางเดินหายใจจากเสียงปอดได้สูงกว่าแบบจำลองอ้างอิง ทั้งค่าความถูกต้องและความสมดุลของตัวชี้วัดรายกลุ่ม

หลังจากสรุปค่าพารามิเตอร์ที่เหมาะสมกับงานวิจัย กำหนดการทดลองหลักของแบบจำลองออโตเอนโคเดอร์แบบหลายงาน โดยใช้อัตราการเรียนรู้ เท่ากับ 0.0001 โครงสร้างเข้ารหัสและถอดรหัส 2-2 และชั้นเชื่อมโยงสมบูรณ์ จำนวน 3 ชั้น วิเคราะห์ความสามารถในการจำแนกกลุ่มของ

คุณลักษณะด้วย t-SNE เพื่อเปรียบเทียบ (ก) สัญญาณเสียงต้นฉบับ (ข) คุณลักษณะจาก CNN-2L และ (ค) คุณลักษณะจาก MTAE

แสดงการลดมิติของข้อมูลด้วยเทคนิค t-SNE เปรียบเทียบการกระจายตัวของข้อมูล 4 กลุ่ม โดยใช้ตัวอย่างจาก 1 ใน 5 ส่วน การวิเคราะห์ภาพดังกล่าวสามารถตีความได้ในเชิงคุณภาพของตัวแทนคุณลักษณะ กล่าวคือ หากตัวแทนคุณลักษณะสามารถจำแนกกลุ่มได้ดี จะปรากฏเป็นกลุ่มที่มีความหนาแน่นภายในกลุ่มสูง และมีระยะห่างระหว่างกลุ่มมากขึ้น ขณะที่การทับซ้อนกันของกลุ่มสะท้อนความยากในการสร้างขอบเขตจำแนก



รูปที่ 4.1 การกระจายภาพ t-SNE ของคุณลักษณะแบบต่าง ๆ

สัญญาณต้นฉบับ การกระจายตัวของข้อมูลสัญญาณต้นฉบับมีลักษณะปะปนและทับซ้อนกันสูงระหว่างกลุ่ม ซึ่งแสดงว่าในสัญญาณเสียงปอดต้นฉบับมีความแปรปรวนมากจากหลายปัจจัย เช่น ระดับความดัง ความยาวรอบการหายใจ สัญญาณรบกวน และความแตกต่างระหว่างบุคคล ส่งผลให้รูปแบบเชิงโครงสร้างที่จำแนกโรคแต่ละชนิดยังไม่เด่นชัด การที่กลุ่มยังจำแนกไม่ออกในภาพนี้จึงเป็นเหตุผลให้เห็นว่าจำเป็นต้องมีขั้นตอนการสกัดคุณลักษณะหรือการเรียนรู้ตัวแทนที่เหมาะสมก่อนการจำแนก

คุณลักษณะจาก CNN-2L เมื่อพิจารณาคุณลักษณะที่สกัดโดยแบบจำลองอ้างอิง พบว่าเกิดการรวมกลุ่มของข้อมูลมากขึ้นเมื่อเทียบกับสัญญาณต้นฉบับ มีบางกลุ่มที่เริ่มจับตัวเป็นกลุ่ม และมีทิศทางในการจำแนกจากกลุ่มอื่นดีขึ้น มีพื้นที่รอยต่อ (Boundary overlap) ซึ่งตีความได้ว่าแบบจำลองอ้างอิงสามารถเรียนรู้คุณลักษณะพื้นฐานของสเปกตรัมได้บางส่วน แต่ยังไม่สามารถดึงคุณลักษณะ

จำเพาะที่จำแนกกลุ่มที่คล้ายกันได้อย่างชัดเจน ผลภาพนี้สอดคล้องกับผลตัวเลขในตารางที่ 4.5 ที่แบบจำลองอ้างอิงให้ค่าความถูกต้องรวมต่ำ ซึ่งบ่งชี้ว่าคุณลักษณะที่ได้จากแบบจำลองอ้างอิงยังมีความสามารถในการจำแนกจำกัดเมื่อเจอกลุ่มข้อมูลที่มีลักษณะเสียงใกล้เคียงกัน

คุณลักษณะของ MTAE ที่สกัดจากแบบจำลองออโตเอนโคเดอร์แบบหลายงาน พบว่ากลุ่มของข้อมูลมีความชัดเจนขึ้นอย่างเห็นได้ชัด โดยแนวโน้มสำคัญ คือ ความกระชับภายในกลุ่มเพิ่มขึ้น (Intra-class compactness) และความแยกจากกันระหว่างกลุ่มมากขึ้น (Inter-class separation) เมื่อเทียบกับรูปที่ 4.1 (ข) ซึ่งแสดงว่าแบบจำลองสามารถเรียนรู้ตัวแทนที่จัดระเบียบเชิงโครงสร้างให้กลุ่มเดียวกันอยู่ใกล้กันมากขึ้น และผลกลุ่มต่างกันให้ออกห่างกันมากขึ้น

แบบจำลองออโตเอนโคเดอร์แบบหลายงาน มีลักษณะการเรียนรู้ที่ควบคุมให้แบบจำลองสร้างตัวแทนเชิงแฝงที่ดีทั้งต่อการสร้างใหม่ และต่อการจำแนกร่วมกัน ทำให้ตัวแทนที่เรียนรู้มีความทั่วไปและทนทานต่อความแปรปรวนของสัญญาณมากกว่าแบบจำลองที่เน้นจำแนกเพียงอย่างเดียว ผลภาพจากเทคนิค t-SNE จึงสนับสนุนข้อค้นพบว่าแบบจำลองออโตเอนโคเดอร์แบบหลายงานให้คุณลักษณะที่เหมาะสมต่อการจำแนกโรคทางเดินหายใจจากเสียงปอดได้ดีกว่า

โดยสรุปผลภาพเทคนิค t-SNE ชี้ให้เห็นว่า แบบจำลองออโตเอนโคเดอร์แบบหลายงานสามารถสร้างตัวแทนคุณลักษณะที่มีการรวมกลุ่มภายในกลุ่มมากขึ้นและเพิ่มการแยกจากกันระหว่างกลุ่มได้ดีกว่าแบบจำลองอ้างอิง ส่งผลให้การจำแนกโรคจากเสียงปอดทำได้แม่นยำขึ้น ทั้งนี้ผลเชิงภาพดังกล่าวมีความสอดคล้องกับผลเชิงตัวเลขจากตารางที่ 4.5 และรูปที่ 4.2 เมตริกซ์สับสนที่แสดงการลดความสับสนระหว่างกลุ่มโรค

4.6 ผลการดำเนินงานในส่วนการเรียนรู้หลายงานร่วมกับซัพพอร์ตเวกเตอร์แมชชีน

แบบจำลองออโตเอนโคเดอร์แบบหลายงานร่วมกับซัพพอร์ตเวกเตอร์แมชชีนทำหน้าที่เป็นตัวจำแนกขั้นสุดท้ายบนตัวแทนเชิงแฝงที่ได้จากการเรียนรู้หลายงาน พารามิเตอร์สำคัญของซัพพอร์ตเวกเตอร์แมชชีนที่มีผลต่อรูปแบบของขอบเขตการตัดสินใจ คือ พารามิเตอร์แกมมา ซึ่งควบคุมความกว้างของฟังก์ชันเคอร์เนลแบบเรเดียล และส่งผลโดยตรงต่อความซับซ้อนของไฮเปอร์เพลนที่ใช้ในการจำแนกข้อมูล

ด้วยเหตุผลเชิงทฤษฎีและหลักฐานจากงานวิจัยก่อนหน้านี้ งานวิจัยนี้จึงเลือกกำหนดค่าพารามิเตอร์แกมมา เท่ากับ 0.1 สำหรับซัพพอร์ตเวกเตอร์แมชชีน ตามแนวทางที่ได้รับการยอมรับใน

วรรณกรรม เพื่อหลีกเลี่ยงการเพิ่มความซับซ้อนของกระบวนการทดลองและเพื่อให้การประเมินประสิทธิภาพของวิธีการแบบจำลองอัตโนมัติแบบหลายงานร่วมกับชีพพอร์ตเวกเตอร์แมชชีน มุ่งเน้นไปที่คุณภาพของตัวแทนเชิงแฝงที่ได้จากการเรียนรู้หลายงานเป็นหลัก เมื่อใช้ค่าพารามิเตอร์แกมมาค่าดังกล่าว ชีพพอร์ตเวกเตอร์แมชชีนสามารถกำหนดขอบเขตการตัดสินใจที่มีความคมชัด โดยอาศัยการจำแนกกลุ่มโครงสร้างของตัวแทนเชิงแฝงจากอัตโนมัติแบบหลายงาน ดังที่แสดงในรูปที่ 4.2

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

(ก) แบบจำลอง CNN-2L

Actual label	Pneumonia	55	7	1	3
	Normal	0	385	5	25
	COPD	0	8	91	3
	Asthma	1	44	2	310
		Pneumonia	Normal	COPD	Asthma
Predicted label					

(ข) แบบจำลอง MTAE

Actual label	Pneumonia	60	4	0	2
	Normal	0	389	4	22
	COPD	0	4	91	7
	Asthma	1	34	2	320
		Pneumonia	Normal	COPD	Asthma
Predicted label					

(ค) แบบจำลอง MTAE-SVM

รูปที่ 4.2 เมตริกซ์สับสนของแบบจำลอง

ตารางที่ 4.6 ผลการประเมินประสิทธิภาพของแบบจำลองในการจำแนกโรคปอด

แบบจำลอง	ความถูกต้อง (เปอร์เซ็นต์)	ความแม่นยำ (เปอร์เซ็นต์)	ความไว (เปอร์เซ็นต์)	ค่าคะแนน F1 (เปอร์เซ็นต์)
CNN-2L	70.64 ± 8.66	70.60 ± 10.43	65.40 ± 8.44	65.20 ± 8.70
MTAE	89.47 ± 4.21	92.05 ± 3.15	88.72 ± 6.63	89.96 ± 5.07
MTAE-SVM	91.49 ± 3.03	93.45 ± 2.32	91.22 ± 4.17	92.13 ± 3.14

หมายเหตุ ± ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD)

ตารางที่ 4.6 แสดงผลการประเมินประสิทธิภาพของแบบจำลอง ซึ่งรายงานเป็นค่าเฉลี่ย พร้อมส่วนเบี่ยงเบนมาตรฐานจากการทดสอบแบบการตรวจสอบไขว้แบบห้าส่วน ผลการทดลองพบว่าแบบจำลองออโตเอนโคเดอร์แบบหลายงานร่วมกับซัพพอร์ตเวกเตอร์แมชชีนให้ค่าประสิทธิภาพสูงสุดในทุกตัวชี้วัด โดยมีค่าความถูกต้อง 91.49 ± 3.03 เปอร์เซ็นต์ และมีค่าความแปรปรวนอยู่ในระดับต่ำ สะท้อนถึงความถูกต้องและความเสถียรของแบบจำลองในการจำแนกข้อมูลหลายกลุ่มโรค

ผลดังกล่าวแสดงให้เห็นว่าการผสานตัวแทนเชิงแฝงจากโครงสร้างแบบจำลองออโตเอนโคเดอร์แบบหลายงานเข้ากับตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีนช่วยเพิ่มความสามารถในการจำแนกข้อมูลระหว่างกลุ่มได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม การประเมินผลในภาพรวมยังไม่สามารถแสดงสมรรถนะของแบบจำลองในแต่ละกลุ่มโรคได้อย่างครบถ้วน ดังนั้น ตารางที่ 4.7 จึงนำเสนอผลการเปรียบเทียบประสิทธิภาพในระดับรายกลุ่มโรค เพื่อให้เห็นพฤติกรรมของแบบจำลอง

ตารางที่ 4.7 การเปรียบเทียบผลการจำแนกโรคปอดรายกลุ่มโรค

วิธีการสกัด คุณลักษณะ และตัวจำแนก	ACC_{Normal} (เปอร์เซ็นต์)	$ACC_{Pneumonia}$ (เปอร์เซ็นต์)	ACC_{Asthma} (เปอร์เซ็นต์)	ACC_{COPD} (เปอร์เซ็นต์)	OA (เปอร์เซ็นต์)
MFCC and CNN-2L	78.50	62.50	67.40	52.44	70.64
MFCC and MTAE	92.77	83.33	86.83	89.22	89.47
MFCC and MTAE-SVM	93.73	90.91	89.64	89.22	91.49

จากตารางที่ 4.7 แสดงผลการเปรียบเทียบประสิทธิภาพของวิธีการสกัดคุณลักษณะ และตัวจำแนกในการจำแนกโรคปอดแบบแยกรายกลุ่ม โดยพิจารณาค่าความถูกต้องในการจำแนกของแต่ละกลุ่มโรค ได้แก่ กลุ่มภาวะปกติ (ACC_{Normal}) กลุ่มโรคปอดอักเสบ ($ACC_{Pneumonia}$) กลุ่มโรคหอบหืด (ACC_{Asthma}) และกลุ่มโรคปอดอุดกั้นเรื้อรัง (ACC_{COPD}) รวมถึงค่าความถูกต้องรวม (Overall Accuracy: OA) จากผลการทดลองพบว่าวิธีแบบจำลองออโตเอนโคเดอร์แบบหลายงานเข้ากับตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีนให้ค่าความถูกต้องรวมสูงสุดที่ 91.49 เปอร์เซ็นต์ และให้ผลการจำแนกที่ดีในทุกกลุ่มโรคเมื่อเทียบกับวิธีอื่น

เมื่อพิจารณาในระดับรายกลุ่มโรค พบว่าแบบจำลองออโตเอนโคเดอร์แบบหลายงานเข้ากับตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีนสามารถเพิ่มความถูกต้องในการจำแนกกลุ่มปอดอักเสบและกลุ่มหอบหืดได้อย่างชัดเจน โดยค่าความถูกต้องของกลุ่มโรคปอดอักเสบเพิ่มจาก 83.33 เปอร์เซ็นต์ เป็น 90.91 เปอร์เซ็นต์ และกลุ่มโรคหอบหืดเพิ่มจาก 86.83 เปอร์เซ็นต์ เป็น 89.64 เปอร์เซ็นต์ ขณะที่ค่าความถูกต้องรวมเพิ่มขึ้นจากแบบจำลองออโตเอนโคเดอร์แบบหลายงาน 89.47 เปอร์เซ็นต์ เป็น 91.49 เปอร์เซ็นต์ ผลลัพธ์ดังกล่าวสะท้อนให้เห็นว่าการใช้ซัพพอร์ตเวกเตอร์แมชชีนช่วยปรับขอบเขตการตัดสินใจเหมาะสมยิ่งขึ้นบนคุณลักษณะที่เรียนรู้จากแบบจำลองออโตเอนโคเดอร์แบบหลายงาน ส่งผลให้แบบจำลองสามารถแยกแยะโรคปอดแต่ละประเภทได้อย่างมีประสิทธิภาพมากขึ้น โดยเฉพาะในกลุ่มโรคที่มีลักษณะสัญญาณใกล้เคียงกัน

4.7 ผลการเปรียบเทียบกับแนวทางที่มีอยู่เดิมตามงานวิจัยอ้างอิง

หัวข้อนี้นำเสนอการเปรียบเทียบวิธีที่นำเสนอทั้งแบบสัมประสิทธิ์เซปสตรีมความถี่เมลร่วมกับแบบจำลองออโตเอนโคเดอร์แบบหลายงาน และสัมประสิทธิ์เซปสตรีมความถี่เมลร่วมกับแบบจำลองออโตเอนโคเดอร์แบบหลายงานผสานตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีนกับแนวทางที่มีอยู่เดิม ประเมินการจำแนก 4 กลุ่ม แสดงค่าความถูกต้องรายกลุ่มและความถูกต้องรวม เพื่อให้เห็นความสามารถในการแข่งขันของวิธีที่นำเสนอเมื่อเทียบกับงานวิจัยก่อนหน้า

ในตารางที่ 4.8 แนวทางที่นำเสนอให้ค่าความถูกต้องรวมสูงกว่าแนวทางที่มีอยู่เดิมแบบจำลองออโตเอนโคเดอร์แบบหลายงานที่มีกระบวนการสกัดคุณลักษณะสัมประสิทธิ์เซปสตรีมความถี่เมลมีค่าความถูกต้องรวม เท่ากับ 89.47 เปอร์เซ็นต์ และแบบจำลองออโตเอนโคเดอร์แบบหลายงานร่วมกับตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีนที่มีกระบวนการสกัดคุณลักษณะสัมประสิทธิ์เซปสตรีมความถี่เมลมีค่าความถูกต้องรวม เท่ากับ 91.49 เปอร์เซ็นต์ ซึ่งสูงกว่า โครงข่าย

ประสาทเทียมแบบคอนโวลูชันสองชั้นที่มีกระบวนการสกัดคุณลักษณะสัมประสิทธิ์เซปสตรัมความถี่ เมลมีค่าความถูกต้องรวม (Aykanat et al., 2017) เท่ากับ 70.64 เปอร์เซ็นต์ และสูงกว่าการแปลง เวฟเลตเชิงประจักษ์ (Empirical Wavelet Transform: EWT) ร่วมกับการประมวลผลด้วยชุดตัวกรอง (Filter Bank-based Processing: FBP) และการจำแนกด้วยไลต์กราเดียนต์บูสตี้งแมชชีน (Light Gradient Boosting Machine: LGBM) (Tripathy et al., 2022) ที่มีค่าความถูกต้องรวม เท่ากับ 77.13 เปอร์เซ็นต์ สะท้อนว่าแบบจำลองออโตเอนโคเดอร์แบบหลายงานสามารถสกัดคุณลักษณะที่เหมาะสมต่อการจำแนกโรคได้ดีกว่าวิธีแบบดั้งเดิม

ตารางที่ 4.8 การเปรียบเทียบผลการจำแนก 4 กลุ่มกับแนวทางที่มีอยู่เดิม

วิธีสกัด คุณลักษณะ และตัวจำแนก	ACC_{Normal} (เปอร์เซ็นต์)	$ACC_{Pneumonia}$ (เปอร์เซ็นต์)	ACC_{Asthma} (เปอร์เซ็นต์)	ACC_{COPD} (เปอร์เซ็นต์)	OA (เปอร์เซ็นต์)
MFCC+CNN-2L	78.50	62.50	67.40	52.44	70.64
EWT-FBP + LGBM	–	–	–	–	77.13
MFCC + MTAE	92.77	83.33	86.83	89.22	89.47
MFCC + MTAE- SVM	93.73	90.91	89.64	89.22	91.49

เมื่อเปรียบเทียบระหว่างแบบจำลองออโตเอนโคเดอร์แบบหลายงานกับแบบจำลองออโตเอนโคเดอร์แบบหลายงานร่วมกับตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีนพบว่า ค่าความถูกต้องรวมเพิ่มจาก 89.47 เปอร์เซ็นต์ เป็น 91.49 เปอร์เซ็นต์ (เพิ่มขึ้น 2.02 เปอร์เซ็นต์) ซึ่งสอดคล้องกับคำอธิบายในหัวข้อก่อนหน้านี้ว่าการใช้ตัวจำแนกซัพพอร์ตเวกเตอร์แมชชีน ช่วยเพิ่มความไวและความสามารถในการตัดสินใจบนคุณลักษณะเชิงแฝง โดยการเพิ่มระยะห่างระหว่างคุณลักษณะกับไฮเปอร์เพลนของซัพพอร์ตเวกเตอร์แมชชีน และให้ผลดีขึ้นกว่าการใช้ตัวจำแนกของแบบจำลองออโตเอนโคเดอร์แบบหลายงานเพียงอย่างเดียว (Pathonsuwan, W., Phapatanaburi, K., Buayai, P., Jumphoo, T., Anchuen, P., Uthansakul, M., and Uthansakul, P., 2022)

ในเชิงรายกลุ่มจุดที่เพิ่มเติมที่สุด คือ กลุ่มโรคปอดอักเสบ เพิ่มจาก 83.33 เปอร์เซ็นต์ (แบบจำลองออโตเอนโคเดอร์แบบหลายงาน) เป็น 90.91 เปอร์เซ็นต์ (แบบจำลองออโตเอนโคเดอร์

แบบหลายงานร่วมกับตัวจำแนกแบบซัพพอร์ทเวกเตอร์แมชชีน) หรือเพิ่ม 7.58 เปอร์เซนต์ แสดงให้เห็นว่าตัวจำแนกแบบซัพพอร์ทเวกเตอร์แมชชีน มีบทบาทสำคัญต่อการลดความสับสนในกลุ่มที่ทับซ้อนกับกลุ่มอื่น ขณะที่กลุ่มโรคปอดอุดกั้นเรื้อรัง มีค่าคงที่ 89.22 เปอร์เซนต์ ทั้งสองวิธี แสดงว่าความยากในการจำแนกบางกลุ่มอาจยังขึ้นกับลักษณะสัญญาณของโรคนั้น ๆ และควรพิจารณาร่วมกับเมตริกซ์สับสนเพื่ออธิบายคู่กลุ่มที่สับสน