



ระบบตรวจจับและพยากรณ์การรั่วของท่อลมในไลน์การผลิตท่อโดยใช้เทคโนโลยีเอ็ดจ์
คอมพิวเตอร์และการเรียนรู้เชิงลึก



วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์
ปีการศึกษา 2566
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

ระบบตรวจจับและพยากรณ์การรั่วของท่อลมในไลน์การผลิตท่อโดยใช้เทคโนโลยีเอ็ดจ์
คอมพิวเตอร์และการเรียนรู้เชิงลึก



วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์
ปีการศึกษา 2566
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

วิทยานิพนธ์ เรื่อง "ระบบตรวจจับและพยากรณ์การรั่วของท่อลมในไลน์การผลิตท่อโดยใช้เทคโนโลยี
เอ็ดจ์คอมพิวเตอร์และการเรียนรู้เชิงลึก"

ของ ธนกฤต ไกรสิงห์

ได้รับการพิจารณาให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการสอบวิทยานิพนธ์
(รองศาสตราจารย์ ดร.สาคร เมฆรักษาวิช)

..... ประธานที่ปรึกษาวิทยานิพนธ์
(ผู้ช่วยศาสตราจารย์ ดร.วินัย วงษ์ไทย)

..... กรรมการที่ปรึกษาวิทยานิพนธ์
(ดร.เอมอัสนา นรินตสุวรรณ์)

..... กรรมการผู้ทรงคุณวุฒิภายใน
(ผู้ช่วยศาสตราจารย์ ดร.ธนธร พ่อคำ)

..... กรรมการผู้ทรงคุณวุฒิภายใน
(ดร.ณัฐพล คุ่มใหญ่โต)

อนุมัติ

.....
(รองศาสตราจารย์ ดร.กรรองกาญจน์ ชูทิพย์)

คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง	ระบบตรวจจับและพยากรณ์การรั่วของท่อลมในไลน์การผลิตท่อโดยใช้เทคโนโลยีเอ็ดจ์คอมพิวเตอร์และการเรียนรู้เชิงลึก
ผู้วิจัย	ธนกฤต ไกรสิงห์
ประธานที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.วินัย วงษ์ไทย
กรรมการที่ปรึกษา	ดร.เอมอชนา นรินทร์สุขรัตน์
ประเภทสารนิพนธ์	วิทยานิพนธ์ วท.ม. วิทยาการคอมพิวเตอร์, มหาวิทยาลัยนเรศวร, 2566
คำสำคัญ	รั่วซึมของลมอัดอากาศ, การตรวจจับการรั่วซึมของท่อลม, เทคนิคการเรียนรู้เชิงลึก, ระบบท่อลม, ต้นไม้การตัดสินใจ, ป่าไม้สุ่ม, เครื่องมือเวกเตอร์สนับสนุน, 1D-CNN

บทคัดย่อ

การรั่วซึมของลมเกิดจากชิ้นส่วนของระบบท่อลมในเครื่องจักรที่ใช้ลมในการทำงานเกิดเสียหายหรือลัดลมเหลวได้ ข้อบกพร่องที่สำคัญนี้เป็นสาเหตุของการสูญเสียประมาณครึ่งหนึ่งของผลผลิตจากสายการผลิตในภาคการผลิตสิ่งของแต่ละสายงานการผลิต เทคนิคการเรียนรู้เชิงลึกสามารถนำมาใช้ในการตรวจจับการรั่วซึมของระบบท่อลมได้ ผู้วิจัยขอเสนอโมเดลการเรียนรู้เชิงลึกสำหรับการตรวจจับการรั่วซึมของชิ้นส่วนของระบบท่อลมพิเศษโดยใช้ระบบเซ็นเซอร์ตรวจวัดความแรง หลังจากการฝึกสอนโมเดลครั้งแรก ผู้วิจัยได้ทำการเพิ่มข้อมูลที่ผู้วิจัยได้รวบรวมมาและนำมาใช้ในการฝึกสอนโมเดลทั้ง 4 เทคนิคอีกครั้ง ทำให้ผู้วิจัยสามารถจำลองพฤติกรรมธรรมชาติของเส้นทางจริงได้และทำให้เกิดการเพิ่มข้อมูลที่ได้รับมาซึ่งนำไปใช้ในกระบวนการฝึกสอนและทดสอบที่ปรับปรุงขึ้นเพื่อสร้างโมเดลที่ดีขึ้น นอกจากนี้ โมเดลทั้งหมดที่ถูกฝึกสอนและนำข้อมูลที่เพิ่มเติมแล้วไปประยุกต์ใช้ได้ให้ผลลัพธ์ที่แม่นยำอย่างสูง ส่วนร่วมหลักของผู้วิจัยในวงการคือวิธีการประเมินความแม่นยำของโมเดลและอัลกอริทึมที่เรียบง่ายซึ่งผู้คนสามารถนำมาใช้เป็นพื้นฐานในการสร้างแอปพลิเคชันที่ใช้โมเดลเหล่านี้พร้อมกับผลลัพธ์การประเมินของโมเดล ผู้วิจัยได้ค้นพบและการเสนอผลงานของผู้วิจัยมุ่งเน้นในการให้ข้อมูลที่ผ่านการทดสอบอย่างเข้มงวดให้แก่วิศวกรและโรงงาน เพื่อช่วยป้องกันการเกิดขัดข้องในระบบท่อลมที่เกิดจากการรั่วซึมของลมอัดอากาศ ผู้วิจัยขอเสนอว่าส่วนร่วมเหล่านี้เป็นสิ่งใหม่และตามความรู้ของผู้วิจัยยังไม่เคยรายงานในวรรณกรรมก่อนหน้านี้ ดังนั้นวิทยานิพนธ์ฉบับนี้เป็นส่วนร่วมที่เกี่ยวข้องและสำคัญงานโรงงานที่ใช้ระบบลมอัดอากาศ

Title	A SYSTEM FOR DETECTION AND PREDICTION OF LEAKS IN AIR PIPE USING EDGE COMPUTING AND DEEP LEARNING
Author	Thanakrit Kraising
Advisor	Assistant Professor Winai Wongthai, Ph.D.
Co-Advisor	Aimaschana Nirauntasukrat, Ph.D
Academic Paper	M.S. Thesis in Computer Science, Naresuan University, 2023
Keywords	Compressed air leak, Pneumatic line leakage detection, Deep learning, Pneumatic pipe system, Decision Tree, Random Forest, support Vector Machine, 1D-CNN

ABSTRACT

Gas leaks from Fittings of a pneumatic pipe system result in the breakdown or failure of the system. This significant fault causes about half of the losses of output from production lines in the manufacturing sector. Deep Learning methods can be used to detect gas leakage of the pneumatic pipe system. We propose a Deep Learning model for the detection of air leaks from pneumatic pipe system Fittings using an accelerometer sensor system. We trained four models with four Machine Learning techniques with the data generated from our experimental pneumatic pipe. After the first training, we augmented the collected data and used it to train all four models again. We were able to mimic the natural behavior of the actual line and thereby augment the collected data which was used in an enhanced training and testing process to create a better model. Secondly, all the trained models in which the augmented data was applied yielded highly accurate results. This was especially demonstrated by our proposed Deep Learning model which yielded the highest accuracy of 99.2%. Our main contribution to the field is our method of evaluating the accuracy of the model and the simple algorithm that one may use as a basis for building applications based on the model, together with the model's evaluation results. Our findings and contribution provide well-tested information to engineers and companies to avoid breakdowns in pneumatic pipe systems caused by air leaks. We claim that these contributions are new, and to the

best of our knowledge have not previously been reported in the literature, thus are relevant and important contributions to the field.



ประกาศคุณูปการ

ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงในความกรุณาของ ผู้ช่วยศาสตราจารย์ ดร.วินัย วงษ์ไทย ประธานที่ปรึกษาวิทยานิพนธ์ที่ได้อุดหนุนให้เวลาอันมีค่ามาเป็นทีปรึกษา พร้อมทั้งให้คำแนะนำตลอดเวลาในการทำวิทยานิพนธ์ฉบับนี้ และกราบขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์อันประกอบไปด้วย ผู้ช่วยศาสตราจารย์ ดร.วินัย วงษ์ไทย ดร.เอ็มอัสนา นิรันตสุขรัตน์ ผู้ช่วยศาสตราจารย์ ดร.ธนธร พ่อคำ และดร.ณัฐพล คุ่มใหญ่โต กรรมการผู้ทรงคุณวุฒิภายใน ที่ได้กรุณาให้คำแนะนำตลอดจนแก้ไขข้อบกพร่องของวิทยานิพนธ์ด้วยความเอาใจใส่ จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้อย่างสมบูรณ์และทรงคุณค่า

เหนือสิ่งอื่นใดขอกราบขอบพระคุณ บิดา มารดาและเพื่อนของผู้วิจัยที่ทำให้กำลังใจและให้การสนับสนุนในทุกๆ ด้านอย่างดีที่สุดเสมอมา

คุณค่าและคุณประโยชน์อันพึงจะมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบและอุทิศแด่ผู้มีพระคุณทุกๆ ท่าน ผู้วิจัยหวังเป็นอย่างยิ่งว่า งานวิจัยนี้จะเป็นประโยชน์ต่อผู้ที่สนใจนำไปใช้ประโยชน์ไม่มากนักน้อย หากมีข้อบกพร่องประการใดที่อาจจะเกิดขึ้นภายในวิทยานิพนธ์ ผู้วิจัยขอน้อมรับเพื่อเป็นประโยชน์ในการพัฒนางานวิจัยต่อไป

ธนกฤต ไกรสิงห์

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	จ
ประกาศคุุณูปการ	ช
สารบัญ.....	ซ
สารบัญตาราง.....	ฅ
สารบัญรูป.....	ฉ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ปัญหางานวิจัย	4
1.3 วัตถุประสงค์ของงานวิจัย	4
เพื่อพัฒนาโมเดลสำหรับตรวจจับและพยากรณ์การรั่วของท่อลมในไลน์การผลิตท่อ โดยใช้เทคโนโลยีเอ็ดจ์คอมพิวเตอร์และการเรียนรู้เชิงลึก.....	4
เพื่อนำโมเดลที่ได้นำมาใช้กับแอปพลิเคชันจำลองของระบบตรวจจับและพยากรณ์ การรั่วของท่อลมในไลน์การผลิตท่อ.....	4
1.4 ขอบเขตของงานวิจัย	4
ขอบเขตการวิจัยด้านข้อมูล	4
ขอบเขตการวิจัยด้านระบบ	5
ขอบเขตการวิจัยด้านผู้ใช้งาน.....	5
ขอบเขตการวิจัยด้านเทคโนโลยี.....	5
1.5 คุุณูปการของงานวิจัย.....	5

ผู้วิจัยได้พบว่าโมเดลทั้งหมดที่ผลิตมาก่อนนี้มีระดับความแม่นยำที่ไม่เพียงพอ โดยพบว่าข้อมูลการทดสอบที่ผู้วิจัยบันทึกจากระบบทดสอบของผู้วิจัยไม่เหมือนกับข้อมูลการทำงานปกติและธรรมชาติของข้อมูลจริงที่ถูกเก็บรวบรวมจากไลน์โรงงาน แล้วผู้วิจัยได้ดำเนินการโดยการเปรียบเทียบข้อมูลการทดสอบของผู้วิจัยกับข้อมูลการผลิตจากไลน์การผลิตในโรงงาน และด้วยการเพิ่มข้อมูลการผลิตจากไลน์การผลิตเข้ากับข้อมูลการทดสอบของผู้วิจัยผู้วิจัยได้รับระดับความแม่นยำที่สูงขึ้นอย่างมีนัยสำคัญจากโมเดลการเรียนรู้เชิงลึกของผู้วิจัย 6

ที่อ้างถึงคือโมเดลการเรียนรู้เชิงลึก ที่ผู้วิจัยเสนอได้บรรลุความแม่นยำถึง 99.2% ส่วนหนึ่งของโมเดลการเรียนรู้เชิงลึกใหม่นี้เป็นอัลกอริทึมใหม่ที่ผู้วิจัยพัฒนาขึ้น นี่เป็นอัลกอริทึมที่เรียบง่าย แต่มีประสิทธิภาพที่มีผลในการประมวลผลของโมเดลการเรียนรู้เชิงลึก อย่างมีนัยสำคัญ..... 6

อิงจากโมเดลนี้ ส่วนสนับสนุนที่สามที่อ้างถึงคือ ตรรกะของโมเดลการเรียนรู้เชิงลึกและอัลกอริทึมใหม่ที่ผู้วิจัยพัฒนาขึ้นสามารถนำมาใช้ในการสร้างแอปพลิเคชันในอนาคตเพื่อตรวจจับการรั่วซึมที่เกิดขึ้นจริงและการรั่วซึมที่อาจเกิดขึ้นโดยใช้ระบบเซนเซอร์แอคเซลโรเมตร งานวิจัยที่ผู้วิจัยจะทำในอนาคตคือการสร้างระบบเซนเซอร์แอคเซลโรเมตรโดยใช้โมเดลการเรียนรู้เชิงลึก ใหม่และอัลกอริทึมที่เรียบง่าย 6

1.6 ประโยชน์ที่ได้รับ..... 6

ได้ต้นแบบระบบที่ใช้ในการหารอยรั่วในไลน์การผลิตแบบอัตโนมัติที่ทดสอบและสามารถใช้งานให้เป็นประโยชน์จริงกับภาคเอกชน (ตามเงื่อนไข ในหัวข้อ 1.6.2)..... 6

งานวิจัยนี้สามารถเป็นต้นแบบเพื่อปรับไปใช้กับรอยรั่วของท่อลม และท่อน้ำในไลน์การผลิตในโรงงานอื่น ๆ ได้เป็นจำนวนมาก..... 6

สามารถนำระบบ DPED ไปผลิต เพื่อให้ใช้งานได้ในโรงงานอื่น ๆ ให้ได้ตามมาตรฐานเชิงพาณิชย์ 6

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 งานวิจัยที่เกี่ยวข้อง	7
ทฤษฎีที่เกี่ยวข้อง	8
บทที่ 3 วิธีดำเนินการวิจัย.....	11
ระบบท่อลมในแบบงานการผลิตและปัญหาที่เกี่ยวข้อง.....	11
ข้อมูลสันสะท้อนและการตรวจจับความเสื่อมสภาพของเครื่อง	13
การออกแบบการทดลอง.....	13
บทที่ 4 ผลการวิจัย	15
การทดลองเก็บข้อมูล.....	15
ความสำคัญของกระบวนการฝึกโมเดล RF	34
เมทริกซ์ความสับสนของโมเดล 1D-CNN	35
แอปพลิเคชัน.....	37
บทที่ 5 บทสรุป.....	41
ข้อเสนอแนะ	41
บทสรุปวิทยานิพนธ์	43
ภาคผนวก.....	48
อภิธานศัพท์	53
บรรณานุกรม	2
ประวัติผู้วิจัย	4

สารบัญตาราง

หน้า

ตาราง 1 สรุปงานวิจัยที่เกี่ยวกับการตรวจจับการรั่วซึมลมจากพิตติงในระบบท่อลมปีบอัด	7
ตาราง 2 การเชื่อมต่อท่อลม พิตติงและวาล์ว	14
ตาราง 3 ตัวอย่างข้อมูลดิบของคลาส 0 พิตติงต่อกันแน่น วาล์วปิด	16
ตาราง 4 ตัวอย่างข้อมูลดิบของคลาส 1 พิตติงต่อกันแน่น วาล์วเปิด	17
ตาราง 5 ตัวอย่างข้อมูลดิบของคลาส 2 พิตติงกันต่อเกือบแน่น วาล์วปิด	18
ตาราง 6 ตัวอย่างข้อมูลดิบของคลาส 3 พิตติงต่อกันเกือบแน่น วาล์วเปิด	19
ตาราง 7 ตัวอย่างข้อมูลดิบของคลาส 4 พิตติงต่อกันหลวม วาล์วปิด	20
ตาราง 8 ตัวอย่างข้อมูลดิบของคลาส 5 พิตติงต่อกันหลวม วาล์วเปิด	21
ตาราง 9 การปรับแต่งไฮเปอร์พารามิเตอร์ของแต่ละเทคนิค	28
ตาราง 10 ตารางผลการฝึกอบรมโมเดลกับข้อมูลดิบ	29
ตาราง 11 ตารางเปรียบเทียบจำนวนข้อมูลดิบและข้อมูลปรับแต่ง	33
ตาราง 12 ผลของความแม่นยำระหว่างข้อมูลดิบและข้อมูลที่ปรับแต่ง	34

สารบัญรูป

	หน้า
รูป 1 ระบบท่อก่อนการเพิ่มส่วนประกอบการทดลอง.....	12
รูป 2 ระบบท่อดลองที่ถูกเพิ่มส่วนประกอบเพิ่มเติม.....	12
รูป 3 ขั้นตอนทั้ง 7 ของกระบวนการทดลอง.....	15
รูป 4 การแสดงข้อมูล (วาล์วปิด) (Data Plotting (valve closed)).....	25
รูป 5 โครงสร้างของ 1D-CNN ที่นำมาใช้ในกระบวนการฝึกอบรม.....	27
รูป 6 ตัวอย่างของข้อมูลที่แบ่งออกเป็น 5 ส่วน	31
รูป 7 การสุ่มข้อมูล และการแบ่งส่วนของข้อมูล	32
รูป 8 การรวมข้อมูลและกำหนดคลาส	32
รูป 9 ตัวอย่างข้อมูลที่ได้รับการปรับแต่งของแต่ละคลาส	33
รูป 10 คุณลักษณะ ความสำคัญของกระบวนการฝึกอบรมต้นไม้ป่าสุ่ม	35
รูป 11 คอนฟิชั่นเมตริก: การทำนายผลการทดสอบโมเดลด้วยเทคนิค 1D-CNN	36
รูป 12 สถาปัตยกรรมส่วนแอปพลิเคชัน	37
รูป 13 ตัวอย่างแอปพลิเคชันที่แสดงค่าของเซนเซอร์สามแกน.....	38
รูป 14 แอปพลิเคชันที่แสดงถึงค่าการทำนายเปรียบเทียบกับค่าจริง.....	38
รูป 15 แอปพลิเคชันแสดงข้อมูลดิบของสถานการณ์ล้มเหลวหรือมีลมรั่ว	39
รูป 16 แอปพลิเคชันแสดงผลการทำนายของโมเดลของสถานการณ์ล้มเหลวที่ได้มา.....	39
รูป 17 การติดตั้งอุปกรณ์การทดลองเก็บข้อมูล.....	49
รูป 18 การติดตั้งอุปกรณ์รับข้อมูลจากเซนเซอร์และเขียนลงเมมโมรี่การ์ด	49
รูป 19 วาล์วเปิด	50

รูป 20 วาล์วปิด.....	50
รูป 21 การติดตั้งเซนเซอร์กับฟิตติ้ง.....	51
รูป 22 การติดตั้งภาพรวมของการทดลองทั้งหมด	52



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในหัวข้อนี้จะกล่าวถึงปัญหา ที่มาและความสำคัญของงานวิจัยนี้ โดยโรงงานผลิตท่อและอุปกรณ์ ท่อพอลิเอทิลีนความหนาแน่นสูงท่อ (High Density Polyethylene หรือ HDPE) และ ท่อพอลิเอทิลีนความหนาแน่นต่ำ (Low density polyethylene หรือ LDPE) ได้เกิดปัญหาการรั่วของลมในไลน์การผลิตของเครื่องผลิตท่ออยู่บ่อยครั้ง ทางทีมวิศวกรของทางโรงงานที่ผลิตท่อได้กล่าวถึงปัญหาที่พบในไลน์การผลิตท่อ ว่าเครื่องผลิตท่อ นั้นมีส่วนประกอบของเครื่องปั๊มลม (Air Pump Industrial) และบ่อยครั้งได้มีการรั่วของสายยางที่ใช้เป่าลมในการผลิต เมื่อเกิดการรั่วของลม ทางโรงงานได้มีการวัดแรงดันลมที่จุดไหน ไว้สำหรับวัดแรงดันลมเมื่อแรงดันลมสูงเพิ่มขึ้นเกินที่กำหนดไว้ หมายถึง เกิดรอยรั่วของท่อลมในไลน์การผลิตท่อ จากนั้นทางโรงงานต้องหยุดการผลิตท่อของเครื่องผลิตนั้น และหลังจากนั้นทีมวิศวกรจะหารอยรั่วของท่อลมด้วยการใช้มือจับท่อลมและเดินตามสายท่อลมทั้งไลน์การผลิตเพื่อหารอยรั่วของลม เมื่อพบรอยรั่ว ทีมวิศวกรจึงได้ทำการซ่อมบำรุงเพื่อให้เครื่องผลิตกลับมาทำงานได้ตามปกติ ในบางครั้งทีมวิศวกรต้องเสียเวลาในการหารอยรั่วถึง 1 วันเต็ม ๆ และต้องหยุดการผลิต 1 วัน จึงทำให้เกิดความเสียหายกับโรงงานผลิตท่อ ทั้งด้านบุคลากร เวลา และมูลค่าการผลิต

ดังนั้น จากปัญหาที่กล่าวมาในหัวข้อข้างต้นนี้สามารถแก้ไขได้โดยมีงานวิจัยเป็นต้นแบบในการพัฒนา “ระบบตรวจจับและพยากรณ์การรั่วของท่อลมในไลน์การผลิตท่อโดยใช้เทคโนโลยีเอจคอมพิวติ้งและการเรียนรู้เชิงลึก” (A System for Detection and Prediction of Leaks in Air Pipe using Edge Computing and Deep Learning หรือ DPED) ขึ้นมาเพื่อแก้ปัญหาที่กล่าวมา โดยงานวิจัยที่สำรวจมาจะอธิบายไว้ในหัวข้อ 0 งานวิจัยที่เกี่ยวข้อง

ในภาคอุตสาหกรรมสายการผลิตหลาย ๆ ส่วนประกอบพึ่งพากับลมบีบอัดตามที่ถูกพูดถึงโดย Guenther และคณะ [1] ผู้วิจารณ์ได้ระบุว่า ในขณะที่ลมบีบอัดเป็นประโยชน์ การจัดหาลมบีบอัดผ่านระบบท่อลมพิเศษต่างๆ มีค่าใช้จ่ายสูง เช่นกัน ระบบเช่นนี้บ่อยครั้งถูกรบกวนด้วยการรั่วซึมของลมที่ลดประสิทธิภาพของระบบ ทำให้เกิดการขัดข้องในกระบวนการผลิต

หลักฐานแสดงให้เห็นว่าข้อบกพร่องในระบบเหล่านี้เป็นสาเหตุของการสูญเสียประมาณครึ่งหนึ่งของผลผลิตและผลผลิตในภาคอุตสาหกรรมเหล่านี้ ที่น่ากังวลอีกประการคือการรั่วซึมเหล่านี้บ่อยครั้งพบในท่อหลักในสายการผลิต เช่นกับงานวิจัยของ Zhang และคณะ [2] ได้ระบุว่ามิพินลำนท่อลมบีบอัดในโรงไฟฟ้าผลิตพลังงานนิวเคลียร์และโรงงานเคมี และขัดข้องในระบบลมบีบอัดในอุตสาหกรรมเหล่านี้โดยเฉพาะนั้นเป็นปัญหาที่มีความ

เสียงสูงอย่างมาก ผู้วิจัยได้สัมภาษณ์กับวิศวกรในโรงงานเพื่อปรึกษาเกี่ยวกับปัญหาทางเทคนิคในท่อลมบีบอัด และได้ให้ข้อมูลทั่วไปที่ดีเกี่ยวกับปัญหาในจุดหรือตำแหน่งต่าง ๆ ในสายการผลิตของโรงงาน

คำศัพท์ทางเทคนิคที่ใช้ในงานวิจัยนี้ประกอบด้วย:

- จุด (Point): จุดประกอบด้วยท่อสองชิ้นที่เชื่อมต่อกันด้วยฟิตติ้งสองทาง
- ฟิตติ้ง (Fitting): อุปกรณ์ที่ใช้เชื่อมต่อท่อเข้าด้วยกันในแบบสองทาง
- ถังรับอากาศ (Receiver): ส่วนประกอบของจุดที่เชื่อมต่อด้านหนึ่งกับท่อและด้านอีกด้านเชื่อมต่อกับเครื่องอัดอากาศ ตัวเครื่องรับจะรับลมบีบอัดที่เกิดจากเครื่องอัดลมและส่งผ่านไปยังไลน์
- เครื่องอัดอากาศ (Compressor): อุปกรณ์ที่ให้ลมบีบอัดให้กับท่อลมบีบอัด
- ไลน์ (Line): ส่วนทั้งหมดของท่อลมบีบอัดที่เริ่มต้นจากเครื่องอัดลมและตลอดทางที่มีจุดหลาย ๆ จุด
- ท่อ (Pipe): ท่อที่ส่งลมบีบอัดจากเครื่องอัดอากาศไปยังจุดต่าง ๆ บนไลน์

วิศวกรที่ผู้วิจัยได้ปรึกษาได้แจ้งให้ผู้วิจัยทราบว่าปัญหาที่สำคัญคือลมสามารถรั่วไหลจากฟิตติ้งที่ตั้งอยู่ที่จุดต่าง ๆ บนไลน์ได้ จำนวนของฟิตติ้งอาจมีจำนวนมากขึ้นขึ้นกับความยาวของไลน์ เงื่อนไขที่แตกต่างกันอาจทำให้เกิดการรั่วซึมเหล่านี้ รวมถึงการเสื่อมสภาพของฟิตติ้งที่อยู่บนไลน์จากเวลาผ่านไป การรั่วซึมประเภทนี้ยังพบเห็นได้ในฟิตติ้งของท่อน้ำพลาสติกที่เสี่ยงต่อการเสื่อมสภาพจากหลายสาเหตุที่แตกต่างกัน [3] ไม่เพียงแต่การค้นหานี้ใช้เวลานาน เป็นที่รู้กันอย่างแพร่หลายว่าการหยุดใช้งานหรือปิดระบบไปเป็นเวลานานอาจเกิดขึ้น แต่การขัดข้องที่มีเวลา ยาวนานอาจก่อให้เกิดปัญหาด้านการบำรุงรักษาที่มีค่าใช้จ่ายสูงและการสูญเสียผลผลิตเช่นกัน เหตุการณ์เช่นนี้เป็นเรื่องที่ไม่สามารถหลีกเลี่ยงได้ เนื่องจากระบบการตรวจจับถูกเริ่มทำงานหลังจากการรั่วซึมเกิดขึ้นแล้ว [4] วิธีหนึ่งในการป้องกันการขัดข้องเหล่านี้คือการทำการคาดการณ์การรั่วซึมโดยการตรวจจับเหตุการณ์การรั่วไหลตั้งแต่เริ่มต้นเกิดขึ้นและด้วยการพิสูจน์จุดที่รั่วไหลอย่างรวดเร็วและแม่นยำ จึงต้องการระบบเพื่อให้สามารถทำความเข้าใจเหตุการณ์เหล่านี้ในทันทีและเป็นที่เร่งด่วน

ในแง่นี้ เนื่องจากความเร่งด่วนผู้วิจัยขอเสนอโมเดลการเรียนรู้เชิงลึก (Deep Learning หรือ DL) สำหรับการตรวจจับการรั่วซึมของลมจากฟิตติ้งในระบบท่อลมบีบอัด โดยใช้ระบบเซนเซอร์วัดความเร่งสามแกน การเรียนรู้เชิงลึก เป็นเทคนิคการเรียนรู้ของเครื่องที่ใช้ระบบประสาทเทียมในการสร้างโมเดลการเรียนรู้เชิงลึก ซึ่งเป็นเทคนิคการเรียนรู้ของเครื่อง (Machine Learning หรือ ML) รายละเอียดทั้งหมดเกี่ยวกับเทคนิคนี้มีอยู่ในส่วนที่ 2.2 โดยสรุปโดยสั้น สำหรับการทดลองที่ผู้วิจัยรายงานในวิทยานิพนธ์นี้ ผู้วิจัยสร้างระบบทดสอบของท่อลมบีบอัดที่ประกอบด้วยส่วนประกอบหลักสองอย่างคือ ท่อและฟิตติ้ง ระบบนี้ถูกติดตั้งในห้องปฏิบัติการของผู้วิจัย และผู้วิจัยได้รวบรวมข้อมูลจากระบบทดสอบนี้ การกำหนดค่าจริงและฟังก์ชันของระบบนี้ โดยจะอธิบายในส่วนเนื้อหาเรื่องวัสดุและวิธีการด้านล่างนี้

การสั่นสะเทือนที่สม่ำเสมอและสม่ำเสมอจากฟิตติ้ง และตรวจจับค่าการสั่นสะเทือนรวมถึงบันทึกและใช้ในการระบุว่าฟิตติ้งใดเข้าสู่สถานการณ์ใกล้เคียงกับการล้มเหลวหรือสถานการณ์ล้มเหลว ซึ่งทำให้สามารถตรวจจับการรั่วซึมจากฟิตติ้งได้เร็ว หรือป้องกันการรั่วซึมก่อนที่จะเกิดขึ้น การขยายส่วนของโมเดลการเรียนรู้เชิงลึกในอนาคตจะ

ช่วยให้สามารถตรวจจับตำแหน่งจุดที่เกิดความล้มเหลว และนำทีมงานซ่อมบำรุงไปยังจุดโดยไม่จำเป็นต้องทำการ
ตรวจสอบด้วยตาอย่างลำบาก



1.2 ปัญหางานวิจัย

ผู้วิจัยกำลังพิจารณาสองสถานการณ์คือ 'สถานการณ์ใกล้ล้มเหลว' และ 'สถานการณ์ล้มเหลว' วิศวกรในสายการผลิตจำเป็นต้องให้ความสำคัญกับทั้งสองสถานการณ์เหล่านี้ของการรั่วซึมลมบีบอัดไปพร้อมกัน เมื่อใช้วิธีของผู้วิจัย ผู้วิจัยสามารถสร้างแอปพลิเคชันที่แจ้งเตือนวิศวกรเกี่ยวกับสองสถานการณ์: 'สถานการณ์ใกล้ล้มเหลว' หรือ 'สถานการณ์ล้มเหลว' ที่จะให้ข้อมูลแก่วิศวกรเพื่อป้องกันการขัดข้องของไลน์ ซึ่งจะช่วยป้องกันปัญหาการบำรุงรักษาที่มีค่าใช้จ่ายสูงและการขัดข้องในการผลิต

ทั้งนี้ในการจำลองสถานการณ์การเก็บข้อมูลที่เกิดขึ้นให้ตรงกับสภาพความเป็นจริงนั้นค่อนข้างเกิดขึ้นได้ยากจากการได้รับคำปรึกษาจากวิศวกร นั่นก็คือในโลกความเป็นจริงได้มิได้มีการเปิดวาล์วปล่อยลมยาวตลอด หรือปิดตลอด จะมีการเปิดและปิดอยู่เป็นระยะ ๆ นั้นทำให้ผู้วิจัยต้องหาอุปกรณ์เพิ่มให้การรับค่าจากวาล์วว่าเป็นสถานเปิดหรือปิดอยู่

ผลการวิจัยของผู้วิจัยไม่จำกัดเฉพาะโรงงานที่ได้ทดสอบเดียวเท่านั้น แต่สามารถเป็นประโยชน์ต่อโรงงานที่มีปัญหาการรั่วซึมของลมเหล่านี้โดยทั่วไป

1.3 วัตถุประสงค์ของงานวิจัย

เพื่อพัฒนาโมเดลสำหรับตรวจจับและพยากรณ์การรั่วของท่อลมในไลน์การผลิตท่อโดยใช้เทคโนโลยีเอ็ดจ์คอมพิวเตอร์และการเรียนรู้เชิงลึก

เพื่อนำโมเดลที่ได้นำมาใช้กับแอปพลิเคชันจำลองของระบบตรวจจับและพยากรณ์การรั่วของท่อลมในไลน์การผลิตท่อ

1.4 ขอบเขตของงานวิจัย

ขอบเขตการวิจัยด้านข้อมูล

ข้อมูลที่ใช้ในการเรียนรู้แรงสั่นสะเทือนของท่อลมขณะที่เครื่องผลิตท่อชุดจำลอง ที่สร้างขึ้น ณ มหาวิทยาลัยนครสวรรค์

- ข้อมูลได้จากการทดลอง เพื่อไปสร้างเป็นต้นแบบของลักษณะการสั่นของท่อลม และดูลักษณะการรั่วของท่อลม เพื่อนำไปให้เครื่องเรียนรู้ถึงลักษณะของการรั่วของท่อลม

- การทดลองในวิทยานิพนธ์นี้จะทดลองกับ 1 เซนเซอร์แอคเซลโรเมตรที่ 1 จุดติดตั้งเท่านั้น

ขอบเขตการวิจัยด้านระบบ

- ระบบ DPED ต้องมีโมเดลที่เรียนรู้แรงสั่นสะเทือนเพื่อวิเคราะห์รอยร้าวของท่อลมด้วยวิธีการ 1D-CNN
- ระบบ DPED ต้องพยากรณ์สถานการณ์ล่วงหน้าถึงการรั่วของท่อลม ด้วยเทคนิคการเรียนรู้เชิงลึก
- ระบบ DPED ต้องตรวจหารอยร้าวของท่อลมอัตโนมัติ

ขอบเขตการวิจัยด้านผู้ใช้งาน

- โรงงานที่มีส่วนเกี่ยวข้องการใช้งานท่อลมขนาดเล็กในการผลิต เช่น โรงงานผลิตท่อ HDPE อื่น ๆ ที่ใช้ท่อลมในการผลิต สามารถนำระบบไปใช้งานได้จริง

ขอบเขตการวิจัยด้านเทคโนโลยี

1. Edge computing
ใช้ในการประมวลผลข้อมูลที่ได้รับมา โดยจะทำการประมวลให้ใกล้กับแหล่งข้อมูลมากที่สุด
2. Python
เป็น programming language ที่ใช้ในการพัฒนาระบบขึ้นมา ตั้งแต่รับข้อมูลจาก sensor และประมวลผลข้อมูล
3. Accelerometer sensor mpu-9250/6050
เป็นอุปกรณ์ที่ใช้ในการวัดความเร่งในแนวแกน x, y, z ซึ่งสามารถปรับใช้ในการหาแรงสั่นสะเทือนได้ โดยจะใช้เพียงพารามิเตอร์จากเซนเซอร์ดังกล่าวเพียงค่าแกน x, แกน y, และ แกน z เท่านั้น

1.5 คุณูปการของงานวิจัย

มี 3 คุณูปการที่เกิดจากการศึกษาของผู้วิจัยในวิทยานิพนธ์นี้

ผู้วิจัยได้พบว่าโมเดลทั้งหมดที่ผลิตมาก่อนนี้มีระดับความแม่นยำที่ไม่เพียงพอ โดยพบว่าข้อมูลการทดสอบที่ผู้วิจัยบันทึกจากระบบทดสอบของผู้วิจัยไม่เหมือนกับข้อมูลการทำงานปกติและธรรมชาติของข้อมูลจริงที่ถูกเก็บรวบรวมจากไลน์โรงงาน แล้วผู้วิจัยได้ดำเนินการโดยการเปรียบเทียบข้อมูลการทดสอบของผู้วิจัยกับข้อมูลการผลิตจากไลน์การผลิตในโรงงาน และด้วยการเพิ่มข้อมูลการผลิตจากไลน์การผลิตเข้ากับข้อมูลการทดสอบของผู้วิจัยผู้วิจัยได้รับระดับความแม่นยำที่สูงขึ้นอย่างมีนัยสำคัญจากโมเดลการเรียนรู้เชิงลึกของผู้วิจัย

ที่อ้างถึงคือโมเดลการเรียนรู้เชิงลึก ที่ผู้วิจัยเสนอได้บรรลุความแม่นยำถึง 99.2% ส่วนหนึ่งของโมเดลการเรียนรู้เชิงลึกใหม่นี้เป็นอัลกอริทึมใหม่ที่ผู้วิจัยพัฒนาขึ้น นี่เป็นอัลกอริทึมที่เรียบง่าย แต่มีประสิทธิภาพที่มีผลในการประมวลผลของโมเดลการเรียนรู้เชิงลึก อย่างมีนัยสำคัญ อิงจากโมเดลนี้ ส่วนสนับสนุนที่สามที่อ้างถึงคือ ตรรกะของโมเดลการเรียนรู้เชิงลึก และอัลกอริทึมใหม่ที่ผู้วิจัยพัฒนาขึ้นสามารถนำมาใช้ในการสร้างแอปพลิเคชันในอนาคตเพื่อตรวจจับการรั่วซึมลมที่เกิดขึ้นจริงและการรั่วซึมที่อาจเกิดขึ้นโดยใช้ระบบเซนเซอร์แอคเชลโรเมตร งานวิจัยที่ผู้วิจัยจะทำในอนาคตคือการสร้างระบบเซนเซอร์แอคเชลโรเมตรโดยใช้โมเดลการเรียนรู้เชิงลึก ใหม่และอัลกอริทึมที่เรียบง่าย

1.6 ประโยชน์ที่ได้รับ

ได้ต้นแบบระบบที่ใช้ในการหารอยรั่วในไลน์การผลิตแบบอัตโนมัติที่ทดสอบและสามารถใช้งานให้เป็นประโยชน์จริงกับภาคเอกชน (ตามเงื่อนไข ในหัวข้อ 0)

งานวิจัยนี้สามารถเป็นต้นแบบเพื่อปรับไปใช้กับรอยรั่วของท่อลม และท่อน้ำในไลน์การผลิตในโรงงานอื่น ๆ ได้เป็นจำนวนมาก

สามารถนำระบบ DPED ไปผลิต เพื่อให้ใช้งานได้ในโรงงานอื่น ๆ ให้ได้ตามมาตรฐานเชิงพาณิชย์

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 งานวิจัยที่เกี่ยวข้อง

รายงานในหัวข้อก่อนหน้าเกี่ยวกับการตรวจจับการรั่วซึมจากท่อน้ำและท่อลมได้ถูกระบุ เช่น [1] ได้นำเสนอการตรวจจับการรั่วซึมลมบีบอัดโดยใช้ระบบเซนเซอร์อัลตราโซนิกสแกนเนอร์ที่ช่วยให้สามารถสแกนท่อเหล็กและตรวจจับการรั่วซึมลมได้

การตรวจจับข้อบกพร่องการรั่วซึมก๊าซในระบบท่อลมบีบอัดโดยใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) ซึ่งเป็นเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ถูกพูดถึงใน [2] โดยในนั้นใช้เครื่องมือวัดเสียงเพื่อวิเคราะห์สัญญาณเสียงที่เกิดขึ้น กระบวนการในการประมาณค่าสเปกตรัมเสียงของเสียงรั่วซึมจากพิตติงในท่อน้ำพลาสติกที่ฝังอยู่ได้ถูกพูดถึงใน [3] ซึ่งใช้เซนเซอร์ไฮโดโฟนสำหรับการวัดเสียงหรือเซนเซอร์แอคเซลโรเมตรที่อยู่ห่างจากจุดรั่วซึมสำหรับการวัดสั่นสะเทือน เรื่องสรุปของงานวิจัยนี้พร้อมทั้งงานของผู้วิจัยถูกนำเสนอในตาราง 1 ตารางแสดงด้านสำคัญของโครงการวิจัยแต่ละงานโดยอิงจาก (1) วัตถุประสงค์หลักของแอปพลิเคชัน, (2) การนำเทคนิคการเรียนรู้เชิงเครื่องมือมาใช้, (3) เซนเซอร์/เครื่องมือที่ใช้, และ (4) สิ่งที่จะตรวจจับ

ตาราง 1 สรุปงานวิจัยที่เกี่ยวข้องกับการตรวจจับการรั่วซึมลมจากพิตติงในระบบท่อลมบีบอัด

งานวิจัย	วัตถุประสงค์หลักของแอปพลิเคชัน	การนำเทคนิคการเรียนรู้เชิงเครื่องมือมาใช้	เซนเซอร์/เครื่องมือที่ใช้	สิ่งที่ตรวจจับ
[1]	เพื่อตรวจจับการรั่วซึมลม	-	อัลตราโซนิก	Steel pipe
[2]	ตรวจจับการรั่วซึมลมบีบอัดโดยใช้เทคนิคการเรียนรู้เชิงลึก	เทคนิคการเรียนรู้เชิงลึก	เครื่องมือวัดเสียง	เสียงรั่วซึม
[3]	การประมาณค่าสเปกตรัมเสียงของเสียงรั่วซึมจากพิตติงในท่อน้ำพลาสติกที่ฝังอยู่	เทคนิคการเรียนรู้เชิงลึก	ไฮโดโฟนหรือเซนเซอร์วัดความเร่งสามแกน	เสียงหรือสั่นสะเทือนที่เกิดจากการรั่วซึม
วิทยานิพนธ์ของผู้วิจัย	ตรวจจับการรั่วซึมลมจากพิตติงในระบบท่อลมบีบอัด	เทคนิคการเรียนรู้เชิงลึกและอัลกอริทึมใหม่	เซนเซอร์วัดความเร่งสามแกน	เสียงหรือสั่นสะเทือนที่เกิดจากการรั่วซึม

จากตาราง 1 ผู้วิจัยได้เห็นว่างานวิจัยที่อ้างถึงก่อนหน้านี้เกี่ยวข้องกับตรวจจับการรั่วซึม โดยมีนักวิจัยเพียงคนเดียวที่ใช้เทคนิคการเรียนรู้ของเครื่อง งานของ [3] เกี่ยวข้องกับการตรวจจับพิตติงและใช้เซนเซอร์วัดความเร่งสามแกนในการวิเคราะห์ท่อน้ำพลาสติกที่ถูกฝังอยู่ ไม่ใช่ท่อลม นอกเหนือจากนี้ [5] ใช้เซนเซอร์วัดความเร่ง

สามแกน MPU6050 เพื่อรวบรวมข้อมูลสั่นสะเทือนที่ใช้ในการตรวจจับการรั่วซึมท่อน้ำ และเซนเซอร์ MPU6050 ได้รับการตรวจสอบความสัมพันธ์ของสั่นสะเทือนจากการคำนวณ [5] เทคนิคการเรียนรู้ของเครื่อง เป็นวิธีที่สำคัญในการทำให้ปัจจัยเทียบเท่ากับปัจจัยประสงค์ของมนุษย์ การเรียนรู้ของเครื่องเป็นโมเดลที่เน้นการระบุโครงสร้างที่ยังไม่ได้ค้นพบ ความสัมพันธ์และฟังก์ชันระหว่างข้อมูลนำเข้าและผลลัพธ์ [6] การเรียนรู้ของเครื่องสามารถจัดการกับข้อมูลที่มีมิติสูงและหลากหลาย ทำให้เป็นเทคนิคที่มีศักยภาพสำหรับการพัฒนาอัลกอริทึมที่คาดเดาได้อย่างมีความสามารถ [7] ตัวอย่างเช่น [8] เพิ่มความแม่นยำในการตรวจจับการรั่วซึมลมด้วยเทคนิคการเรียนรู้ของเครื่อง

ทฤษฎีที่เกี่ยวข้อง

คำจำกัดความพื้นฐานของเทคนิคการเรียนรู้ของเครื่อง ถูกนำมาจาก [9] เทคนิคการเรียนรู้ของเครื่องเป็นสาขาย่อยของปัญญาประดิษฐ์ (Artificial Intelligence, AI) ส่วนเทคนิคการเรียนรู้เชิงลึก เป็นสาขาย่อยของเทคนิคการเรียนรู้ของเครื่อง โครงข่ายประสาทเทียม (Neural Network, NN) เป็นส่วนหลักของอัลกอริทึมในเทคนิคการเรียนรู้เชิงลึก จำนวนหรือความลึกของเลเยอร์ในโครงข่ายประสาทเทียม เป็นสิ่งสำคัญเนื่องจากอัลกอริทึม เทคนิคการเรียนรู้เชิงลึก ต้องมีความลึกมากกว่าสามเลเยอร์ ตามที่ได้กล่าวถึงใน [6] เทคนิคการเรียนรู้ของเครื่องเป็นรูปแบบที่มุ่งเน้นการระบุลักษณะการเชื่อมโยงภายในชุดข้อมูลและวัตถุประสงค์ของข้อมูล ในการข้อมูลนำเข้าและผลลัพธ์ เทคนิคการเรียนรู้ของเครื่องสามารถจัดการข้อมูลที่มีมิติสูงและหลายตัวแปรได้ ซึ่งทำให้มันเป็นวิธีที่มีศักยภาพสำหรับการพัฒนาอัลกอริทึมทำนายที่แม่นยำมากขึ้นได้ [7] การนำเทคนิคการเรียนรู้ของเครื่องมาใช้ตามที่กล่าวถึงใน [8] เพื่อเพิ่มความแม่นยำในการตรวจจับการรั่วของอากาศ โดย [10] ได้กล่าวไว้ว่า มีหลายเทคนิค/อัลกอริทึมของเทคนิคการเรียนรู้ของเครื่องที่สามารถจำแนกได้ในทางสองลักษณะ แบ่งเป็นสองลักษณะ คือ MLs ซึ่งรวมถึงเทคนิคต้นไม้ตัดสินใจ หรือ Decision Tree (DT), ป่าไม้สุ่ม หรือ Random Forest (RF), และเครื่องมือเวกเตอร์สนับสนุน หรือ Support Vector Machine (SVM) และลักษณะที่สองคือเทคนิคการเรียนรู้เชิงลึก ซึ่งได้รวมถึงเทคนิคโครงข่ายประสาทแบบคอนโวลูชัน หรือ Convolutional Neural Networks (CNNs) ผู้วิจัยให้ความสำคัญกับเทคนิคต้นไม้ตัดสินใจ, ป่าไม้สุ่ม, เครื่องมือเวกเตอร์สนับสนุน, โครงข่ายประสาทแบบคอนโวลูชัน ซึ่งผู้วิจัยเรียกว่า ซีเอ็นเอ็น ตอนนี้เป็นเทคนิคพื้นฐานสำหรับการทดลองของผู้วิจัยในการทดลองนี้ผู้วิจัยใช้ทั้ง 4 เทคนิคเหล่านี้เป็นเทคนิคการเรียนรู้เชิงลึกมีสอนสำหรับงานการจำแนก

ต้นไม้ตัดสินใจ เป็นเทคนิคแบ่งแยกสำหรับการทำการถดถอย (Regression) และการจำแนกแบ่งกลุ่ม (Classification) ของข้อมูลตามที่กำหนดไว้ [11] ใช้เทคนิคต้นไม้ตัดสินใจ เพื่อเรียนรู้ลักษณะจากรูปแบบที่สามารถระบุได้ในฐานข้อมูลขนาดใหญ่ ลักษณะและรูปแบบเหล่านี้มีความสำคัญสำหรับการแยกและการสร้างโมเดลสำหรับทำนายข้อมูล และสามารถนำมาใช้ในการฝึกฝนและสร้างโมเดลด้วยเทคนิคต้นไม้ตัดสินใจ ตามที่กำหนดไว้ใน [10] กลุ่มของโมเดลเทคนิคต้นไม้ตัดสินใจ สามารถนำมาใช้ในการฝึกฝนและหลังจากนั้นจะสามารถสร้างโมเดลต้นไม้ป่าสุ่มได้

เครื่องมือเวกเตอร์สนับสนุนเป็นวิธีที่ใช้เทคนิคการเรียนรู้ของเครื่องให้มีประสิทธิภาพ ที่ใช้ในงานแบ่งประเภท โดยมีวัตถุประสงค์เพื่อสร้างโมเดลที่สามารถแบ่งข้อมูลออกเป็นกลุ่มๆ ตามหมวดหมู่ที่กำหนดไว้ โดยมีพื้นฐานทฤษฎีจากการคำนวณเวกเตอร์ซึ่งสามารถแบ่งข้อมูลให้อยู่ในลักษณะเชิงเส้น [10]

อย่างไรก็ตามผู้วิจัยใช้เทคนิคต้นไม้ตัดสินใจ, เทคนิคต้นไม้ป่าสุ่ม, และเครื่องมือเวกเตอร์สนับสนุน สำหรับการวิเคราะห์พื้นฐานของข้อมูลที่ผู้วิจัยรวบรวมมา ร่วมกับ 1D-CNN ผู้วิจัยจะได้กล่าวถึงต่อไปในส่วน 3.1 ผลลัพธ์ตัวอย่างจากการวิเคราะห์นี้เพื่อเข้าใจข้อมูลที่เก็บได้ เช่น ผลทำนายสำหรับกระบวนการฝึกสอนโมเดลของเทคนิคต้นไม้ป่าสุ่ม ก่อนกระบวนการฝึกเทคนิคต้นไม้ตัดสินใจ, เทคนิคต้นไม้ป่าสุ่ม, และเครื่องมือเวกเตอร์สนับสนุน ข้อมูลที่เก็บได้ถูกสกัดออก และผลลัพธ์ของการสกัดนี้เรียกว่า “Input features” หรือ “คุณลักษณะสำหรับนำเข้า” และคุณลักษณะสำหรับนำเข้าเหล่านี้ถูกนำไปใช้ในกระบวนการฝึกเทคนิคต้นไม้ตัดสินใจ, เทคนิคต้นไม้ป่าสุ่ม, และเครื่องมือเวกเตอร์สนับสนุน

กระบวนการความสำคัญของคุณลักษณะ เป็นกระบวนการที่ให้คะแนนกับคุณลักษณะสำหรับนำเข้าตามความมีประสิทธิภาพของคุณลักษณะที่ทำนายตัวแปรตามเป้าหมาย [12]

LeCun et al. [13] กล่าวว่า “เทคนิคการเรียนรู้เชิงลึกเป็นการทำให้โมเดลทางคอมพิวเตอร์ที่ประกอบด้วยชั้นประมวลผลหลาย ๆ ชั้น สามารถเรียนรู้การแทนข้อมูลที่มีระดับของขนาดต่าง ๆ”

การเรียนรู้เชิงลึกผู้วิจัยของยกตัวอย่างของเทคนิคการเรียนรู้เชิงลึกที่ใช้ CNN [10] CNN เป็นเทคนิคที่ถูกลิขิตมานานมานี้ ที่มีความสำคัญสำหรับการแก้ปัญหาการรู้จำแบบรูป [14] โครงสร้างของ CNN ที่ถูกนำมาใช้ในงานวิจัยนี้ถูกนิยามไว้ใน [15]

โครงสร้างการฝึกโมเดล CNN ประกอบด้วยชั้นหลัก 3 ชั้น: ชั้นแรกเป็นชั้นการคอนโวลูชัน (Convolutional layer), ชั้นที่สองชั้นกระจายค่าเฉลี่ย (Global average pooling layer), และชั้นที่สามชั้นการจัดประเภท softmax [15] ชั้นแรกและชั้นที่สองใช้สำหรับการสกัดลักษณะของข้อมูลที่เก็บได้ ชั้นที่สามใช้สำหรับการจัดประเภท หรือผู้วิจัยสามารถเรียกชั้นนี้ว่าชั้นผลลัพธ์ ชั้นนี้สามารถสร้างตัวจัดประเภท fully-connected แบบ softmax ที่มี C นิวรอนที่เท่ากับจำนวนคลาสในชุดข้อมูล สามชั้นนี้เป็นกระบวนการหลักของการทดลองของผู้วิจัย

ผู้วิจัยจะพูดถึงรายละเอียดของการประมวลผลของทุกชั้นในการทดลองในส่วน 2.5.2 การทดลองสร้างข้อมูลการสั่นสะเทือนจากการร่วมนของท่อทดสอบในระบบอัด นักวิจัยจาก [16] ทำทนายว่าข้อมูลการสั่นสะเทือนเช่นของผู้วิจัยสามารถถือว่าเป็นข้อมูลหนึ่งมิติหรือ 1D data ไม่ใช่ข้อมูลสองมิติหรือ 2D data ซึ่งเหมาะสำหรับการประมวลผลภาพ และ 1D-CNN เหมาะสำหรับข้อมูล 1D ดังนั้นผู้วิจัยก็นำ 1D-CNN มาใช้เพื่อสกัดลักษณะนามธรรมจากข้อมูลการสั่นสะเทือนที่เก็บได้ ผู้วิจัยจำแนกข้อมูลการสั่นสะเทือนเหล่านี้เป็น 6 ชั้น ซึ่งถูกระบุรายละเอียดในส่วน 2.5.2 1D-CNN ได้ถูกใช้โดย [17] สำหรับการวินิจฉัยข้อบกพร่องของสัญญาณการสั่นสะเทือนจากลูกปืนม้วน 1D-CNN สามารถสร้างโมเดลที่มีระดับความแม่นยำในการจัดประเภทสูงสุดถึง 95% ความแม่นยำนี้สามารถประเมินได้โดยการคำนวณเมตริกซ์ความสับสน ซึ่งถูกพัฒนาขึ้นเพื่อตรวจสอบข้อบกพร่องในความแม่นยำจากระดับสูงสุดที่เป็นไปได้คือ 100% ยกตัวอย่างเช่น ความแม่นยำอื่น ๆ 5% โมเดลจะจัดประเภทข้อมูลตามตัวจัดประเภทที่กำหนดไว้ เมท

ริกซ์ความสับสนคือการสรุปข้อมูลในรูปแบบตารางสองมิติที่แสดงจำนวนการทำนายที่ถูกต้องและไม่ถูกต้องโดยตัวจัดประเภท มันถูกใช้เพื่อวัดประสิทธิภาพของโมเดลการจัดประเภท มันสามารถใช้ในการประเมินประสิทธิภาพของโมเดลการจัดประเภทผ่านการคำนวณเมตริกการทำงาน เช่น ความแม่นยำ (accuracy), ความแม่นยำเชิงละเอียด (precision), การเรียกคืน (recall), และ F1-score นอกจากนี้ [18] ยังอธิบายถึงวิธีที่เมตริกซ์ความสับสนถูกใช้ในการทบทวนประสิทธิภาพของตัวจัดประเภทต่อข้อมูลทดสอบเมื่อตัวจัดประเภทนี้ทำการจัดประเภท ในเมตริกซ์ความสับสน มิติที่หนึ่งแทนคลาสจริงของวัตถุ ในขณะที่มิติที่สองแทนคลาสที่ทำนายไว้ [19] ส่วน 3.3 จะอธิบายรายละเอียดทั้งหมดเกี่ยวกับวิธีที่ผู้วิจัยใช้เมตริกซ์ความสับสนเพื่อสร้างผลลัพธ์การทดลองหลักของผู้วิจัย



บทที่ 3

วิธีดำเนินการวิจัย

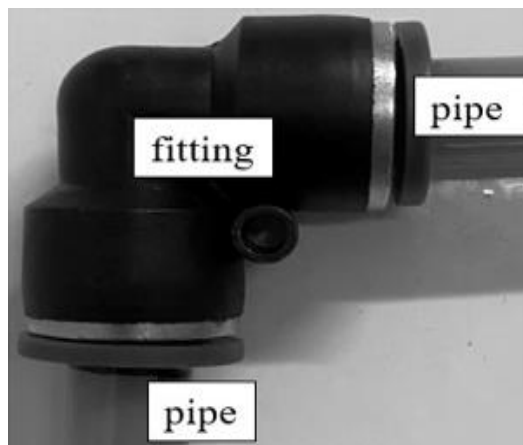
ระบบท่อลมในแบบงานการผลิตและปัญหาที่เกี่ยวข้อง

ในโรงงานหรือโรงงานที่มีการผลิตเป็นส่วนใหญ่ ระบบท่อลมเป็นส่วนสำคัญของการดำเนินงาน เนื่องจากท่อลมบีบอัดนี้เป็นแหล่งให้พลังงานแก่อุปกรณ์ต่าง ๆ ที่ใช้ในการผลิต แต่ปัญหาที่เกิดขึ้นกับระบบท่อลมบีบอัดนี้คือการรั่วซึมของลมที่อาจจะเกิดขึ้นได้ การรั่วซึมลมจากระบบท่อลมทำให้เกิดปัญหาหลายอย่าง เช่น:

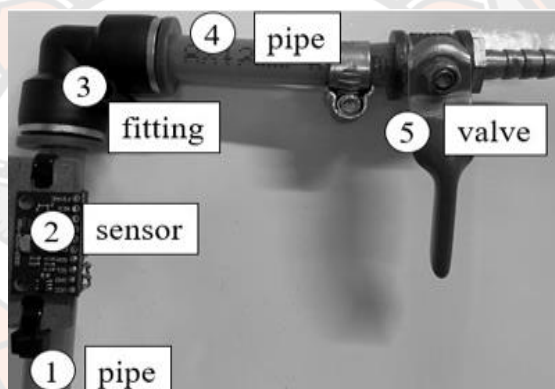
- ขาดการทำงานอย่างเต็มประสิทธิภาพ: การรั่วซึมลมจะทำให้เกิดการสูญเสียของลมที่มีค่าประหยัด จากนั้นอุปกรณ์ที่ขึ้นอยู่กับลมอาจไม่ทำงานอย่างเต็มประสิทธิภาพ
- การขัดข้องในกระบวนการผลิต: การรั่วซึมลมอาจทำให้อุปกรณ์หรือเครื่องจักรที่ใช้ในกระบวนการผลิตขัดข้อง หรือไม่ทำงานอย่างถูกต้อง
- สิ่งแวดล้อม: การรั่วซึมลมอาจมีผลต่อสิ่งแวดล้อม เช่น การสูญเสียพลังงาน การปล่อยก๊าซที่อาจมีผลกระทบต่อสภาพอากาศ
- ความปลอดภัย: การรั่วซึมลมในสถานที่ที่ใช้ก๊าซอันตรายอาจสร้างความเสี่ยงต่อความปลอดภัยของบุคคลในพื้นที่
- ค่าใช้จ่ายในการบำรุงรักษา: การรั่วซึมลมเมื่อไม่ได้รับการแก้ไขหรือบำรุงรักษาทันทีอาจทำให้เกิดค่าใช้จ่ายในการซ่อมแซมและบำรุงรักษาสูง
- สูญเสียผลผลิต: การรั่วซึมลมอาจทำให้เกิดความหยุดยั้งในกระบวนการผลิตและสูญเสียผลผลิต

การควบคุมและตรวจจ็ับการรั่วซึมลมจากระบบท่อลมบีบอัดเป็นสิ่งสำคัญในการรักษาประสิทธิภาพและความปลอดภัยของกระบวนการผลิต การใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) และเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) สามารถช่วยในการตรวจจ็ับและป้องกันการรั่วซึมลมอย่างมีประสิทธิภาพมากขึ้น ทำให้เกิดประสิทธิภาพในกระบวนการผลิตและประหยัดค่าใช้จ่ายในการบำรุงรักษาและการซ่อมแซม

วิศวกรของโรงงานที่ผู้วิจัยศึกษาได้ให้ข้อมูลทั่วไปเกี่ยวกับท่อและอุปกรณ์ของจุดหรือตำแหน่งในสายผลิตดังรูป 1 และ รูป 2 แสดงระบบท่อก่อนขั้นตอนการออกแบบทดลอง และหลังจากขั้นตอนการออกแบบทดลอง รูป 1 แสดงระบบท่อก่อนขั้นตอนการออกแบบทดลอง และรูป 2 แสดงระบบท่อทดลอง รูป 2 จะถูกพูดถึงอย่างละเอียดในส่วนที่ 0 ของวิทยานิพนธ์นี้



รูป 1 ระบบท่อนก่อนการเพิ่มส่วนประกอบการทดลอง



รูป 2 ระบบท่อทดลองที่ถูกเพิ่มส่วนประกอบเพิ่มเติม

รูป 1 เป็นแสดงจุดที่ประกอบด้วยท่อลมสองแห่งที่เชื่อมต่อกันด้วยฟิตติ้งสองทางหรือคอนเนคเตอร์ หนึ่งท่อเชื่อมต่อกับถังรับอากาศ โดยอีกปลายเชื่อมต่อกับเครื่องอัดอากาศ (ไม่แสดงในรูป) ท่ออีกแห่งนำไปยังเส้นผลิตเครื่องอัดอากาศ (Compressor) ที่ทำงานตลอดเวลาจะสร้างลมบีบอัดที่จะเก็บไว้ในถังรับอากาศ แล้วจึงถูกนำผ่านท่อลมเพื่อส่งลมบีบอัดไปยังอุปกรณ์ลมบีบอัดในเส้นผลิตรูป 2 แสดงเส้นผลิตที่ถูกเพิ่มความสามารถสำหรับการทดสอบสองส่วนประกอบที่ถูกเพิ่มในรูป 2 ประกอบด้วยวาล์ว (ที่ติดป้ายหมายเลข 5) และเซนเซอร์ (ที่ติดป้ายหมายเลข 2) การติดตั้งอุปกรณ์เซนเซอร์นั้นผู้วิจัยได้ทำการทดลองหาระยะที่เหมาะสมของการติดเซนเซอร์นี้ โดยได้ข้อมูลมาให้ใกล้ชิดกับฟิตติ้งที่ต้องการวัดให้มากที่สุด เพื่อที่จะให้เซนเซอร์ตรวจวัดและรับรู้ถึงแรงสั่นสะเทือนได้ดีที่สุด

ปัญหาที่ต้องการแก้ไขคือการลมอาจรั่วออกจากฟิตติ้งเนื่องจากความเสื่อมสภาพของฟิตติ้ง ซึ่งอาจทำให้เกิดความขัดข้องในเส้นผลิต วิศวกรต้องค้นหาที่รั่วหรือที่รั่วตามเส้นทั้งหมดโดยการตรวจสอบและสัมผัสฟิตติ้งแต่ละอันในเส้น ซึ่งอาจมีจำนวนมากถึงร้อยตัวในเส้นที่ยาวมาก กระบวนการค้นหามักใช้เวลาานาน ซึ่งจะเป็นการสูญเสีย

ผลผลิตจากเส้นผลิตนั้น งานวิจัยได้แสดงให้เห็นว่าการร่วมนำทำให้เกิดขาดการผลิตถึงครึ่งหนึ่งของงานการผลิตในภาคการผลิต [1]

ข้อมูลสันสะเทือนและการตรวจจับความเสื่อมสภาพของเครื่อง

การเสื่อมสภาพของเครื่องสามารถตรวจจับได้จากการเปลี่ยนแปลงในการสั่นสะเทือนของเครื่อง [4] เซนเซอร์แอคเซลโรเมตรสามารถวัดค่าความเร่งที่แสดงเป็นค่าบนแกนสามทิศทาง ที่เรียกว่าแกน x , y , และ z ค่าเร่งที่สามแกนเหล่านี้เป็นข้อมูลสันสะเทือน การร่วมนำจากฟิตติ้งทำให้เกิดการสั่นสะเทือนที่แตกต่างจากสันสะเทือนปกติของเส้นที่ทำงาน ในการทดลองผู้วิจัยใช้เซนเซอร์แอคเซลโรเมตรเพื่อวัดสันสะเทือนของระบบท่ออย่างต่อเนื่อง (รูป 1) และบันทึกข้อมูลสันสะเทือนของระบบท่อที่ถูกทดสอบ เสียงสันสะเทือนจากการร่วมนำจะแสดงในรูปแบบที่ไม่เป็นปกติของการสั่นสะเทือนทั่วไป การออกแบบและการทดลองจะถูกกล่าวถึงในรูป 2 ในส่วน 2.5

การใช้ 1D-CNN ถูกพูดถึงในงานวิจัย [17] ที่ใช้ 1D-CNN ในการวินิจฉัยข้อเสียของข้อมูลการสั่นสะเทือนจากปลั๊กปืนห้องกลม ด้วยข้อมูลนี้ผู้วิจัยได้ใช้ความสนใจในการใช้ 1D-CNN ในการทดลองของผู้วิจัยเพื่อจำแนกข้อมูลการสั่นสะเทือนที่เก็บจากระบบทดสอบของผู้วิจัย และข้อมูลการผลิตจากโรงงานที่วิศวกรให้คำปรึกษาผู้วิจัย

การออกแบบการทดลอง

การออกแบบการทดลองนี้เน้นตามส่วนของสายผลิตจริงของโรงงาน เสนอแบบทดลองของระบบทดสอบขนาดเล็กของผู้วิจัยได้รับการตรวจสอบจากวิศวกรของโรงงานนั้น ผู้วิจัยใช้ส่วนประกอบต้นฉบับสองชิ้นและท่อหนึ่งชิ้นจากโรงงานในการออกแบบระบบทดสอบ ผู้วิจัยซื้อส่วนประกอบอื่นจากผู้ค้าที่เหมาะสมโดยมั่นใจว่าข้อมูลคำสั่งจะได้รับการอนุมัติจากวิศวกร เพื่อทดสอบทั้งสองสถานการณ์ คือ ฟิตติ้งที่ทำงานอย่างถูกต้องและฟิตติ้งที่อาจเสื่อมสภาพ จึงใช้ฟิตติ้งหนึ่งชิ้นใหม่และหนึ่งชิ้นที่ใช้แล้วในการทดลอง

รูป 2 แสดงระบบท่อที่ผู้วิจัยใช้ในการทดลองทั้งหมด วาล์ว (5) ควบคุมการเปิดปิดการไหลของลมอากาศที่จำลองการทำงานปกติของระบบผลิต ตัวเซ็นเซอร์ (2) เป็นเซนเซอร์แอคเซลโรเมตรประเภท MPU6050 ที่วัดความสั่นสะเทือนที่เกิดขึ้นจากระบบท่อทดลอง และถูกบันทึกและเก็บไว้ในการ์ดหน่วยความจำเพื่อการวิเคราะห์เพิ่มเติม MPU6050 ได้รับการตรวจสอบและยืนยันว่าเป็นอุปกรณ์ที่ผ่านการทดสอบอย่างดีที่สุดสำหรับการตรวจจับการสั่นสะเทือนที่คำนวณเข้าใช้งาน [5] MPU6050 บันทึกข้อมูลสันสะเทือนด้วยค่าเร่งบนแกนสามทิศทาง (แกน x , แกน y , และ แกน z) ตามแบบธรรมดาของเซนเซอร์วัดความเร่งสามแกน

สำหรับการเชื่อมต่อของส่วนประกอบในการทดลองรูป 2 ยังแสดงการเชื่อมต่อของส่วนประกอบในระบบท่อทดลอง คอมเพรสเซอร์จะเปิดที่ตลอดเวลาและอากาศจะไหลออกจากเรซีฟเวอร์ไปที่ท่อด้านซ้าย (1) จากนั้นผ่านเซนเซอร์แล้วผ่านฟิตติ้งเข้าไปที่ท่อด้านขวา (4) และไปยังเส้นผลิต (ซึ่งไม่เป็นส่วนหนึ่งของการทดสอบของผู้วิจัย) วาล์ว (5) ควบคุมการไหลเข้าสู่เส้นผลิต

มีการทดสอบการเชื่อมต่อทั้งหมดสามแบบสำหรับพิตติงและท่อด้านซ้าย: (i) เมื่อพิตติง (3) ถูกเชื่อมต่อแน่นกับท่อด้านซ้าย (1) นี้จะทำให้เกิดเสียงเงียบ น่าจะเป็นการสันสะเทือนระดับต่ำมาก ที่บ่งบอกว่าไม่มีการรั่วลมจากพิตติง (ii) เมื่อพิตติงถูกเชื่อมต่ออย่างเป็นประกายกลางกับท่อด้านซ้าย (1) เนื่องจากบางประกายที่เกิดจากเสียดสีหรืออาการเสียดสีทำให้เกิดการรั่วลมเล็กน้อยซึ่งจะเป็นเสียงระดับต่ำและอาจมีการสันสะเทือนเพิ่มเติมเกินจากปกติ (iii) เมื่อพิตติงถูกเชื่อมต่ออย่างหลวมกับท่อด้านซ้าย (1) ซึ่งทำให้เกิดเสียงระดับสูงและการสันสะเทือนขนาดใหญ่ที่เกิดจากการรั่วลมมากจากพิตติง ผู้วิจัยใช้พิตติงใหม่สำหรับ (i) และใช้พิตติงที่ใช้แล้วสำหรับ (ii) และ (iii)

ผู้วิจัยได้สร้างคลาสข้อมูลทั้งหมด 6 คลาสสำหรับข้อมูลที่ได้รับการเก็บรวบรวมโดยอิงจากความรู้ของวิศวกร

ตาราง 2 การเชื่อมต่อท่อลม พิตติงและวาล์ว

Configuration	ชื่อคลาส	พิตติง	วาล์ว
(i) ต่อพิตติงกับท่อลมให้แน่น (1)	คลาส 0	ต่อกันแน่น	ปิด
	คลาส 1	ต่อกันแน่น	เปิด
(ii) เมื่อพิตติงถูกเชื่อมต่อกับท่อด้านซ้าย (1) ในระดับความแน่นหนากลาง เนื่องจากมีความผิดปกติที่เป็นจริงหรือเป็นไปได้	คลาส 2	ต่อกันเกือบแน่น	ปิด
	คลาส 3	ต่อกันเกือบแน่น	เปิด
(iii) เมื่อพิตติงถูกเชื่อมต่อกับท่อด้านซ้าย (1) โดยไม่แน่นหนามาก จะทำให้เกิดเสียงระดับสูง และการสันสะเทือนขนาดใหญ่	คลาส 4	ต่อกันหลวม	ปิด
	คลาส 5	ต่อกันหลวม	เปิด

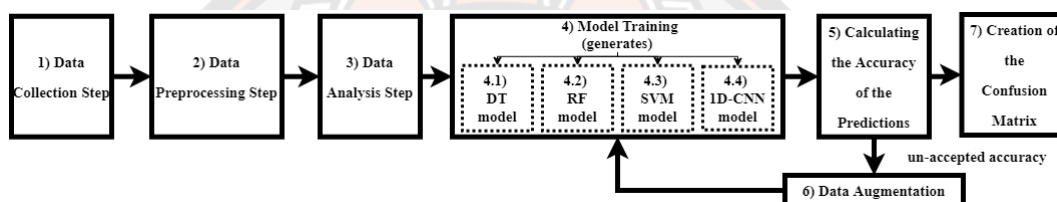
การเชื่อมต่อทั้งสามรูปแบบที่กล่าวถึงข้างต้น เรียกว่า (i), (ii) และ (iii) ถูกจัดหมวดหมู่เป็น: การกำหนดค่าของฮาร์ดแวร์และซอฟต์แวร์สำหรับการทดลองประกอบด้วย AMD Ryzen 5 3600 CPU, RAM 3200 MHz, GPU NVIDIA GeForce GTX 1070, และ Python 3.8.5 รวมถึงไลบรารีการแยกแยะและสกัดลักษณะของชุดข้อมูลชุดเวลาหรือ TSFEL [20], scikit-learn [21], และ Keras [22].

บทที่ 4

ผลการวิจัย

การทดลองเก็บข้อมูล

เจ็ดขั้นตอนของการทดลองถูกแสดงในรูป 3 ซึ่งประกอบด้วย 1) การเก็บข้อมูล, 2) การประมวลผลข้อมูล, 3) การวิเคราะห์ข้อมูล, 4) การฝึกโมเดล, 5) การทำนาย, 6) การเพิ่มข้อมูล, และ 7) การประเมินความแม่นยำของโมเดล



รูป 3 ขั้นตอนทั้ง 7 ของกระบวนการทดลอง

ข้อมูลการสั่นสะเทือนถูกเก็บรวบรวมและจัดประเภทเป็นหกคลาสใน 7 ขั้นตอน:

ขั้นตอนที่ 1 เก็บข้อมูล

ข้อมูลการสั่นสะเทือนถูกเก็บรวบรวมจากเซนเซอร์เร็กเซลโลเมตร (2) ในรูป 2 สำหรับติดตั้งที่เป็นใหม่และใช้แล้ว เพื่อเก็บข้อมูลสำหรับแต่ละคลาส ผู้วิจัยปฏิบัติตามขั้นตอนนี้: เมื่อปิดวาล์วบนติดตั้งที่เชื่อมต่อแน่นหนา ผู้วิจัยเก็บข้อมูลสำหรับคลาส 0 เป็นเวลา 1 นาทีและข้อมูลถูกบันทึกลงในไฟล์ .csv ในระหว่าง 1 นาทีนี้, พิตติ้งถูกสุ่มตัวอย่างทุก 50 มิลลิวินาที, ทำให้ได้ 3,000 จุดข้อมูลในไฟล์ .csv และแต่ละจุดข้อมูลเรียกว่า "timesteps" (ไทม์สเต็ป) จะถูกทำซ้ำ 19 ครั้ง โดยให้ได้ 20 ไฟล์ .csv แยกไฟล์ ซึ่งทำให้ผู้วิจัยได้จำนวนข้อมูล 60,000 จุดสำหรับ Class 0 เช่นเดียวกับนั้น ข้อมูลที่ถูกเก็บสำหรับ Class 2 (การเชื่อมต่อ Fitting อย่างหลวมเล็กน้อยกับวาล์วปิด) โดยให้ได้ 20 ไฟล์ .csv ที่มีข้อมูลของ Class 2 พร้อมกับ 60,000 จุดข้อมูลข้อมูลสำหรับ Class 4 (การเชื่อมต่อ Fitting โดยให้หลวมเล็กน้อยกับวาล์วเปิด) ก็ถูกเก็บรวบรวมไว้อีกด้วย โดยให้ได้ 20 ไฟล์ .csv ที่มี 60,000 จุดข้อมูล จากนั้น ด้วยการเปิดวาล์ว กระบวนการนี้ถูกทำซ้ำสำหรับ Class 1 (การเชื่อมต่อ Fitting แบบแน่นหนา, วาล์วเปิด), Class 3 (การเชื่อมต่อ Fitting แบบมีความแน่นหนากลางพร้อมวาล์วเปิด), และ Class 5 (การเชื่อมต่อ Fitting แบบหลวมเล็กน้อยกับวาล์วเปิด) ทั้งหมดนี้นำไปสู่การสร้างข้อมูลที่ใช้ในการฝึกโมเดลและทดสอบโมเดลของผู้อ่านได้

ตาราง 3 ตัวอย่างข้อมูลดิบของคลาส 0 ฟิตติ้งต่อกันแน่น วาล์วปิด

x_axis	y_axis	z_axis	timestamp	milliseconds
2.49	-11.89	-8.36	2000-1-8T1:23:38	441139
2.5	-11.9	-8.36	2000-1-8T1:23:38	441169
2.5	-11.89	-8.33	2000-1-8T1:23:38	441189
2.52	-11.89	-8.38	2000-1-8T1:23:38	441209
2.51	-11.91	-8.38	2000-1-8T1:23:38	441229
2.5	-11.9	-8.36	2000-1-8T1:23:38	441254
2.49	-11.9	-8.38	2000-1-8T1:23:38	441274
2.5	-11.9	-8.38	2000-1-8T1:23:38	441293
2.51	-11.9	-8.34	2000-1-8T1:23:38	441314
2.49	-11.9	-8.36	2000-1-8T1:23:38	441334
2.5	-11.89	-8.36	2000-1-8T1:23:38	441354
2.49	-11.89	-8.36	2000-1-8T1:23:38	441374
2.5	-11.91	-8.37	2000-1-8T1:23:38	441397
2.5	-11.9	-8.36	2000-1-8T1:23:38	441417
2.49	-11.9	-8.34	2000-1-8T1:23:38	441437
2.5	-11.89	-8.38	2000-1-8T1:23:38	441457
2.5	-11.89	-8.36	2000-1-8T1:23:38	441477
2.51	-11.91	-8.33	2000-1-8T1:23:38	441497
2.49	-11.9	-8.35	2000-1-8T1:23:38	441517
2.49	-11.89	-8.35	2000-1-8T1:23:38	441537
2.51	-11.88	-8.36	2000-1-8T1:23:38	441557
2.5	-11.89	-8.35	2000-1-8T1:23:38	441577
2.49	-11.89	-8.37	2000-1-8T1:23:38	441597
2.5	-11.88	-8.36	2000-1-8T1:23:38	441617
2.5	-11.91	-8.35	2000-1-8T1:23:38	441640
2.5	-11.89	-8.35	2000-1-8T1:23:38	441660
2.48	-11.9	-8.37	2000-1-8T1:23:38	441680
2.52	-11.88	-8.35	2000-1-8T1:23:38	441700
2.48	-11.9	-8.34	2000-1-8T1:23:38	441720
2.5	-11.9	-8.33	2000-1-8T1:23:38	441740
2.49	-11.89	-8.36	2000-1-8T1:23:38	441760

ตาราง 4 ตัวอย่างข้อมูลดิบของคลาส 1 พัดตั้งต่อกันแน่น วาล์วเปิด

x_axis	y_axis	z_axis	timestamp	milliseconds
2.51	-11.86	-8.37	2004-1-8T5:26:8	591279
2.51	-11.88	-8.37	2004-1-8T5:26:8	591309
2.5	-11.88	-8.35	2004-1-8T5:26:8	591329
2.49	-11.88	-8.37	2004-1-8T5:26:8	591349
2.51	-11.9	-8.37	2004-1-8T5:26:8	591368
2.52	-11.88	-8.34	2004-1-8T5:26:8	591388
2.5	-11.88	-8.37	2004-1-8T5:26:8	591408
2.5	-11.9	-8.35	2004-1-8T5:26:8	591428
2.49	-11.89	-8.36	2004-1-8T5:26:8	591448
2.5	-11.89	-8.37	2004-1-8T5:26:8	591468
2.5	-11.89	-8.36	2004-1-8T5:26:8	591488
2.49	-11.88	-8.36	2004-1-8T5:26:8	591508
2.5	-11.89	-8.35	2004-1-8T5:26:8	591528
2.5	-11.87	-8.36	2004-1-8T5:26:8	591551
2.5	-11.87	-8.36	2004-1-8T5:26:8	591571
2.51	-11.87	-8.41	2004-1-8T5:26:8	591591
2.5	-11.88	-8.37	2004-1-8T5:26:8	591611
2.52	-11.88	-8.36	2004-1-8T5:26:8	591631
2.51	-11.89	-8.36	2004-1-8T5:26:8	591651
2.51	-11.87	-8.35	2004-1-8T5:26:8	591671
2.51	-11.87	-8.37	2004-1-8T5:26:8	591691
2.5	-11.89	-8.35	2004-1-8T5:26:8	591711
2.5	-11.91	-8.33	2004-1-8T5:26:8	591731
2.5	-11.9	-8.34	2004-1-8T5:26:8	591751
2.49	-11.91	-8.32	2004-1-8T5:26:8	591771
2.51	-11.89	-8.33	2004-1-8T5:26:8	591794
2.51	-11.9	-8.35	2004-1-8T5:26:8	591814
2.5	-11.89	-8.39	2004-1-8T5:26:8	591834
2.5	-11.9	-8.34	2004-1-8T5:26:8	591854
2.49	-11.87	-8.37	2004-1-8T5:26:8	591874
2.51	-11.9	-8.38	2004-1-8T5:26:8	591894

ตาราง 5 ตัวอย่างข้อมูลดิบของคลาส 2 พัดตั้งกันต่อเกือบแน่น วาล์วปิด

x_axis	y_axis	z_axis	timestamp	milliseconds
2.68	-9.1	-7.9	2004-1-9T11:18:14	190028
2.67	-9.12	-7.93	2004-1-9T11:18:14	190058
2.69	-9.11	-7.92	2004-1-9T11:18:14	190079
2.67	-9.11	-7.92	2004-1-9T11:18:14	190098
2.68	-9.11	-7.89	2004-1-9T11:18:14	190118
2.67	-9.11	-7.92	2004-1-9T11:18:14	190138
2.68	-9.1	-7.92	2004-1-9T11:18:14	190158
2.67	-9.11	-7.94	2004-1-9T11:18:14	190178
2.69	-9.11	-7.93	2004-1-9T11:18:14	190198
2.67	-9.11	-7.92	2004-1-9T11:18:14	190218
2.69	-9.12	-7.93	2004-1-9T11:18:14	190238
2.68	-9.11	-7.91	2004-1-9T11:18:14	190258
2.68	-9.13	-7.94	2004-1-9T11:18:14	190281
2.67	-9.11	-7.95	2004-1-9T11:18:14	190301
2.68	-9.11	-7.95	2004-1-9T11:18:14	190321
2.67	-9.09	-7.91	2004-1-9T11:18:14	190341
2.68	-9.12	-7.93	2004-1-9T11:18:14	190361
2.66	-9.12	-7.91	2004-1-9T11:18:14	190381
2.67	-9.11	-7.92	2004-1-9T11:18:14	190401
2.68	-9.11	-7.93	2004-1-9T11:18:14	190421
2.7	-9.11	-7.93	2004-1-9T11:18:14	190441
2.67	-9.11	-7.95	2004-1-9T11:18:14	190461
2.67	-9.11	-7.9	2004-1-9T11:18:14	190481
2.68	-9.12	-7.91	2004-1-9T11:18:15	190501
2.68	-9.11	-7.93	2004-1-9T11:18:15	190524
2.68	-9.12	-7.9	2004-1-9T11:18:15	190544
2.68	-9.12	-7.94	2004-1-9T11:18:15	190564
2.68	-9.11	-7.92	2004-1-9T11:18:15	190584
2.67	-9.09	-7.93	2004-1-9T11:18:15	190604
2.69	-9.11	-7.93	2004-1-9T11:18:15	190624
2.69	-9.1	-7.92	2004-1-9T11:18:15	190644

ตาราง 6 ตัวอย่างข้อมูลดิบของคลาส 3 พัดตั้งต่อกันเกือบแน่น วาล์วเปิด

x_axis	y_axis	z_axis	timestamp	milliseconds
2.62	-9.08	-7.93	2000-1-8T1:17:32	73363
2.66	-9.06	-7.95	2000-1-8T1:17:32	73392
2.66	-9.04	-7.95	2000-1-8T1:17:32	73412
2.62	-9.04	-7.94	2000-1-8T1:17:32	73431
2.61	-9.01	-7.92	2000-1-8T1:17:32	73451
2.64	-9.03	-7.93	2000-1-8T1:17:32	73471
2.64	-9.03	-7.87	2000-1-8T1:17:33	73492
2.65	-9.04	-7.91	2000-1-8T1:17:33	73512
2.66	-9.08	-7.92	2000-1-8T1:17:33	73532
2.64	-9.06	-7.89	2000-1-8T1:17:33	73551
2.65	-9.05	-7.92	2000-1-8T1:17:33	73571
2.65	-9.1	-7.9	2000-1-8T1:17:33	73591
2.64	-9.1	-7.91	2000-1-8T1:17:33	73611
2.62	-9.08	-7.89	2000-1-8T1:17:33	73634
2.64	-9.07	-7.88	2000-1-8T1:17:33	73654
2.61	-9.06	-7.9	2000-1-8T1:17:33	73674
2.64	-9.08	-7.91	2000-1-8T1:17:33	73693
2.64	-9.05	-7.92	2000-1-8T1:17:33	73713
2.64	-9.05	-7.89	2000-1-8T1:17:33	73733
2.63	-9.03	-7.92	2000-1-8T1:17:33	73753
2.63	-9.04	-7.92	2000-1-8T1:17:33	73773
2.65	-9.02	-7.9	2000-1-8T1:17:33	73793
2.62	-9.07	-7.91	2000-1-8T1:17:33	73813
2.63	-9.06	-7.91	2000-1-8T1:17:33	73833
2.61	-9.08	-7.9	2000-1-8T1:17:33	73853
2.64	-9.08	-7.92	2000-1-8T1:17:33	73876
2.63	-9.06	-7.93	2000-1-8T1:17:33	73896
2.63	-9.11	-7.91	2000-1-8T1:17:33	73916
2.66	-9.05	-7.91	2000-1-8T1:17:33	73936
2.61	-9.04	-7.93	2000-1-8T1:17:33	73956
2.66	-9.04	-7.9	2000-1-8T1:17:33	73976

ตาราง 7 ตัวอย่างข้อมูลดิบของคลาส 4 พัดตั้งต่อกันหลวม วาล์วปิด

x_axis	y_axis	z_axis	timestamp	milliseconds
2.41	-9.52	-8.04	2014-1-15T15:22:11	322572
2.41	-9.51	-8.04	2014-1-15T15:22:11	322602
2.41	-9.47	-8.04	2014-1-15T15:22:11	322621
2.39	-9.53	-8.09	2014-1-15T15:22:11	322641
2.38	-9.5	-8.1	2014-1-15T15:22:11	322661
2.43	-9.31	-8.11	2014-1-15T15:22:11	322681
2.43	-9.47	-8.13	2014-1-15T15:22:11	322701
2.43	-9.47	-8.13	2014-1-15T15:22:11	322721
2.38	-9.43	-8.11	2014-1-15T15:22:11	322741
2.41	-9.3	-8.04	2014-1-15T15:22:11	322762
2.42	-9.52	-8.06	2014-1-15T15:22:11	322782
2.42	-9.48	-8.06	2014-1-15T15:22:11	322807
2.42	-9.54	-8.11	2014-1-15T15:22:11	322830
2.35	-9.6	-8.06	2014-1-15T15:22:11	322850
2.41	-9.4	-8.05	2014-1-15T15:22:11	322870
2.39	-9.52	-8.08	2014-1-15T15:22:11	322890
2.43	-9.3	-8.11	2014-1-15T15:22:11	322910
2.39	-9.38	-8.18	2014-1-15T15:22:11	322930
2.42	-9.33	-8.05	2014-1-15T15:22:11	322950
2.39	-9.29	-8.09	2014-1-15T15:22:11	322970
2.38	-9.48	-8.12	2014-1-15T15:22:11	322990
2.41	-9.35	-8.16	2014-1-15T15:22:11	323010
2.38	-9.43	-8.12	2014-1-15T15:22:11	323030
2.47	-9.5	-8.02	2014-1-15T15:22:11	323050
2.41	-9.59	-8.03	2014-1-15T15:22:11	323073
2.38	-9.51	-8.13	2014-1-15T15:22:11	323093
2.41	-9.6	-8.1	2014-1-15T15:22:11	323113
2.42	-9.54	-8.1	2014-1-15T15:22:11	323133
2.33	-9.55	-8.13	2014-1-15T15:22:11	323153
2.41	-9.5	-8.01	2014-1-15T15:22:11	323172
2.38	-9.33	-8.14	2014-1-15T15:22:11	323192

ตาราง 8 ตัวอย่างข้อมูลดิบของคลาส 5 ฟิตติ้งต่อกันหลวม วาล์วเปิด

x_axis	y_axis	z_axis	timestamp	milliseconds
2.58	-9.32	-8.05	2004-1-8T5:25:46	569637
2.58	-9.33	-8.01	2004-1-8T5:25:46	569667
2.58	-9.37	-8.04	2004-1-8T5:25:46	569687
2.55	-9.36	-8.05	2004-1-8T5:25:46	569707
2.58	-9.39	-8.03	2004-1-8T5:25:46	569726
2.56	-9.38	-8.02	2004-1-8T5:25:46	569746
2.58	-9.44	-8.01	2004-1-8T5:25:46	569766
2.59	-9.46	-8.04	2004-1-8T5:25:46	569786
2.56	-9.42	-8.03	2004-1-8T5:25:46	569807
2.55	-9.42	-8.05	2004-1-8T5:25:46	569827
2.58	-9.45	-8	2004-1-8T5:25:46	569847
2.58	-9.4	-8	2004-1-8T5:25:46	569867
2.61	-9.41	-8.04	2004-1-8T5:25:46	569887
2.56	-9.39	-8.06	2004-1-8T5:25:46	569910
2.56	-9.36	-8.02	2004-1-8T5:25:46	569930
2.6	-9.37	-8.01	2004-1-8T5:25:46	569950
2.58	-9.37	-8.02	2004-1-8T5:25:46	569970
2.58	-9.34	-8.02	2004-1-8T5:25:46	569990
2.59	-9.35	-8.03	2004-1-8T5:25:46	570010
2.58	-9.37	-8.01	2004-1-8T5:25:47	570030
2.56	-9.38	-7.98	2004-1-8T5:25:47	570050
2.6	-9.38	-8.02	2004-1-8T5:25:47	570070
2.56	-9.42	-8.01	2004-1-8T5:25:47	570090
2.56	-9.4	-8.01	2004-1-8T5:25:47	570110
2.58	-9.45	-8.01	2004-1-8T5:25:47	570130
2.6	-9.42	-8.03	2004-1-8T5:25:47	570153
2.58	-9.42	-8.04	2004-1-8T5:25:47	570173
2.56	-9.42	-8.05	2004-1-8T5:25:47	570193
2.6	-9.43	-8.05	2004-1-8T5:25:47	570213
2.57	-9.42	-8.03	2004-1-8T5:25:47	570233
2.58	-9.37	-8.04	2004-1-8T5:25:47	570253

ขั้นตอนที่ 2: ขั้นตอนการประมวลผลข้อมูล

การทำความสะอาดข้อมูลเป็นกระบวนการที่สำคัญในการเตรียมข้อมูลสำหรับการประมวลผล มันมุ่งเน้นการลบข้อมูลที่ขาดหายไป แก้ไขข้อมูลที่ไม่ถูกต้อง และปรับปรุงโครงสร้างข้อมูลเพื่อให้เหมาะสมกับการวิเคราะห์หรือการประมวลผลที่กำหนดเป้าหมายไว้ เช่น เมื่อมีข้อมูลที่สูญหายหรือไม่สมบูรณ์ เราสามารถเติมค่าที่สมบูรณ์ได้โดยใช้ค่าเฉลี่ยหรือค่ากลางของข้อมูลอื่น ๆ ในชุดข้อมูล ในขณะเดียวกัน การลบข้อมูลที่ซ้ำซ้อนหรือไม่จำเป็นออกจากชุดข้อมูลช่วยลดความซับซ้อนของข้อมูลและเพิ่มประสิทธิภาพในการวิเคราะห์ หลังจากการทำความสะอาดเสร็จสิ้น ข้อมูลจะมีคุณภาพที่ดีและพร้อมใช้งานสำหรับการวิเคราะห์และการประมวลผลต่อไป

ข้อมูลที่ได้รับจากขั้นตอนที่ 1 จะนำไปทำความสะอาด จัดเตรียม และปรับปรุงให้เหมาะสมต่อทหขั้นของข้อมูลความเร่ง ขั้นตอนนี้เป็นส่วนสำคัญในการเตรียมข้อมูลสำหรับการฝึกอบรมและวัดผลของโมเดล การทำความสะอาดและการจัดเตรียมข้อมูลที่ดีสามารถส่งผลต่อประสิทธิภาพของโมเดลที่จะถูกสร้างขึ้น การทำความสะอาดข้อมูลนั้นได้มีขั้นตอนดังนี้

- ขั้นตอนการแปลงข้อมูลจากข้อความหรือ String เป็น ตัวเลขทศนิยม หรือ float มีความสำคัญเพราะมันช่วยให้โมเดลเรียนรู้และทำนายข้อมูลได้โดยใช้ข้อมูลที่ถูกต้องและเข้าใจได้ การแปลงข้อมูลจากข้อความเป็นตัวเลขทศนิยมช่วยให้เราสามารถนำข้อมูลที่เป็นตัวเลขและมีค่าต่างๆ ไปใช้ในการฝึกโมเดลได้ โดยไม่จำเป็นต้องกังวลเกี่ยวกับการเข้าใจข้อมูลในรูปแบบของ String ที่ไม่ชัดเจน การทำขั้นตอนนี้เป็นส่วนสำคัญในการเตรียมข้อมูลก่อนนำไปใช้ในการฝึกโมเดลและทำนายผลลัพธ์ในโมเดล การเรียนรู้ของเครื่องให้ถูกต้องและมีประสิทธิภาพที่ดีมากขึ้น โดยเปลี่ยนลักษณะข้อมูลหรือ data type ของแกน x, แกน y, และแกน z ให้เป็นประเภทตัวเลขทศนิยมเพื่อนำไปใช้ในการประมวลผลต่าง ๆ ได้
- ประมวลผลข้อมูล การทำข้อมูลให้อยู่ศูนย์กลาง (zero mean) คือกระบวนการที่ปรับค่าข้อมูลให้มีค่าเฉลี่ยเป็นศูนย์ โดยการลบค่าเฉลี่ยของข้อมูลออกจากข้อมูลทุกตัว การทำ การทำข้อมูลให้อยู่ศูนย์กลาง เป็นกระบวนการที่สำคัญในการปรับปรุงคุณภาพข้อมูลก่อนนำไปใช้ในการฝึกโมเดลหรือการประมวลผลอื่น ๆ ข้อดีของการทำ การทำข้อมูลให้อยู่ศูนย์กลาง รวมถึงการทำให้ค่าเฉลี่ยของข้อมูลเป็นศูนย์ ทำให้เห็นแนวโน้มของข้อมูลได้ชัดเจนขึ้น และช่วยลดความแตกต่างของข้อมูลที่มีค่ามากหรือน้อยออกไป ทำให้โมเดลที่ฝึกด้วยข้อมูลที่มี การทำข้อมูลให้อยู่ศูนย์กลาง มีโอกาสที่จะให้ผลลัพธ์ที่ดีขึ้นกระบวนการการทำ การทำข้อมูลให้อยู่ศูนย์กลาง มักใช้ในการปรับค่าข้อมูลในชุดข้อมูล โดยการลบค่าเฉลี่ยของทุกตัวออกจากข้อมูลทั้งหมด ซึ่งสามารถทำได้โดยคำนวณค่าเฉลี่ยของข้อมูลแต่ละคุณลักษณะแล้วลบค่าเฉลี่ยนั้นออกจากทุกตัวของคุณลักษณะที่เกี่ยวข้อง ในกรณีของรูปแบบข้อมูลที่เป็นเมทริกซ์ การทำ การทำข้อมูลให้อยู่ศูนย์กลาง จะใช้วิธีการหักเฉลี่ยของแต่ละคอลัมน์ออกจากค่าในคอลัมน์นั้นทั้งหมด และในกรณีของข้อมูลชุดเดียว การทำ การทำข้อมูลให้อยู่ศูนย์กลาง จะหมายถึงการลบค่าเฉลี่ยของทุกตัวออกจากชุดข้อมูลทั้งหมด แล้วหลังจากนั้นนำค่าของแต่ละแกน ลบด้วยค่าเฉลี่ยของแกนนั้น ๆ เช่น ค่าแกน x - ค่าเฉลี่ยของแกน x
- ขั้นตอนของการทำ Outlier Treatment เป็นขั้นตอนที่ใช้ในการจัดการกับข้อมูลที่มีค่าที่ผิดปกติหรือเชิงสุ่มที่ไม่เกี่ยวข้องข้อมูลหลักฐานของเรา โดยมักจะใช้เพื่อปรับปรุงคุณภาพข้อมูลและป้องกันความ

ผิดพลาดในการวิเคราะห์ การ Outlier Treatment สามารถทำได้โดยการกำหนดค่าสำหรับค่าที่ผิดปกติให้เป็นค่าที่สมเหตุสมผลหรือใช้เทคนิคการจัดข้อมูลที่ผิดปกติออกไปจากชุดข้อมูล ซึ่งการปรับปรุงนี้ช่วยเพิ่มความถูกต้องและความเชื่อถือในการวิเคราะห์ข้อมูล คำว่า "Outlier Treatment" สามารถแปลเป็น "การจัดการข้อมูลที่ผิดปกติ" โดย Outlier หมายถึงค่าที่อยู่นอกเหนือจากการกระจายตัวของข้อมูลหลักฐาน การทำ Outlier Treatment ก็คือการปรับแก้หรือกำจัดค่าที่ผิดปกติที่อยู่นอกจากชุดข้อมูลเพื่อให้ข้อมูลที่ได้อสมเหตุสมผลและนำไปใช้ในการวิเคราะห์ได้อย่างถูกต้องและเชื่อถือได้มากยิ่งขึ้น ผู้วิจัยได้เปลี่ยนข้อมูลที่เป็นค่าผิดปกติ (outlier) ผู้วิจัยได้เพิ่มเงื่อนไขการทำความสะอาดข้อมูลไปเลยว่า ถ้าค่าของ แกน x, แกน y, และแกน z ในแถวใดที่มีค่ามากกว่า 0.4 และน้อยกว่า -0.4 ให้แปลงเป็นค่า Not a Number (NaN) หรือค่าว่าง

- d) ขั้นตอนการ fillna หรือ "การเติมค่าที่ขาดหายไป" เป็นขั้นตอนที่ใช้ในการจัดการกับค่าที่ขาดหายไป ชุดข้อมูล ซึ่งข้อมูลที่ขาดหายไปอาจจะเป็นเป็นค่าว่าง (NaN) หรือสัญลักษณ์อื่นที่ใช้แทนค่าที่ขาดหายไป เช่น 0 หรือ -1 ขึ้นอยู่กับวิธีการเก็บข้อมูล ในขั้นตอนการเติมค่าที่ขาดหายไปในวิเคราะห์ข้อมูลจะเลือกวิธีการเติมค่าที่ขาดหายไปเพื่อให้ได้ชุดข้อมูลที่สมบูรณ์และสมดุล เช่น ใช้ค่าเฉลี่ยหรือค่ามัธยฐานของข้อมูลที่มีอยู่แล้วในชุดข้อมูล เพื่อเติมลงไปในพื้นที่ที่ขาดหายไป หรือใช้วิธีการอื่น เช่น การเติมค่าที่ขาดหายไปด้วยค่าที่เหมาะสมจากข้อมูลที่ใกล้เคียง หรือการใช้โมเดลเพื่อทำนายค่าที่ขาดหายไปแล้วเติมลงไป ชุดข้อมูล การเติมค่าที่ขาดหายไปเป็นขั้นตอนสำคัญที่ช่วยให้ได้ชุดข้อมูลที่ครบถ้วนและพร้อมสำหรับการวิเคราะห์และการประมวลผลต่อไป การใช้การเติมค่าที่ขาดหายไปโดยกำหนด method='ffill' หมายถึงการเติมค่าที่ขาดหายไปโดยใช้ค่าที่อยู่ก่อนหน้าที่ไม่ขาดหายไป (forward fill) โดยค่าที่อยู่ก่อนหน้าจะถูกนำมาใช้เติมลงไปในพื้นที่ที่ขาดหายไป ชุดข้อมูล นั้นหมายความว่าค่าที่ขาดหายไปจะถูกเติมด้วยค่าที่อยู่ก่อนหน้าที่ไม่ขาดหายไป ชุดข้อมูล

ขั้นตอนนี้รวมถึงการปรับปรุงข้อมูลเพื่อสร้าง ชุดข้อมูลย่อยที่ถูกแบ่งออกมาจากชุดข้อมูลหลัก (ชุดข้อมูลย่อย) หรือ sub-sequences ของข้อมูล 60,000 ช่วงเวลาสำหรับหนึ่งคลาสถูกแบ่งเป็น sub-sequences ขนาด 500 ช่วงเวลา ซึ่งให้ผลลัพธ์เป็นประมาณ 112 ชุดข้อมูลย่อยสำหรับแต่ละคลาส นี้จะถูกเรียกว่า "สัญญาณ" (signals) ดังนั้นผู้วิจัยมีทั้งหมด 676 สัญญาณ (คือประมาณ 112 สัญญาณ \times 6 คลาส) โดยจำนวนทั้งหมดของ 60,000 ช่วงเวลา, เป็นอัตราการเก็บข้อมูลทุก ๆ 50 มิลลิวินาที, และ 500 ช่วงเวลาสำหรับแต่ละสัญญาณได้รับการตกลงจากวิศวกรเพื่อให้มีความเชื่อถือทางวิชาการและมีความเหมาะสมสัญญาณ 676 จำนวนถูกแบ่งสุ่มเป็นชุดข้อมูลการฝึก 80% และชุดข้อมูลทดสอบอีก 20% จึงได้ข้อมูลสำหรับการฝึกจำนวน 540 สัญญาณ และข้อมูลสำหรับการทดสอบอีก 136 สัญญาณ

นอกจากนี้ ในขั้นตอนการประมวลผลข้อมูลผู้วิจัยได้ดำเนินการเตรียมข้อมูลให้อยู่ในรูปของคุณลักษณะทางสถิติ หรือ Statistical feature เพื่อที่จะนำข้อมูลที่ประมวลผลในการหาคุณลักษณะแล้วเป็นการเพิ่มข้อมูลให้กับการเรียนรู้ของเครื่อง 3 เทคนิคที่ได้กล่าวข้างต้นไว้

การทำข้อมูลดิบเป็นข้อมูลคุณลักษณะทางสถิติ เป็นกระบวนการที่เรานำข้อมูลดิบที่เก็บมาจากเซ็นเซอร์หรือแหล่งข้อมูลอื่น ๆ และคำนวณค่าสถิติทางคณิตศาสตร์เพื่อสร้างคุณลักษณะทางสถิติเหล่านั้นขึ้นมา เพื่อนำไปใช้ในการฝึกโมเดลปัญญาประดิษฐ์ หรือการวิเคราะห์ข้อมูลต่อไป

นี่คือคำอธิบายของแต่ละค่าทางสถิติที่ถูกใช้งานบ่อยในการแปลงข้อมูลดิบเป็น statistical feature:

1. Kurtosis: วัดรูปร่างของการกระจายของข้อมูล เพื่อวิเคราะห์ว่ามีการกระจายแบบแหลมเข้าหรือแหลมออก
2. Skewness: วัดความเบ้หมุนของการกระจายของข้อมูล เพื่อวิเคราะห์ว่าข้อมูลมีการเบี่ยงที่ซ้ายหรือขวา
3. IRQ (Interquartile Range): ช่วงระหว่างควอไทล์เปอร์เซ็นต์ของข้อมูล คือค่าที่อยู่ระหว่าง percentiles 25% ถึง 75%
4. Mean: ค่าเฉลี่ยของข้อมูล
5. Median: ค่ากลางของข้อมูล
6. Min: ค่าต่ำสุดของข้อมูล
7. Max: ค่าสูงสุดของข้อมูล
8. Std (Standard Deviation): ค่าเบี่ยงเบนมาตรฐานของข้อมูล เพื่อวัดการกระจายของข้อมูล
9. Var (Variance): ค่าความแปรปรวนของข้อมูล เป็นตัวบ่งชี้ว่าข้อมูลกระจายมากน้อยเพียงใด
10. RMS (Root Mean Square): ค่าเฉลี่ยที่ยกกำลังสองของข้อมูล ที่ใช้สำหรับวัดความสัมพันธ์ระหว่างค่าต่าง ๆ ในช่วงเวลาหรือการจำลองข้อมูล

การแปลงข้อมูลดิบเป็น statistical feature ช่วยให้เราสามารถนำข้อมูลมาวิเคราะห์และใช้ในการฝึกโมเดลปัญญาประดิษฐ์ได้อย่างมีประสิทธิภาพในการทำนายหรือวิเคราะห์ข้อมูลต่อไป

การทำ statistical feature สามารถมีข้อดีดังนี้:

1. การลดข้อมูล: การแปลงข้อมูลเป็น statistical feature ช่วยลดขนาดข้อมูลลงได้ เนื่องจากคุณสมบัติทางสถิติเพียงไม่กี่ตัวสามารถแสดงผลข้อมูลได้ครอบคลุม ซึ่งทำให้ง่ายต่อการจัดเก็บและประมวลผลข้อมูล
2. ลดความซับซ้อนในการตัดสินใจ: การใช้คุณสมบัติทางสถิติเป็นตัวแทนของข้อมูลช่วยลดความซับซ้อนในการตัดสินใจ เนื่องจากข้อมูลจะถูกแทนที่ด้วยค่าที่สามารถตีความได้ง่ายและสื่อความหมายได้ชัดเจน

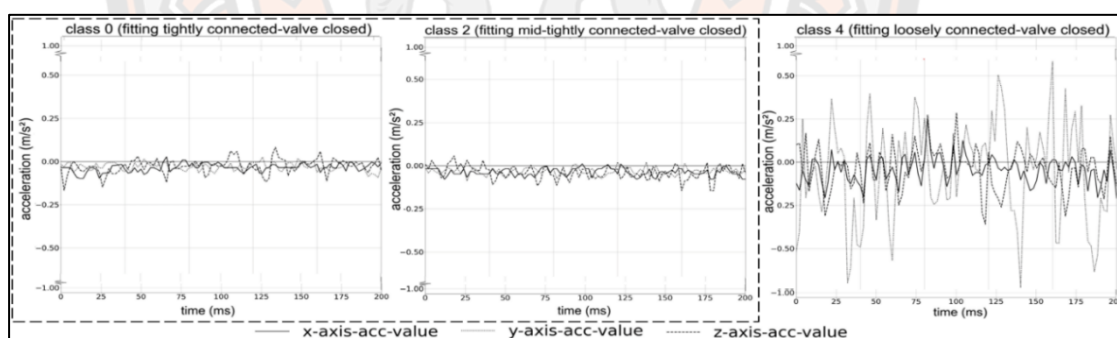
3. การเปรียบเทียบข้อมูลง่ายขึ้น: การใช้คุณสมบัติทางสถิติทำให้ง่ายต่อการเปรียบเทียบข้อมูลระหว่างกัน โดยสามารถใช้ค่าเฉลี่ยหรือค่าทางสถิติอื่น ๆ เพื่อเปรียบเทียบความแตกต่างของข้อมูล

4. ลดความสูญเสียข้อมูล: การทำ statistical feature ช่วยลดความสูญเสียข้อมูลที่อาจเกิดขึ้นในข้อมูลที่มี ความเป็นเดียวกัน โดยการสร้างค่าทางสถิติที่สามารถแทนที่ข้อมูลได้ ซึ่งทำให้สามารถเก็บรักษาข้อมูลสำคัญได้ใน ข้อมูลที่ลดขนาดลง

5. เพิ่มประสิทธิภาพในการฝึกโมเดล: การใช้ statistical feature ช่วยลดความซับซ้อนของข้อมูลและช่วยให้ โมเดลเรียนรู้ได้ง่ายขึ้น โดยลดการฝึกโมเดลที่มีข้อมูลมากมายและซับซ้อนลง ซึ่งทำให้เพิ่มประสิทธิภาพและความ ถูกต้องของโมเดล

ขั้นตอนที่ 3: ขั้นตอนการวิเคราะห์ข้อมูล

สัญญาณเหล่านี้ ซึ่งเป็นข้อมูลหลักสำหรับการทดสอบของผู้วิจัย ได้รับการวิเคราะห์ด้วยกระบวนการที่ ไม่ใช่เทคนิคเรียนรู้ของเครื่องจักรเช่นการแสดงผลข้อมูล นี้ช่วยให้ผู้วิจัยรู้จักพฤติกรรมของข้อมูลที่ผู้วิจัยเก็บและ ประมวลผลไว้ ผู้วิจัยจึงสามารถแยกคลาสทั้งหมดจากกันได้ด้วยการแสดงผลข้อมูลหลักนี้ ในจุดนี้ผู้วิจัยสามารถ ตรวจสอบข้อมูลทางภาพได้และผู้วิจัยพบว่ามีความแตกต่างมากระหว่างคลาส ตามที่แสดงในรูปที่ 3 และ 4 ผู้วิจัย พบว่าผู้วิจัยไม่สามารถแยกแยะระหว่างคลาส 0 และคลาส 2 ได้



รูป 4 การแสดงข้อมูล (วาล์วปิด) (Data Plotting (valve closed))

Figures 3 และ 4 เป็นตัวอย่างของข้อมูลที่เก็บรวบรวมและประมวลผลในแต่ละคลาสของทั้ง 6 คลาส แต่ละคลาสมีกราฟเส้นทางสามเส้นในแต่ละรูปภาพ แกน x ของแต่ละกราฟแทนช่วงเวลาจาก 0 ms ถึง 200 ms. แกน Y แทนค่าการสั่นสะเทือนที่สร้างจากระบบท่อทดลอง (รูป 2) ที่วัดโดยเซ็นเซอร์ ช่วงของค่าสั่นสะเทือนคือ -1 ถึง +1 ในหน่วยเมตรต่อวินาทีกำลังสอง (m/s^2) นี่เป็นผลมาจากการทำให้ค่าสั่นสะเทือนแท้จริงเป็นระดับความปกติ กราฟ ทั้ง 3 เส้นคือค่าการเร่งแรงในแนวแกน X (x-axis-acc-value), Y (y-axis-acc-value), และ Z (z-axis-acc-value) ตามลำดับ

จากรูปที่ 3 ผู้วิจัยสามารถเห็นว่าทั้งสามเส้นของกราฟของ Class 0 (การเชื่อมต่อ Fitting แบบแน่นหนา พร้อมวาล์วปิด) แตกต่างมากจากสามเส้นของกราฟของ Class 4 (การเชื่อมต่อ Fitting แบบหลวมเล็กน้อยพร้อมวาล์ว

ปิด) อย่างไรก็ตาม ผู้วิจัยไม่สามารถแยกข้อมูลของ Class 0 จากข้อมูลของ Class 2 ได้ เนื่องจากรูปร่างของกราฟของทั้งสองคลาสมีลักษณะเกือบเหมือนกัน นั่นหมายความว่าเนื่องจากผู้วิจัยไม่สามารถแยกแยะระหว่างสถานการณ์ปกติของการเชื่อมต่อแน่นหนาจากสถานการณ์ของการเชื่อมต่อแน่นหนาที่เกือบจะเสี่ยงที่จะเสียได้ ผู้วิจัยจึงต้องสร้างแอปพลิเคชันเพื่อแจ้งให้วิศวกรทราบเมื่อ (i) การเชื่อมต่อแน่นหนา, ซึ่งเป็นสถานการณ์ปกติ (Class 0), และ (ii) การเชื่อมต่อแน่นหนา (Class 4), ซึ่งแทนทั้งสถานการณ์ที่เสียหายและการเสี่ยงที่จะเสียได้คือ (จะเสียได้)

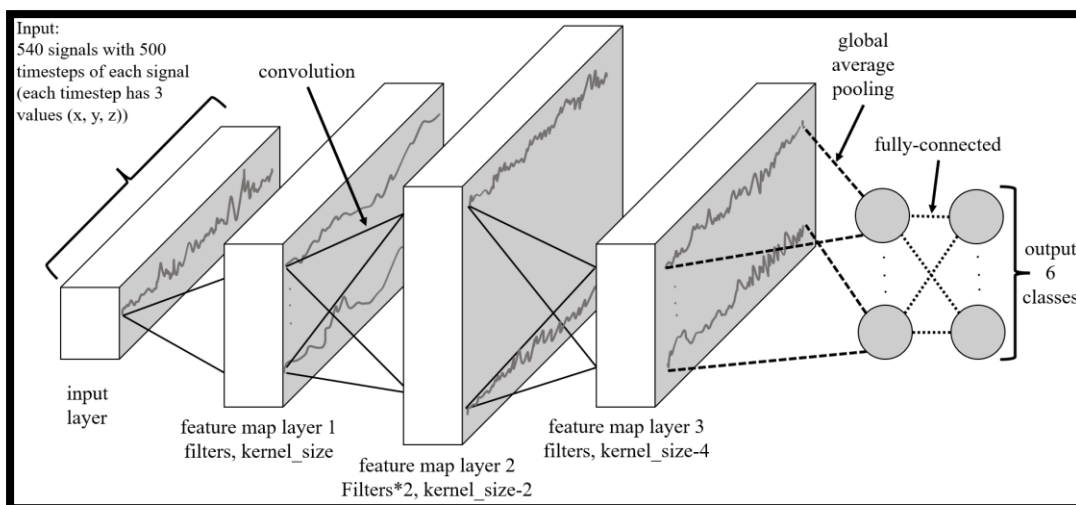
ดังนั้นจึงสามารถจำแนกความรุนแรงของการรั่วไหลของอากาศจากข้อต่อได้เพียงสองส่วนเท่านั้น

(Class 0) การเชื่อมต่อของ Fitting แบบแน่นหนา หรือสถานการณ์ปกติ และ (Class 4) การเชื่อมต่อที่หลวมหรือสถานการณ์ล้มเหลว เพื่อให้แอปพลิเคชันมีประโยชน์มากขึ้น แอปพลิเคชันจำเป็นต้องเตือนวิศวกรเกี่ยวกับสถานการณ์ความล้มเหลวก่อนที่จะเกิดขึ้น ในบริบทที่กำหนด สถานการณ์ที่เกิดขึ้นก่อนที่สถานการณ์ความล้มเหลวจะเกิดขึ้นคือเมื่อ Fitting ถูกเชื่อมต่อแบบมีความแน่นหนากลาง ซึ่งแทนสถานการณ์ใกล้เคียงล้มเหลว (Class 2) ดังนั้นผู้วิจัยต้องการเพียงแค่จำแนก Class 0 จาก Class 2 และเพื่อที่จะทำได้นั้นผู้วิจัยต้องใช้เทคนิคของการเรียนรู้ของเครื่อง (ML) ดังนั้นผู้วิจัยได้ตรวจสอบและเลือกเทคนิค ML ที่เหมาะสมสำหรับการจำแนกข้อมูลที่ผ่านการประมวลผลแล้วของผู้วิจัยอีกครั้ง ผู้วิจัยค้นคว้าได้ตัดสินใจ, Random Forest (RF), Support Vector Machine (SVM), และ 1D-CNN เพื่อจำแนกข้อมูลที่ผ่านการประมวลผลแล้วของผู้วิจัย

ขั้นตอนที่ 4: การฝึกโมเดล

ผู้วิจัยทำการฝึกโมเดลด้วยเทคนิคการเรียนรู้ของเครื่องทั้งสี่เทคนิคที่ผู้วิจัยเลือกไว้ คือ ต้นไม้ตัดสินใจ, เทคนิคป่าไม้สุ่ม, เครื่องมือเวกเตอร์สนับสนุน, และ 1D-CNN ข้อมูลที่ผ่านการประมวลผลแล้วถูกนำเข้ากระบวนการสกัดลักษณะทางสถิติด้วย TSFEL ลักษณะที่สกัดออกมานั้นถูกนำมาใช้เป็นข้อมูลนำเข้าสำหรับโมเดลเทคนิคต้นไม้ตัดสินใจ เทคนิคป่าไม้สุ่ม และ เครื่องมือเวกเตอร์สนับสนุนข้อมูลที่ไม่ได้ทำการเตรียมจัดทำเป็นข้อมูลทางสถิติจะให้กับโมเดล 1D-CNN โดยตรง โมเดล 1D-CNN จะนำข้อมูลตรงเข้ามาในรูปแบบที่ได้จากขั้นตอนการประมวลผลที่ผ่านการจัดสรรคุณลักษณะทางสถิติด้วย TSFEL และใช้เป็นข้อมูลป้อนเพื่อทำการฝึกโมเดล เพื่อวัตถุประสงค์ในการทำ Cross-validation, ข้อมูล validation เป็นส่วนหนึ่งของข้อมูลฝึกใช้ 10% ที่เหมือนกันสำหรับทุกเทคนิค

โมเดลการฝึก CNN ประกอบด้วย 3 ประเภทหลักของเลเยอร์: เลเยอร์ที่ทำการคอนโวลูชัน, เลเยอร์ที่ทำการ Global Average Pooling, และ เลเยอร์ที่ทำการ Softmax Classification [15]. เลเยอร์ที่หนึ่งและสองของ CNN ใช้สำหรับการสกัดลักษณะของข้อมูลที่ถูกเก็บรวมไว้ ทำให้มีทั้งหมด 3 เลเยอร์หลัก ที่ใช้ในการสร้างโมเดล CNN ที่นำมาใช้ในการวิจัยของผู้วิจัย เลเยอร์นี้สามารถสร้างตัวแยกประเภทแบบ softmax fully-connected classifier ที่มีจำนวน neuron ทั้งหมด c โดยที่ c คือจำนวนคลาสในชุดข้อมูลของผู้วิจัย เลเยอร์ทั้งสามนี้รวมกันเป็นกระบวนการหลักของการทดลองของผู้วิจัย



รูป 5 โครงสร้างของ 1D-CNN ที่นำมาใช้ในกระบวนการฝึกอบรม

รูป 5 แสดงโครงสร้างของ 1D-CNN ที่นำมาใช้ในกระบวนการฝึกอบรมในทดสอบของผู้วิจัย ปรับเปลี่ยนมาจาก [15] สถาปัตยกรรมมีทั้ง Input และ Output Input ที่ใส่เข้าไปที่ชั้น input คือสัญญาณ 676 สัญญาณที่มี 500 ไทม์สตีปของแต่ละสัญญาณ (แต่ละไทม์สตีปมีค่าแรงของแกน x, y, และ z ทั้ง 3 แกน) สถาปัตยกรรมนี้ให้ผลลัพธ์เป็นการกระจายความน่าจะเป็นเกี่ยวกับคลาสทั้ง 6 คลาสที่เป็นไปได้ (คลาส 0, คลาส 1, คลาส 2, คลาส 3, คลาส 4, และ คลาส 5) ในชุดข้อมูลการฝึกอบรม สามารถเรียกการกระจายนี้ว่าตัวจัดการเชื่อมโยงอย่างนุ่มนวลหรือตัวจัดการเชื่อมโยงของ 1D-CNN ที่ฝึกอบรมแล้ว

จากรูป 5 ผู้วิจัยยังสามารถเห็นได้ว่าสถาปัตยกรรมประกอบด้วย (i) ทั้งหมดสามชั้นการประมวลผลทางคอนโวลูชัน (ชั้น 1-3), (ii) ชั้นการทำความเข้าใจโดยเฉลี่ยทั่วโลก (ชั้น 4), และ (iii) ชั้นการจัดแยกแบบ softmax (ชั้น 5) ผู้วิจัยได้นำโครงสร้างและการตั้งค่าของเลเยอร์การประมวลผลทางด้านคอนโวลูชันทั้งหมดและพารามิเตอร์ไฮเปอร์พารามิเตอร์ของเลเยอร์เหล่านั้น (ตามที่แสดงในตาราง 9 การปรับแต่งไฮเปอร์พารามิเตอร์ของแต่ละเทคนิค) มาจากงานก่อนหน้า [15] ดังนั้น ผู้วิจัยได้ตั้งค่าเริ่มต้นที่แน่นอนสำหรับ 'filters' และค่าอื่น ๆ สำหรับขนาดของ kernel ในเลเยอร์ที่ 1 จากนั้นในเลเยอร์ที่ 2 ผู้วิจัยได้เพิ่มค่า 'filters' เป็นครึ่งหนึ่งของค่าจากเลเยอร์ที่ 1 (filters*2) และหารค่าขนาดของ kernel จากค่าในเลเยอร์ที่ 1 (kernel_size-2) ในที่สุด ในเลเยอร์ที่ 3 ได้ตั้งค่า filters ให้เป็นค่าเดียวกับในเลเยอร์ที่ 1 และ kernel size ถูกลดลง 4 จากค่าในเลเยอร์ที่ 1 (kernel_size-4)

ตาราง 9 การปรับแต่งไฮเปอร์พารามิเตอร์ของแต่ละเทคนิค

เทคนิคการเรียนรู้ของเครื่อง	ไฮเปอร์พารามิเตอร์	ช่วงของข้อมูล
เทคนิคต้นไม้ตัดสินใจ	criterion	['gini', 'entropy']
	splitter	['best', 'random']
	max_depth	[10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None]
	min_samples_split	[2, 5, 10]
	min_samples_leaf	[1, 2, 4]
เทคนิคป่าไม้สุ่ม	n_estimators	[200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000]
	max_features	['auto', 'sqrt']
	max_depth	[10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None]
	min_samples_split	[2, 5, 10]
	min_samples_leaf	[1, 2, 4]
เครื่องมือเวกเตอร์สนับสนุน	bootstrap	[True, False]
	C	[2.0, 2.8, 3.6, 4.4, 5.2, 6.0, 6.800, 7.600, 8.4, 9.2, 10.0]
	kernel	['linear', 'poly', 'rbf', 'sigmoid']
	gamma	[0.1, 0.19, 0.28, 0.37, 0.4599, 0.5499, 0.64, 0.73, 0.82, 0.9099, 1.0]
	probability	[True, False]
	degree	[3, 4, 5, 6, 7, 8, 9, 10]
	coef0	[2.0, 2.8, 3.6, 4.4, 5.2, 6.0, 6.8, 7.6, 8.4, 9.2, 10.0]
1D-CNN	filters	[256, 128, 64, 32, 16, 8]
	kernel_size	[15, 13, 11, 9, 7]

ตารางที่ 2 แสดงการปรับไฮเปอร์พารามิเตอร์ของแต่ละเทคนิค ML สำหรับ DT, RF และ SVM รายละเอียดของไฮเปอร์พารามิเตอร์ทั้งหมดถูกดัดแปลงจาก [23] และไฮเปอร์พารามิเตอร์สำหรับ 1D-CNN ถูกดัดแปลงจาก [15] ขั้นตอนที่จะสร้างโมเดล DT, RF, SVM และ 1D-CNN ที่ได้รับการฝึกอบรมโมเดล ดังที่แสดงโดยกล่องจุดสีจุดในรูปที่ 2 ผลลัพธ์ของขั้นตอนนี้ระบุไว้ในส่วนที่ 3 และอภิปรายในส่วนที่ 4

ขั้นตอนที่ 5: การคำนวณความแม่นยำของการทำนาย

ในขั้นตอนที่ 5, การทำนายจากรูปแบบที่ฝึกอบรมไว้สี่แบบถูกนำมาเป็นข้อมูลนำเข้าและระดับความแม่นยำของแต่ละโมเดลถูกสร้างขึ้น ผู้วิจัยพบว่าความแม่นยำของทั้ง 4 โมเดลที่ฝึกอบรมคือ DT (81.61%), RF (88.23%), SVM (87.50%), และ 1D-CNN (86.30%) ถูกพิจารณาว่าไม่ยอมรับได้ทั้งหมด ความแม่นยำที่ไม่ยอมรับนี้เกิดจากพฤติกรรมที่ไม่ธรรมชาติของข้อมูลที่ผู้วิจัยเก็บได้ ซึ่งผู้วิจัยสามารถระบุได้เมื่อเปรียบเทียบข้อมูลทดสอบกับข้อมูลจริงจากเส้นผลิตของโรงงาน ในขั้นตอนที่หกของการทดลองนี้ รายละเอียดทั้งหมดเกี่ยวกับพฤติกรรมข้อมูลที่ไม่ธรรมชาติและธรรมชาติจะถูกพูดถึง

ตาราง 10 ตารางผลการฝึกอบรมโมเดลกับข้อมูลดิบ

เทคนิคการเรียนรู้ของเครื่อง	ความแม่นยำกับข้อมูลดิบ
เทคนิคต้นไม้ตัดสินใจ	81.61%
เทคนิคป่าไม้สุ่ม	88.23%
เครื่องมือเวกเตอร์สนับสนุน	87.50%
1D-CNN	86.30%

ในขั้นตอนที่หกนี้เพื่อเพิ่มความแม่นยำในการทำนาย ผู้วิจัยจำเป็นต้องเพิ่มข้อมูลที่ได้มาจากขั้นตอนที่ 1 และข้อมูลที่ได้จากขั้นตอนที่ 2 ให้มีความใกล้เคียงกับข้อมูลจริงที่สุด

ขั้นตอนที่ 6: การเพิ่มและปรับแต่งข้อมูล

ในขั้นตอนที่ 6 นี้จะทำการเพิ่มข้อมูล (data augmenting) ของข้อมูลที่ผ่านมากระบวนการทำข้อมูลพร้อม (preprocessed data) จำไว้ว่า คลาส 0 และ คลาส 1 รวมถึงข้อมูลจาก Fitting ที่เชื่อมต่อแน่นหนา, คลาส 2 และ คลาส 3 รวมถึงข้อมูลจากฟิตติ้งที่เชื่อมต่อแน่นกลาง, และ คลาส 4 และ คลาส 5 รวมถึงข้อมูลจากฟิตติ้งที่เชื่อมต่อแน่นหนา โลงเล็กน้อย ข้อมูลที่ถูกประมวลผลแล้วถูกแบ่งเป็นข้อมูลเซตการฝึกฝน 80% และข้อมูลเซตการทดสอบ 20%

จากขั้นตอนที่ 4 (การฝึกโมเดล) และถูกแสดงในรูปที่ 5 ทั้ง Input และ Output และ Input ที่เข้าไปที่เลเยอร์นำเข้าของข้อมูลการฝึกฝนคือข้อมูลเซตการฝึกฝน ข้อมูลนี้ประกอบไปด้วยสัญญาณ 3,000 สัญญาณที่แต่ละสัญญาณประกอบไปด้วย 500 ไทม์สเต็ป

กระบวนการปรับแต่งข้อมูล

การทำ data augmenting เป็นกระบวนการที่ใช้เพื่อสร้างข้อมูลใหม่โดยใช้ข้อมูลที่มีอยู่แล้ว เพื่อเพิ่มประสิทธิภาพในการฝึกโมเดลปัญญาประดิษฐ์ โดยการสร้างข้อมูลใหม่จะใช้เทคนิคต่างๆ เช่น การแปลงภาพด้วยการหมุนหรือเลื่อนภาพ การเพิ่มสัญญาณรบกวนบนภาพ เปลี่ยนแปลงความสว่าง หรือการทำสเกล เป็นต้น ในบางกรณี เช่น การทำงานกับข้อมูลตำแหน่งเชิงพื้นที่ เราอาจใช้เทคนิคการเพิ่มข้อมูลโดยการสร้างตำแหน่งที่ไม่แท้จริง เพื่อเพิ่มประสิทธิภาพในการฝึกโมเดล เช่น การสร้างรูปภาพที่มีการปรับเปลี่ยนตำแหน่งของวัตถุหรือการเพิ่มเสียงที่มีการเจริญเติบโต เป็นต้น การทำ data augmenting ช่วยลดโอกาสในการเกิด overfitting โดยทำให้โมเดลสามารถทำนายข้อมูลที่มีลักษณะที่หลากหลายมากยิ่งขึ้น ซึ่งทำให้มีประสิทธิภาพในการทำนายข้อมูลใหม่ที่ไม่เคยเห็นมาก่อนได้ดีขึ้น

ในงานวิทยานิพนธ์นี้ได้ทำปรับแต่งข้อมูลเพื่อสนับสนุนและช่วยเหลือกระบวนการเก็บข้อมูลที่ไม่สามารถเก็บข้อมูลให้ตรงตามความเป็นจริงว่าข้อมูลส่วนใดนี้ได้มีการเปิดวาล์วหรือปิดวาล์วอยู่ ดังนั้นผู้วิจัยจึงได้เลือกวิธีการทำปรับแต่งข้อมูลเพื่อสร้างข้อมูลที่มีให้คล้ายกับสภาพความเป็นจริง

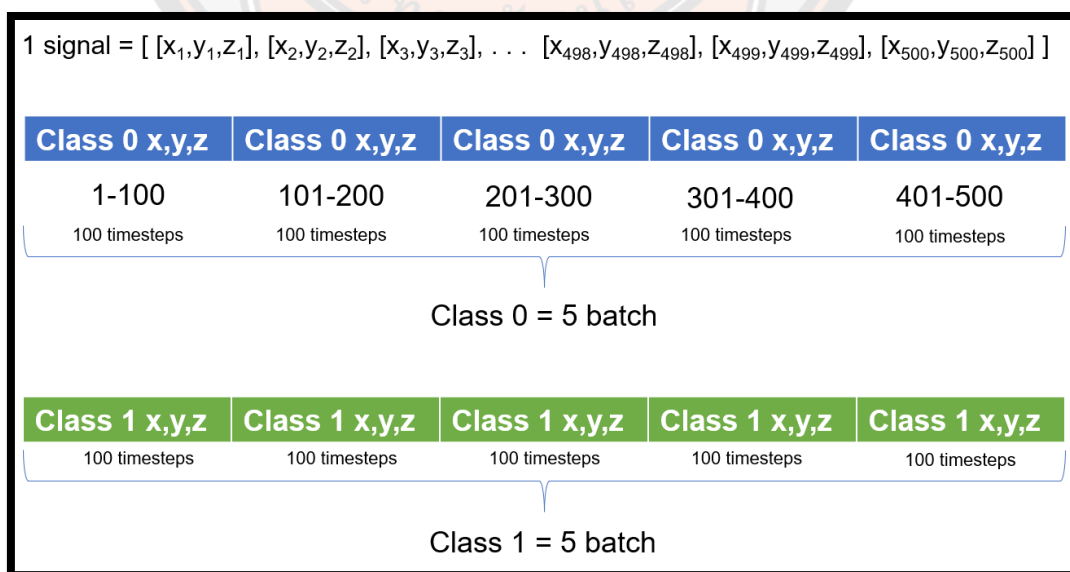
เพื่ออธิบายกระบวนการปรับแต่งข้อมูลแต่ละสัญญาณแทนข้อมูล x , y , และ z ของ 100 ระยะเวลาของคลาสที่แทนด้วยสตริง 5 หลัก ถ้าผู้วิจัยสมมติว่าผู้วิจัยมีสัญญาณ '00000' เป็นหนึ่งข้อมูลสัญญาณของ Class 0 ผู้วิจัยจะเรียกสัญญาณนี้ว่า 'signal s0' ซึ่งมี 500 ขั้นตอนหรือก็คือ 5 'ชุด' ของ 100 ขั้นตอน ดังนั้น ตัวเลข 0 แรกในสัญญาณ s0 นี้แทนข้อมูล x , y , และ z ของ Class 0 ในชุดแรกของ 100 ขั้นตอน ข้อมูลของชุดนี้ที่มี 100 ขั้นตอนสามารถเข้าใจได้ว่าเป็น $[x_0, y_0, z_0], [x_1, y_1, z_1], \dots, [x_{99}, y_{99}, z_{99}]$ ดังนั้นข้อมูลของชุดที่สองสามารถเข้าใจได้ว่าเป็น $[x_0, y_0, z_0], [x_1, y_1, z_1], \dots, [x_{99}, y_{99}, z_{99}]$ ของชุดที่สอง ข้อมูลของชุดที่สองของ 100 ชุดนี้สามารถมองเห็นได้เช่นกันเป็น $[x_{100}, y_{100}, z_{100}], [x_{101}, y_{101}, z_{101}], \dots, [x_{199}, y_{199}, z_{199}]$ การเข้าใจว่าแต่ละสัญญาณ s0 หรือ '00000' มีสัญญาณสำหรับ 500 ช่วงเวลา และทุกช่วงเวลามีข้อมูล x , y , และ z ที่แตกต่างกัน ทั้งนี้ ข้อมูลทั้งหมดของสัญญาณ s0 หรือ '00000' แทนช่วงเวลา 500 และสามารถมองเห็นได้เป็น $[x_0, y_0, z_0], [x_1, y_1, z_1], \dots, [x_{499}, y_{499}, z_{499}]$ และข้อมูลนี้เป็นข้อมูลเพียงหนึ่งสัญญาณจากสัญญาณทั้งหมดของข้อมูลการฝึกอบรมและสัญญาณนี้เป็นของ Class 0 เท่านั้น

ดังนั้นข้อมูลที่รวบรวมได้ซึ่งขณะนี้ใช้เป็นข้อมูลการฝึกอบรมจากทั้ง 6 คลาส จะอยู่ในรูปแบบของสัญญาณที่แสดงเป็น เช่น 00000, 11111, 22222, 33333, 44444 และ 55555

ข้อความที่อธิบายถึงสัญญาณ s0 ด้านบนสามารถใช้กับทุกสัญญาณในชุดทั้งหมดสัญญาณ แต่ละสัญญาณเป็นตัวแทนของพฤติกรรมของข้อมูลที่แตกต่างกันออกไปในสถานการณ์ต่างๆ ของการทดสอบครับ และสัญญาณหกลักษณะเหล่านี้แทนพฤติกรรมที่ไม่ธรรมชาติของข้อมูลที่ถูกผู้วิจัยเก็บมาจากเส้นทางการทดสอบ หากผู้วิจัยกังวลว่าสองสัญญาณของข้อมูลที่เก็บมาอยู่ในสถานการณ์เดียวกัน ผู้วิจัยสามารถมีคู่สามคู่ของสัญญาณทั้งหมดสัญญาณในข้อมูลที่เก็บมาได้ ทุก 100 ช่วงเวลาของสัญญาณจากสัญญาณหกลักษณะสัญญาณสามารถถูกลบ/เพิ่มเติมเพียงแคในสามสถานการณ์ต่อไปนี้ ดังนั้นหลังจากกระบวนการเพิ่มข้อมูลของสัญญาณหกลักษณะแล้ว ข้อมูลสามารถถือว่ามีพฤติกรรมที่เป็นธรรมชาติแล้ว นั่นคล้ายกับพฤติกรรมของข้อมูลจริงบนเส้นผลิต และพร้อมที่จะเป็นข้อมูลนำเข้าสำหรับขั้นตอนที่ 4 (การฝึกโมเดล) อีกครั้ง

ดังนั้น เช่น สองสัญญาณ (00000, 11111) ของคลาส 0 และคลาส 1 ซึ่งทั้งสองเป็นข้อมูลสำหรับสถานการณ์ของการเชื่อมต่อแน่นแน่นของท่อสามารถจะถูกสลับ/เพิ่มขึ้นด้วยขั้นตอน (a) และ (b) (a) ทั้งสองสัญญาณสามารถถูกสลับกัน ดังนั้นผลลัพธ์คือสองสัญญาณ '00000, 11111' สามารถถูกสลับเป็น '00011, 10101' และ (b) ผลลัพธ์ของสัญญาณ '00011' ที่ถูกสลับไปนั้นจะถือเป็นสัญญาณเพิ่มเติมใหม่ของคลาส 0 ขั้นตอน (b) นั้นเกิดขึ้นเนื่องจากผลรวมขององค์ประกอบ '0' ใน '00011' มีค่าเท่ากับ 3 ซึ่งมากกว่าผลรวมขององค์ประกอบ '1' ใน '00011' ซึ่งมีค่าเท่ากับ 2 นั้นถูกต้อง ผลลัพธ์ที่ได้จากการสับเปลี่ยนของ '10101' จะถือเป็นสัญญาณเพิ่มเติมของ Class 1 นั้นเพราะผลรวมของตัวเลข '1' ใน '10101' คือ 3 ซึ่งมากกว่าผลรวมของตัวเลข '0' ใน '10101' ที่เท่ากับ 2 ดังนั้นผู้วิจัยจึงจัดสรร '10101' เป็นสัญญาณเพิ่มเติมของคลาส 1 ในขั้นตอน (c) ผู้วิจัยใช้ขั้นตอน (a) และ (b) กับข้อมูลที่ได้รับการประมวลผลล่วงหน้าของผู้วิจัยสำหรับคลาส 0 และ 1 จนกว่าผู้วิจัยจะมีสัญญาณเพิ่มเติม 500 สัญญาณสำหรับแต่ละคลาส ต่อมาในขั้นตอน (d) ผู้วิจัยใช้ขั้นตอน (a) ถึง (c) สำหรับข้อมูลชุดการฝึกของผู้วิจัยของคลาส 2 และ 3, และคลาส 4 และ 5 ดังนั้น ข้อมูลที่ผู้วิจัยเก็บไว้เดิมทั้งหมด เช่น '00000, 11111, 22222, 33333, 44444, และ 55555' สามารถเพิ่มเติมเป็น เช่น '00011, 11100, 22233, 33322, 44554, และ 45554' ได้

ดังนั้น ผู้วิจัยจึงมีสัญญาณใหม่และถูกเพิ่มเติมทั้งหมด 3000 สัญญาณ (500 สัญญาณต่อคลาสของหกคลาส) เป็นข้อมูลชุดการฝึกใหม่ สำหรับเซตข้อมูลทดสอบ ผู้วิจัยใช้ขั้นตอน 1, 2, 3 และ 4 อีกครั้งกับข้อมูลเหล่านี้ ซึ่งทำให้ได้สัญญาณทดสอบเพิ่มเติมอีก 3000 สัญญาณเป็นข้อมูลเซตทดสอบใหม่ ทั้งสองเซตข้อมูลใหม่จากนั้นถูกนำเข้าไปในขั้นตอนที่ 4 (การฝึกโมเดล) อีกครั้ง ข้อมูลที่เพิ่มเติม (6000 สัญญาณที่ถูกเพิ่มเข้าไป) ควรจะมีพฤติกรรมที่ใกล้เคียงกับข้อมูลจริงที่ได้มาจากเส้นผลิต อย่างไรก็ตาม ในขั้นตอนที่ 5 (การคำนวณความแม่นยำของการพยากรณ์) ข้อมูลที่ถูกเพิ่มเข้ามาใหม่อาจจะมีลักษณะเดียวกันกับชุดข้อมูลเดิมเนื่องจากความสับสนที่เกิดขึ้นในกระบวนการเพิ่มเติมนี้



รูป 6 ตัวอย่างของข้อมูลที่แบ่งออกเป็น 5 ส่วน

Example (3:2) → Class 0 = 100 timesteps 3 batch, Class 1 = 100 timesteps 2 batch.

Random data Class 0 from raw data

100 timesteps, 3 batch



Random data Class 1 from raw data

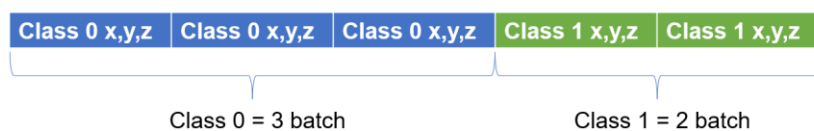
100 timesteps, 2 batch



รูป 7 การส่งข้อมูล และการแบ่งส่วนของข้อมูล

Example (3:2) → Class 0 = 100 timesteps 3 batch, Class 1 = 100 timesteps 2 batch.

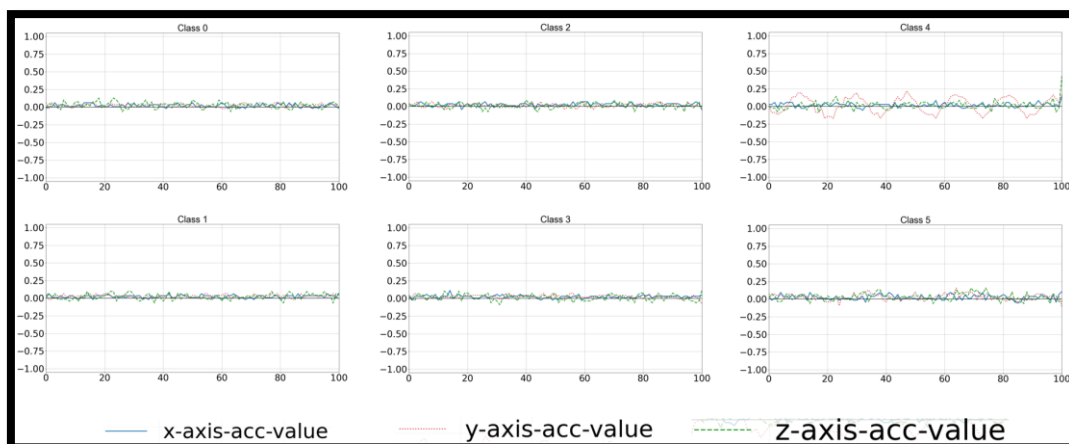
Concatenate data 2 Class (Class 0 and Class 1)



Labeling



รูป 8 การรวมข้อมูลและกำหนดคลาส



รูป 9 ตัวอย่างข้อมูลที่ได้รับการปรับแต่งของแต่ละคลาส

ตาราง 11 ตารางเปรียบเทียบจำนวนข้อมูลดิบและข้อมูลปรับแต่ง

ประเภทข้อมูล	ข้อมูลสำหรับฝึกโมเดล	ข้อมูลสำหรับทดสอบโมเดล
ข้อมูลดิบ	(540, 500, 3)	(136, 500, 3)
ข้อมูลปรับแต่ง	(3,000, 500, 3)	(3,000, 500, 3)

ในขั้นตอนที่ 4 (การฝึกโมเดล) ผู้วิจัยพบว่าความแม่นยำของโมเดลทั้งสี่ตัว คือ DT (81.61%), RF (88.23%), SVM (87.50%), และ 1D-CNN (86.30%) ตามที่ได้กล่าวไว้ข้างต้น ไม่ได้نب่ถึงได้ในระดับที่ยอมรับได้ ดังนั้น ผู้วิจัยทำซ้ำขั้นตอนที่ 4 (การฝึกโมเดล) เพื่อฝึกโมเดลทั้งสี่ตัวอีกครั้งด้วยข้อมูลการเพิ่มเติมที่ใหม่ ผู้วิจัยทำซ้ำขั้นตอนที่ 5 (การคำนวณความแม่นยำของการทำนาย) บนข้อมูลที่เพิ่มเติมอีกครั้งและได้ความแม่นยำอยู่ที่ 99.2% เฉพาะจากโมเดล 1D-CNN ที่ฝึกใหม่ ดังนั้น ผู้วิจัยสามารถข้ามขั้นตอนที่ 6 (การเพิ่มข้อมูล) เนื่องจากโมเดลใหม่มีความแม่นยำที่ยอมรับได้ ในที่สุดผู้วิจัยสามารถไปสู่ขั้นตอนที่ 7 ได้

จากส่วนที่ 2.5 และก่อนการเพิ่มข้อมูล ผู้วิจัยพบว่าทุกรูปแบบที่ได้รับการฝึกอบรมของ DT, RF, SVM, และ 1D-CNN มีความแม่นยำที่ไม่ยอมรับทั้งหมด ผู้วิจัยพบว่าความแม่นยำที่ไม่ยอมรับเกิดขึ้นจากพฤติกรรมที่ไม่ปกติของข้อมูลที่ผู้วิจัยเก็บเมื่อเปรียบเทียบกับข้อมูลจริงจากในการทำงานจริง ผู้วิจัยจึงดำเนินขั้นตอนการเพิ่มข้อมูลที่จากนั้นนำข้อมูลที่เพิ่มเติมมาฝึกฝนที่ใช้ในการได้ราคาของแบบจำลองสี่แบบใหม่อีกครั้ง ความแม่นยำของแบบจำลองทั้งเก่าและใหม่ (แสดงเป็น old%, new%, และ difference%) คือ:

ตาราง 12 ผลของความแม่นยำระหว่างข้อมูลดิบและข้อมูลที่ปรับแต่ง

เทคนิคการเรียนรู้ของเครื่อง	ความแม่นยำกับ ข้อมูลดิบ	ความแม่นยำกับ ข้อมูลที่ปรับแต่ง	ความแม่นยำที่ เพิ่มขึ้น
DT	81.61%	89.23%	7.62%
RF	88.23%	89.56%	1.33%
SVM	87.50%	93.13%	5.63%
1D-CNN	86.30%	99.20%	12.90%

แบบจำลอง 1D-CNN ใหม่มีความแม่นยำสูงสุดและความแตกต่างสูงสุด (+12.9%, 99.2%) ดังนั้นแบบจำลองนี้เหมาะที่จะนำมาใช้ในการตรวจจับรั้วลมจากชิ้นส่วนการเชื่อมต่อในระบบท่อลมบนท่อลมโดยใช้ระบบเซ็นเซอร์แอคเซลโรมิเตอร์สำหรับวิศวกร นำไปใช้ในการติดตั้งแอปพลิเคชันนี้จะถูกพิจารณาในส่วนที่ 4

ขั้นตอนที่ 7: การสร้างคอนฟิวชันเมตริก

ขั้นตอนนี้สร้างคอนฟิวชันเมตริกของโมเดลซึ่งช่วยให้ผู้วิจัยรู้ค่าที่ทำนายผิดพลาดอยู่ที่ไหน โดยเฉพาะ การประเมินความแม่นยำที่ร้อยละ 99.2 ถูกดำเนินการโดยการคำนวณคอนฟิวชันเมตริกเพื่อระบุเหตุผลของความขาดหายของ 0.8% จากความแม่นยำร้อยละ 100

ค่าเหล่านี้สามารถนำมาใช้เป็นแนวทางสำหรับแอปพลิเคชันอื่น ๆ ตามที่จะถูกพูดถึงในส่วนที่ 4 ของรายงาน แอปพลิเคชันเช่นนี้สามารถใช้ในการตรวจจับรอยร้าวของอากาศจากอุปกรณ์ในระบบท่อลมที่ใช้ในการขับเคลื่อนด้วยลม โดยใช้ระบบเซ็นเซอร์แรงเหวี่ยงและแจ้งเตือนวิศวกรเกี่ยวกับรอยร้าวที่พบ

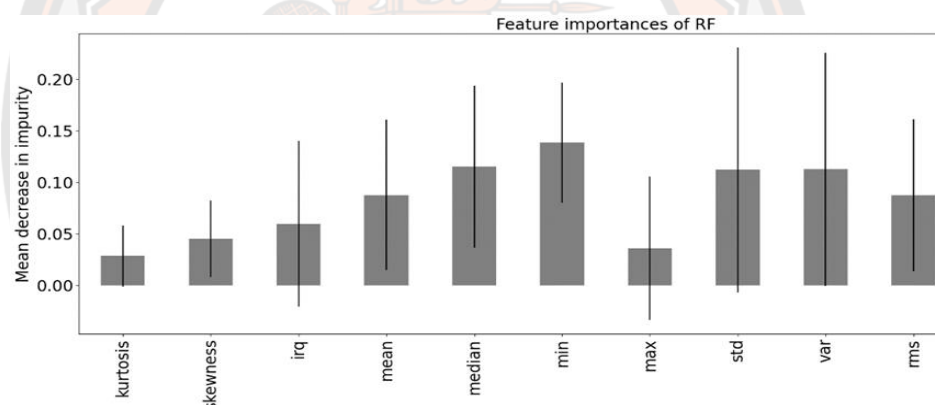
สรุปผลลัพธ์ของขั้นตอนที่ 4, 5, และ 7 โดยละเอียดเนื่องจากตามที่กล่าวในส่วนที่ 2.5.1 ว่าข้อมูลทดลองถูกจัดหมวดหมู่เป็นหกชั้น

ความสำคัญของกระบวนการฝึกโมเดล RF

พวกผู้วิจัยได้ใช้เทคนิคต้นไม้ตัดสินใจ, Random Forests (RF), และ Support Vector Machines (SVM) สำหรับการวิเคราะห์เบื้องต้นของข้อมูลที่ผู้วิจัยได้รับมา รวมทั้งกับ Convolutional Neural Networks (CNN) ด้วย ในส่วนนี้ผู้วิจัยจะศึกษาผลลัพธ์ตัวอย่างที่ได้จากการวิเคราะห์นี้เพื่อเข้าใจข้อมูลที่เก็บได้ โดยเฉพาะผลลัพธ์ที่ได้จากระบบการฝึก RF ที่เรียกว่า "Feature Importances" การวิเคราะห์ Feature Importances ที่ถูกกล่าวถึงในส่วนที่ 2.2 เป็นวิธีที่ให้คะแนนให้กับ

'input features' (คุณลักษณะข้อมูลเข้า) ตามความมีประสิทธิภาพของคุณลักษณะที่ตั้งใจจะทำนายตัวแปรเป้าหมาย [12]

การนำวิธีนี้ไปใช้จะได้ชุดคุณลักษณะที่ละค่า mean decrease in impurity [24] สำหรับแต่ละคุณลักษณะ (feature) ข้อมูลที่ได้จาก Feature Importances สามารถนำมาใช้ในการประเมินความสำคัญของคุณสมบัติแต่ละประการ (feature) ในการทำนายสถานะของตัวระบบที่อลมได้ ซึ่งค่า mean decrease in impurity ที่สูงขึ้นแสดงว่าคุณสมบัตินั้นมีผลต่อการทำนายมากที่สุด ค่า Mean Decrease in Impurity (MDI) ที่มีค่าสูงแสดงถึงความสำคัญของคุณลักษณะนั้น ๆ มากขึ้นในการสร้างแบบจำลองต้นไม้ตัดสินใจหรือในรูปแบบของ Random Forest ที่ใช้ MDI เป็นตัววัดความสำคัญของคุณลักษณะ (feature) ทั้งหมด ซึ่งหมายความว่าคุณลักษณะนั้นมีผลมากในการทำนายข้อมูลที่นำเข้าไปในโมเดล Figure 6 (A) ที่แสดงบางส่วนที่สำคัญของค่าความสำคัญของคุณลักษณะจากการปรับใช้วิธี Feature Importances ในขั้นตอนการฝึก RF ค่าความสำคัญของคุณลักษณะนี้แสดงประสบการณ์ที่ได้จากการคำนวณค่าความสำคัญของคุณลักษณะ คุณลักษณะที่สำคัญที่สุด 4 อันดับแรกของ RF model คือ minimum (min), median, variance (var), และ standard deviation (std) features



รูป 10 คุณลักษณะ ความสำคัญของกระบวนการฝึกอบรมต้นไม้ป่าสุ่ม

เมทริกซ์ความสับสนของโมเดล 1D-CNN

ในส่วนนี้จะอธิบายกระบวนการคำนวณคอนฟิวชันเมตริกเพื่อประเมินความแม่นยำในการจำแนกของโมเดล 1D-CNN ใหม่ของผู้วิจัย ในการสนทนาเกี่ยวกับคอนฟิวชันเมตริกในส่วนที่ 2.2, 2.5.2 และ 3.2, ความแม่นยำ 99.2% ของโมเดล 1D-CNN ใหม่ของผู้วิจัยได้ถูกประเมินโดยการคำนวณคอนฟิวชันเมตริกซึ่งถูกใช้ในการตรวจสอบเหตุผลของความต่างความแม่นยำ 0.8% โมเดลของผู้วิจัยใช้ตัวจำแนก (classifier) เพื่อทำการจัดประเภทข้อมูล และคอนฟิวชันเมตริกถูกนำมาใช้เพื่อตรวจสอบประสิทธิภาพของตัวจำแนกนั้นในขณะที่ข้อมูลทดสอบถูกประมวลผลและจัดประเภท

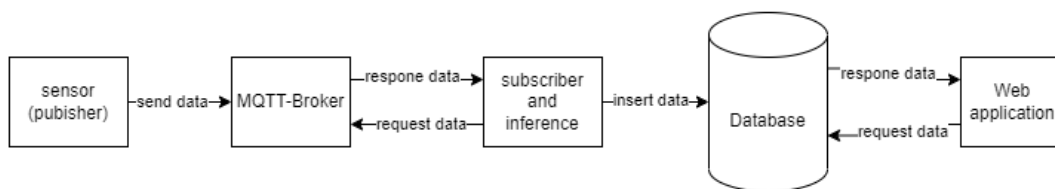
True labels/classes	0	1	2	3	4	5
0	480	8	9	0	0	0
1	0	503	0	0	0	0
2	3	0	475	5	0	0
3	0	0	8	509	0	0
4	0	0	0	0	498	4
5	0	0	0	0	0	498
Predicted labels/classes	0	1	2	3	4	5

รูป 11 คอนฟูชันเมตริก: การทำนายผลการทดสอบโมเดลด้วยเทคนิค 1D-CNN

คอนฟูชันเมตริกนี้มีสองมิติ: คลาสหรือป้ายกำกับจริงของวัตถุและป้ายกำกับที่ทำนายไว้ รูปที่ 6 (B) แสดง คอนฟูชันเมตริกของตัวจัดประการในโมเดล 1D-CNN ใหม่ของผู้วิจัย จากภาพที่ 6 (B) แสดงเมตริกซ์ความสับสนของตัวจำแนกของโมเดล 1D-CNN ใหม่ บนแกน X แสดงป้าย/คลาสที่ทำนายได้ ซึ่งรวมถึง Class 0 ถึง Class 5 แกน Y คือป้าย/คลาสที่แท้จริง รวมถึง Class 0 ถึง Class 5 ผลรวมของเซลล์ทั้งหมดที่ไม่เท่ากับศูนย์คือ 3,000 นี่คือนับจำนวนของสัญญาณทั้งหมดของข้อมูลจากการทดลอง ในการวิเคราะห์เมตริกนี้ ผู้วิจัยสามารถมองที่แกน X และแกน Y ของ Class 0 ก่อน นั่นหมายถึงมีสัญญาณทั้งหมด 480 ที่ถูกจำแนกระดับของ Class 0 ถูกต้องทั้งหมด 480 ทำให้มีความแม่นยำที่ 100% สำหรับ Class 0 ในทดลองนี้ นั่นแสดงว่าโมเดล 1D-CNN ของผู้วิจัยทำนายคลาส 0 ถูกต้อง 480 ครั้งจากทั้งหมด 483 ครั้ง ซึ่งในที่นี้มีความแม่นยำอยู่ที่ 99.38% สำหรับ Class 0 ในการทดลองนี้ จากนั้นผู้วิจัยสามารถมองที่แกน X ที่คลาส 0 อีกครั้งและดูที่แกน Y ที่คลาส 2 จำนวนที่ปรากฏระหว่างการมองที่ทั้งสองขั้นตอนคือ 3 จำนวนนี้แสดงถึงการทำนายผิดคลาส 0 (เป็นคลาส 2) 3 ครั้ง

จากเมตริกนี้ ผู้วิจัยสามารถมองที่แกน X และแกน Y ที่คลาส 4 จำนวนที่ปรากฏระหว่างการมองที่ทั้งสองขั้นตอนคือ 498 และเป็นการทำนายที่ถูกต้องที่สุด/ดีที่สุด ดังนั้น คลาสที่ดีที่สุดคือการทำนายของคลาส 4 ซึ่งมีการทำนายที่ถูกต้อง 498 ครั้งและไม่มีการทำนายที่ผิดพลาดเลย การทำนายที่ผิดพลาดที่น้อยที่สุดคือการทำนายของคลาส 2 ที่มีการทำนายที่ถูกต้อง 475 ครั้ง แต่มีการทำนายที่ผิด 17 ครั้ง (8+9) โดยรวมโมเดลมีการทำนายที่ถูกต้องของคลาส 0 ถึงคลาส 5 คือ 480, 503, 475, 509, 489, และ 498 ครั้ง และการทำนายที่ผิดพลาดของคลาส 0 ถึงคลาส 5 คือ 3, 8, 17 (8+9) 5, 0, และ 4 ครั้งตามลำดับ

แอปพลิเคชัน



รูป 12 สถาปัตยกรรมส่วนแอปพลิเคชัน

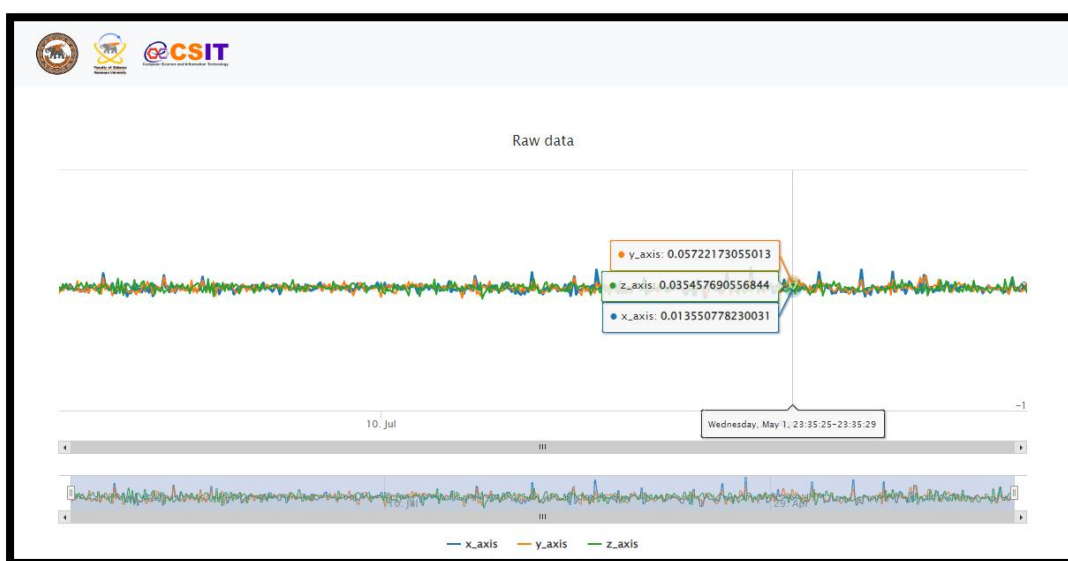
วิศวกรเน้นที่สองประเภทของสถานการณ์รั่วของข้อต่อ 1. สถานการณ์การรั่วของข้อต่อที่มีการเชื่อมแน่นปานกลางและข้อต่อที่หลวม 2. สถานการณ์ที่มีการเชื่อมแน่นอย่างกลางคือเมื่อข้อต่อเชื่อมแน่นอย่างกลาง (คลาส 2 และคลาส 3) ส่วนสถานการณ์ที่ไม่เชื่อมแน่นอย่างชัดเจนคือเมื่อข้อต่อเชื่อมไม่แน่นอย่างชัดเจน (คลาส 4 และคลาส 5) ถ้าผู้วิจัยสมมติว่าโดยปกติแล้ววิศวกรไม่ได้ให้ความสนใจกับสถานการณ์ปกติ (หรือข้อต่อมีการเชื่อมต่ออย่างแน่นหนา) ดังนั้น จากผลลัพธ์ในส่วน 3 และการอภิปรายเกี่ยวกับคลาส 4 และคลาส 5 ที่นี้ ผู้วิจัยสามารถสร้างแอปพลิเคชันที่สามารถแจ้งเตือนวิศวกรของเส้นผลิตถึงสองประเภทของรอยรั่วได้ แอปพลิเคชันสามารถแจ้งเตือนวิศวกรเกี่ยวกับสองสถานการณ์ของรอยรั่ว ด้วยข้อความแจ้งเตือน 2 ประเภทดังนี้ ข้อแรกคือ “near-failure” ซึ่งเกี่ยวกับสถานการณ์ของการต่อที่เชื่อมขัดอยู่ระหว่างกัน (คลาส 2 และคลาส 3) และข้อสองคือ “fail” ซึ่งเกี่ยวกับสถานการณ์ของการต่อที่หลวมพอดี (คลาส 4 และคลาส 5)

การใช้แอปพลิเคชันเป็นไปตามขั้นตอนที่กล่าวถึงได้รับการอ้างอิงจากการใช้เครื่องมือตรวจวัดที่ติดอยู่ที่เส้นผลิต ซึ่งสร้างข้อมูล/สัญญาณจริง (ด้วยค่าต่อลำดับเวลาสำหรับแต่ละช่วงเวลาอย่างต่อเนื่อง) สำหรับการใช้ในแอปพลิเคชัน จากนั้น แอปพลิเคชันสามารถส่งข้อมูลแต่ละช่วงเวลาของ 500 ข้อมูลไปยังโมเดลที่เสนอไว้ในเวลาจริง โมเดลจะจำแนกข้อมูลช่วงเวลา 500 ตัวอย่างนี้เป็นหนึ่งในคลาส 0 ถึงคลาส 5 ได้ ดังนั้นผลการจำแนกของข้อมูลช่วงเวลา 500 ข้อมูลนี้สามารถเป็นค่าเพียงหนึ่งจาก 0, 1, 2, 3, 4 หรือ 5 ซึ่งแทนคลาส 0 ถึงคลาส 5 ได้ หากแอปพลิเคชันสามารถรับข้อมูลช่วงเวลา 5000 ข้อมูลและส่งข้อมูลไปยังโมเดล โมเดลจะให้ผลการจำแนกกลับค่าผลลัพธ์ (เช่น 0, 1, 2, 3, 4, หรือ 5) ของข้อมูลช่วงเวลา 5000 ข้อมูลนี้

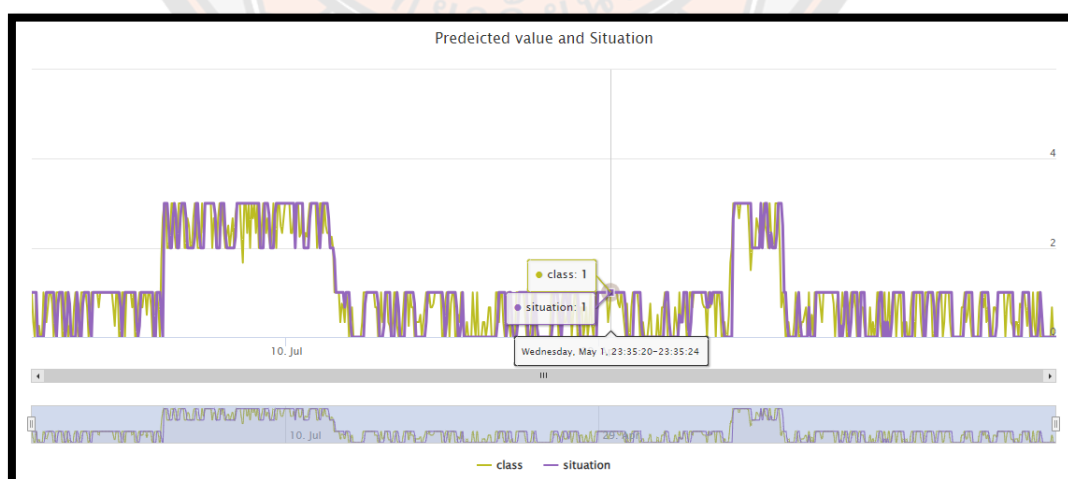
ผลลัพธ์จากโมเดลจะช่วยให้แอปพลิเคชันสามารถระบุถึงสถานะของระบบท่อได้ ว่าเป็นการเหลื่อมแบบ “near-failure” หรือ “fail” ในกรณีของการเชื่อมต่อที่หลวมพอดีหรือหลวมไปแล้ว นอกจากนี้ยังสามารถแจ้งเตือนผู้ดูแลเพื่อดำเนินการแก้ไขข้อบกพร่องในระบบได้อย่างทันทั่วทั้ง ตัวอย่างของผลลัพธ์จำแนกทั้งสิบค่าคือ 3330123401 ดังนั้นแต่ละตัวเลขใน 3330123401 คือผลลัพธ์จำแนกจากโมเดลของแต่ละช่วงเวลา 500 ข้อมูลในชุดข้อมูลช่วงเวลา 5000 ข้อมูลทั้งหมด

สถานการณ์แรกของแอปพลิเคชันคือเมื่อค่าผลการจำแนกสิบค่ามีค่าเป็น 3330123401 ดังนั้นจะมีเลข 0 (Class 0) หรือ 1 (Class 1) อยู่สี่ครั้งใน 3330123401, 2 (Class 2) หรือ 3 (Class 3) อยู่ 5 ครั้งใน 3330123401, และ 4 (Class

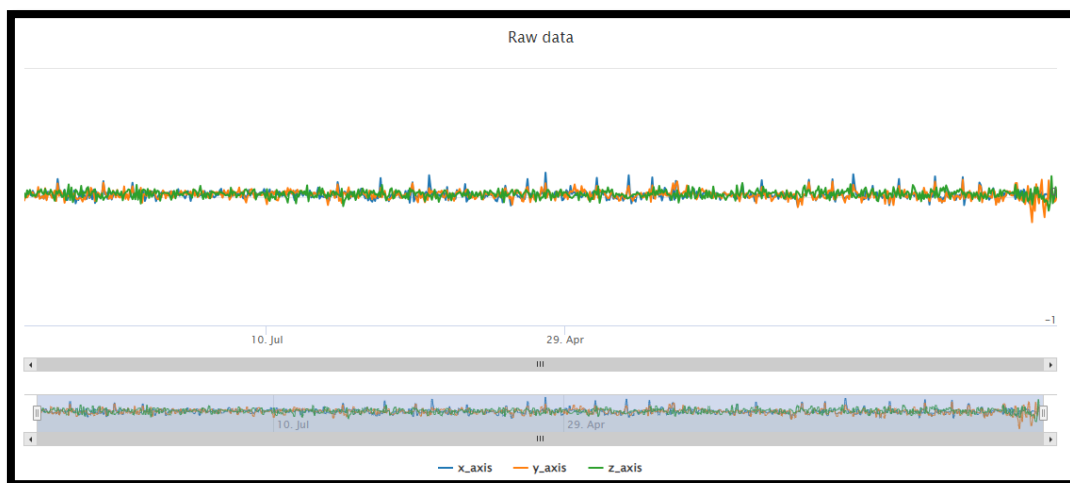
4) หรือ 5 (Class 5) อยู่ 1 ครั้งใน 3330123401 ดังนั้นค่าผลรวมของการเกิดของ Class 0 และ Class 1 คือ 4, Class 2 และ Class 3 คือ 5, และ Class 4 และ Class 5 คือ 1 สัดส่วนของจำนวนการเกิดของ Class 0 และ Class 1 คือ 4/10, Class 2 และ Class 3 คือ 5/10, และ Class 4 และ Class 5 คือ 1/10 โดยที่มีข้อสังเกตว่าข้อความการแจ้งเตือนที่สร้างขึ้นจากแอปพลิเคชันสำหรับวิศวกรสามารถถูกกระตุ้นได้ เช่นจากสามสัดส่วนนั้นมีค่ามากกว่าหรือเท่ากับ 5/10 หรือ 0.5 ดังนั้นในกรณีนี้ ข้อความจะเป็น “near-failure” (สถานการณ์การเชื่อมต่ออย่างหนึ่งกับความพอดี (Class 2 และ Class 3)) เนื่องจากสัดส่วนของ Class 2 และ Class 3 คือ 5/10 (เหนือ) และมากกว่าหรือเท่ากับ 0.5



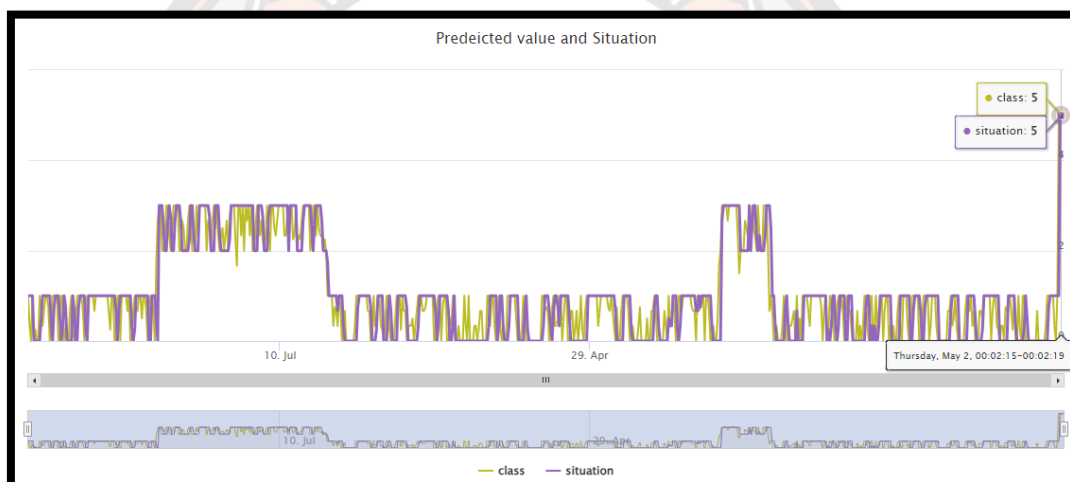
รูป 13 ตัวอย่างแอปพลิเคชันที่แสดงค่าของเซนเซอร์สามแกน



รูป 14 แอปพลิเคชันที่แสดงถึงค่าการทำนายเปรียบเทียบกับค่าจริง



รูป 15 แอปพลิเคชันแสดงข้อมูลดิบของสถานการณ์ล้มเหลวหรือมีลมรั่ว



รูป 16 แอปพลิเคชันแสดงผลการทำนายของโมเดลของสถานการณ์ล้มเหลวที่ได้มา

ในกรณีนี้ ข้อความที่สร้างขึ้นจากแอปพลิเคชันสำหรับวิศวกรบนเส้นผลิตจะเป็น 'ความเสียหาย' (สถานการณ์การเชื่อมต่ออย่างหนึ่งในความหลวมและไม่สามารถใช้งานได้ (Class 4 และ Class 5)) เนื่องจากสัดส่วนของ Class 4 และ Class 5 เท่ากับ $\frac{6}{10}$ และมากกว่าหรือเท่ากับ 0.5 ดังนั้น ค่าเฉลี่ยของจำนวนครั้งของ Class 0 และ Class 1 คือ $\frac{1}{10}$, Class 2 และ Class 3 คือ $\frac{3}{10}$, และ Class 4 และ Class 5 คือ $\frac{6}{10}$ ดังนั้น ตามเงื่อนไขการแจ้งเตือนเดียวกันดังกล่าว ในกรณีนี้ ข้อความที่แสดงคือ 'fail' (สถานการณ์ที่เข้ามั้นแน่นอย่างแข็งแกร่ง (Class 4 และ Class 5)) เนื่องจากสัดส่วนของ Class 4 และ Class 5 เป็น $\frac{6}{10}$ ซึ่งมากกว่าหรือเท่ากับ 0.5 ซึ่งเป็นค่าเกณฑ์ที่กำหนดไว้

โมเดล 1D-CNN ของผู้วิจัยได้จำแนกข้อมูลที่เก็บมาจากระบบท่อทดลองของผู้วิจัยเป็น Class 0, Class 1, Class 2, Class 3, Class 4, และ Class 5 ด้วยความแม่นยำอยู่ที่ 99.2% ดังนั้น โมเดล 1D-CNN จึงมีประสิทธิภาพอย่างมากสำหรับการตรวจจับรอยรั่วของอากาศจากอุปกรณ์คอมพรีเซอร์ในระบบท่อลมโดยใช้ระบบเซ็นเซอร์แรงเหวี่ยง การตรวจจับนี้

สามารถเปิดใช้งานแอปพลิเคชันให้แจ้งเตือนวิศวกรเกี่ยวกับสามสถานการณ์ของอุปกรณ์ในระบบท่อของบริษัทด้วย ข้อความประเภทหลัก 3 ประเภท คือ “normal” (หรืออุปกรณ์เชื่อมต่อแน่นหรือ Class 1 และ Class 2), “near-failure” (หรืออุปกรณ์เชื่อมต่อกลางหนาหรือ Class 2 และ Class 3) และ “fail” (หรืออุปกรณ์เชื่อมต่อหลวมหรือ Class 4 และ Class 5) เหมือนกับการวิเคราะห์ของผู้วิจัยแสดงให้เห็น ข้อมูลที่จัดหมวดหมู่ในคลาส 2 และ 3 เทียบเท่ากับข้อมูลในคลาส 4 และ 5 ซึ่งทำให้แอปพลิเคชันสามารถมองให้สองสถานการณ์เหล่านี้เป็นเหมือนกันได้ นั่นเป็นเพราะว่าการทำให้ท่อตักฟุ้งขาดทุกวันมักจะถูกนำเสนอโดยสถานการณ์ของการใช้งานที่เกือบล้มเหลว ข้อความแจ้งเตือนเหล่านี้สามารถช่วยให้วิศวกรสามารถค้นหาประกาศการรั่วไหลได้เร็วขึ้นและลดการเกิดขัดข้องของไลน์ลงได้ ดังนั้นจึงสามารถหลีกเลี่ยงสายการผลิตของบริษัทนี้ได้ หวังว่าจะเป็นประโยชน์แก่บริษัทอื่น ๆ ที่มีปัญหาเดียวกันเกี่ยวกับการรั่วไหล การรั่วไหลของอากาศบีบอัดสามารถทำให้เกิดขาดทุนในภาคการผลิตถึงครึ่งหนึ่ง [1]



บทที่ 5

บทสรุป

ในบทนี้จะกล่าวถึงข้อเสนอแนะพร้อมด้วยแนวทางในการปรับปรุงวิธีที่นำเสนอในวิทยานิพนธ์ รวมถึงบทสรุปของวิทยานิพนธ์เพื่อการประยุกต์ใช้และการพัฒนางานวิจัยในอนาคต

ข้อเสนอแนะ

5.1.1. เพิ่มเครื่องมือหรืออุปกรณ์เซนเซอร์ในการตรวจวัดหาความผิดปกติ

ผู้วิจัยขอเสนอให้ใช้อุปกรณ์ ไจโรสโคปเซนเซอร์ หรือ gyroscope sensor เพื่อเป็นตัวช่วยในการตรวจจับพฤติกรรมความผิดปกติของฟิตติ้ง เนื่องด้วยเป็นอุปกรณ์ที่ใช้แรงโน้มถ่วงของโลกในการตรวจวัดความผิดปกตินี้

5.1.2 ในการนำโมเดลที่ได้จากการฝึกฝนไปใช้กับฟิตติ้งชนิดอื่น ๆ

ผู้วิจัยแนะนำว่าควรทดลองเก็บข้อมูลและฝึกสอนโมเดลสำหรับฟิตติ้งประเภทนั้น ๆ ไป เช่น โมเดลที่ผู้วิจัยได้ดำเนินการนั้นใช้ได้กับฟิตติ้งข้อ 2 ทาง แต่ถ้าต้องการโมเดลสำหรับตรวจหารอยร้าวในฟิตติ้ง 3 ทาง ต้องเก็บข้อมูลและฝึกสอนโมเดลใหม่ สำหรับประเภทที่ต้องการเป็นประเภทไป

5.1.3 เมื่อมีการใช้งานระบบรวมถึงโมเดลไปเป็นช่วงเวลาหนึ่งควรให้มีการฝึกสอนโมเดลใหม่

กระบวนการฝึกสอนโมเดลใหม่นั้นสามารถทำได้ด้วยการเพิ่มกระบวนการ ฝึกสอนโมเดลใหม่ หรือ Retrain model เพื่อที่จะให้ใช้ข้อมูลที่ได้เก็บมาในการใช้งานจริงนั้น เป็นข้อมูลสำหรับประมวลและฝึกสอนโมเดลไปได้เลย

5.1.4 ควรมีการนำระบบที่ผู้วิจัยได้นำเสนอไปใช้งานกับสายการผลิตจริง

ทั้งนี้เพื่อเป็นการทดสอบเก็บข้อมูลการใช้งานจริง ๆ และรวมถึงต้องมีการคำนวณต้นทุนของวัสดุอุปกรณ์ต่าง ๆ เช่น เซนเซอร์ต่อจุดที่ต้องการวัด หรืออุปกรณ์ที่ต้องการนำมาใช้ประมวลผลจริง ๆ นั้น ต้องมีความเหมาะสมกับงานแบบใด เช่น เทคโนโลยีเอ็ดจ์คอมพิวเตอร์นั้น ควรใช้เป็น Jetson Nano

5.1.5 กระบวนการปรับปรุงให้ข้อมูลนั้นละเอียดขึ้น

เนื่องจากในวิทยานิพนธ์นี้ผู้วิจัยได้แบ่งข้อมูลออกเป็น 5 ส่วน ส่วนละ 100 ไทม์สเต็ป เห็นสมควรให้มีจำนวนข้อมูลมีน้อยลงไปกว่านี้ เพื่อความยืดหยุ่นของการปรับปรุงข้อมูลให้คล้ายกับความเป็นจริงมากขึ้น

5.1.6 เพิ่มกระบวนการลดความผิดพลาดของคอนฟิวชั่นเมตริกและแอปพลิเคชัน

ในจำนวนข้อมูลที่นำมาใช้พิจารณาผลการทำนายของโมเดลนั้นควรเพิ่มขนาดของข้อมูลให้มากกว่าที่ผู้วิจัยได้ใช้งาน โดยการใช้โมเดลที่ซับซ้อนมากขึ้น เพิ่มกระบวนการวิศวกรรมคุณลักษณะ และปรับปรุงพารามิเตอร์ที่มากกว่าที่ผู้วิจัยใช้ เป็นต้น

5.1.7 เพิ่มคุณลักษณะของข้อมูลที่มากกว่าข้อมูลทางสถิติ

ด้วยข้อมูลที่เกิดขึ้นในงานวิจัยนี้เป็นข้อมูลที่เกิดขึ้นบ่อย ๆ จึงควรเพิ่มคุณลักษณะที่เป็นสัญญาณความถี่ และนำไปใช้ในการฝึกสอนโมเดล หรือคุณลักษณะอื่น ๆ นอกจากนี้



บทสรุปวิทยานิพนธ์

โมเดล 1D-CNN ของผู้วิจัยเป็นแนวทางในการจัดการข้อกังวลของการจำแนกประเภทของคลาส 0 และคลาส

2

แบบจำลอง 1D-CNN ของผู้วิจัยสามารถเป็นแนวทางที่เหมาะสมสำหรับการแก้ไขปัญหาในการจำแนก Class 0 (การเชื่อมต่อ Fitting แน่นหนา, วาล์วเปิด) และ Class 2 (การเชื่อมต่อ Fitting แบบแน่นหนากลางพร้อมวาล์วเปิด) ผู้วิจัยได้พูดถึงข้างต้นว่าผู้วิจัยไม่สามารถแยกแยะสถานการณ์ของคลาสเหล่านี้ได้ นั่นหมายถึงผู้วิจัยจะมีเพียงสองสถานการณ์เท่านั้น คือปกติหรือผิดปกติ ดังนั้น แอปพลิเคชันจะสามารถเตือนวิศวกรเกี่ยวกับสถานการณ์ปกติและผิดปกติของท่อต่อได้ แต่ไม่ใช่แอปพลิเคชันที่มีประโยชน์ ที่ผู้วิจัยต้องการคือการเตือนวิศวกรเกี่ยวกับสถานการณ์ 'ปกติ' (Class 0) ต่อจากนั้น ผู้วิจัยจำเป็นต้องแยกแยะ Class 0 จาก Class 2 ในแอปพลิเคชันของผู้วิจัย ผู้วิจัยจะใช้โมเดล 1D-CNN ที่ผู้วิจัยได้สร้างขึ้นในขั้นตอนก่อนหน้านี้เพื่อแก้ไขปัญหา ในหัวข้อนี้จะอธิบายถึงว่าโมเดล 1D-CNN ของผู้วิจัยเป็นแนวทางที่เหมาะสมในการจำแนก Class 0 จาก Class 2 อย่างไร ในการทดลองนี้ โมเดลนี้สามารถจำแนก Class 0 จาก Class 2 ได้ด้วยความแม่นยำถึง 98.78% และด้วยความแม่นยำนี้ โมเดลถือว่าเป็นโมเดลที่เหมาะสม ดังนั้น สามารถนำไปปฏิบัติในการตรวจบรอยรั่วของอากาศจากท่อในระบบด้วยระบบเซ็นเซอร์เร็กเซลโลมิเตอร์ และแจ้งเตือนวิศวกรเกี่ยวกับบรอยรั่วโดยใช้โมเดลนี้ได้ ได้รับข้อมูลจากเมทริกซ์ความสับสนในส่วน 3.3 และรูปที่ 5B ว่าการทำนายของคลาส 0 ได้ทำนายผิด 3 ครั้ง ค่าทำนายผิด 3 ครั้งสามารถเห็นได้เป็น '3' เมื่อแกน x ที่ 0 และแกน y ที่ 2 จากรูปที่ 5B ค่าทำนายผิด 3 ครั้งเป็น Class 2 สำหรับค่าทั้งสามค่าแทนที่จะเป็น Class 0 ค่าทำนายผิดทั้งหมดสามค่านี้เป็นประมาณ 0.62% ของทุกค่าทำนายที่ถูกต้องและผิดของ Class 0 ที่นี้ 0.62% ถูกคำนวณจาก '3 ค่าทำนายผิด / (480 ค่าทำนายถูก + 3 ค่าทำนายผิด) * 100' โปรดทราบว่า จากรูปที่ 5B, '480 ค่าทำนายถูก' คือจำนวนจากแกน x ที่ 0 และแกน y ที่ 0

การทำนายสำหรับ Class 2 มีค่าเทียบเท่ากับเท็จ 9 ครั้ง ค่าทำนายเทียบเท่านี้สามารถเห็นได้เป็น '9' เมื่อแกน x ที่ 2 และแกน y ที่ 0 จากรูปที่ 6B ค่าทำนายเทียบเท่า 9 ค่านี้เป็น Class 0 ทั้ง 9 ค่าแทนที่จะเป็น Class 2 ค่าทำนายที่ผิดเพียงเหล่านี้เป็นประมาณ 1.83% ของทุกรายการค่าทำนายที่ถูกและผิดของ Class 2 ที่นี้ ดังนั้น, 1.83% ถูกคำนวณจาก '9 ค่าทำนายที่ผิด / (475 ค่าทำนายที่ถูก + 9 ค่าทำนายที่ผิด + 8 ค่าทำนายที่ผิด) * 100' Note that, from Figure 6 (B), '8 false' is the number from X-axis at 2 and Y-axis at 3. จากรูปที่ 3 ผู้วิจัยมีปัญหากับการจำแนกประเภท Class 0 และ Class 2 เมื่อดำเนินการจำแนกประเภทโดยไม่มี ML แม้ว่าจะใช้เทคนิค DT, RF, SVM และ 1D-CNN ML ก็ตาม อย่างไรก็ตาม, จากข้อมูลใน Figure 5B ในขั้นตอนนี้, 1D-CNN สามารถสร้างโมเดลที่มีความแม่นยำของค่าทำนายที่ถูกต้องของ Class 0 ที่ 99.38% (100% - 0.62%) และ 98.17% (100% - 1.83%) ของ Class 2, ตามลำดับ หลังจากที่คุณได้บอกถึงความแม่นยำของโมเดลที่ 98.78% ในการจำแนกคลาส 0 จากคลาส 2 ของคุณ น่าจะเป็นข้อมูลที่มีประสิทธิภาพมากที่สุดที่โมเดลสามารถทำนายคลาสที่ถูกต้องในสถานการณ์ที่ท่านกำลังใช้งาน นี่เป็นผลลัพธ์ที่ดีและแสดงให้เห็นว่าโมเดล 1D-CNN ของคุณเหมาะสมที่จะใช้ในการจำแนก Class 0 และ Class 2 ในข้อมูลของคุณอย่างมีประสิทธิภาพ ดังนั้นโมเดล 1D-CNN ของผู้วิจัยจึงเป็นแนวทางที่เหมาะสมสำหรับการจัดการข้อกังวลของการจำแนกประเภทของคลาส 0 และคลาส 2 และสำหรับการสร้างแอปพลิเคชันที่เป็นไปได้ซึ่งจะกล่าวถึงในหัวข้อถัดไป

Class 4 และ Class 5 จะแทนสถานการณ์ที่คลาส 4 และคลาส 5 แทบจะเหมือนกันคือสถานการณ์การติดตั้งที่หลวมพอดี และแสดงถึงการใช้งานที่เป็นไปได้คลาส 4 และคลาส 5 จะอยู่ในสถานการณ์ที่มีการติดตั้งที่หลวมเหมือนกัน ในส่วนนี้จะถูกพูดถึงว่าทำไมคลาส 4 และคลาส 5 ถึงถือว่าอยู่ในสถานการณ์การติดตั้งที่หลวมเหมือนกัน สถานการณ์ที่เหมือนกันนี้สามารถเปิดโอกาสให้มีการใช้แอปพลิเคชันที่เป็นไปได้ จากส่วน 4.1 ข้างต้น โมเดลสามารถจัดการกับปัญหาการจำแนกคลาส 0 และคลาส 2 และการสร้างแอปพลิเคชันที่เป็นไปได้ นอกจากนี้จากส่วนที่ 3.2 โมเดลสามารถจัดกลุ่มคลาส 0, คลาส 1, คลาส 2, คลาส 3, คลาส 4, และ คลาส 5 ได้ด้วยความแม่นยำถึง 99.2% ดังนั้น โมเดลสามารถให้คำแนะนำในการตรวจจบบรรยากาศจากการต่อท่อในระบบท่อลมโดยใช้ระบบเซ็นเซอร์แรงเร่ง, จากสามสถานการณ์คือ 'ปกติ' (คลาส 0 และคลาส 1), 'ใกล้ล้มเหลว' (คลาส 2 และคลาส 3), และ 'ล้มเหลว' (คลาส 4 และคลาส 5) ในส่วนที่ 2.5.1 ผู้วิจัยตั้งชื่อข้อมูลการรั่วท่อเป็น 6 ชั้น คลาส 4 คือ “การเชื่อมต่อที่หลวมพอดี - วาล์วปิด” และคลาส 5 คือ “การเชื่อมต่อที่หลวมพอดี - วาล์วเปิด” ดังนั้น คลาส 4 และคลาส 5 ทั้ง / คลาส คือสถานการณ์การเชื่อมต่อที่หลวมพอดี โมเดล 1D-CNN ของผู้วิจัยสามารถแยกแยะความแตกต่างของสถานการณ์ที่หลวม (คลาส 4 และคลาส 5) จากสถานการณ์ที่มีความหนาแน่นปานกลาง (คลาส 2 และคลาส 3) ซึ่งจะกล่าวถึงในหัวข้อ 3.2 ในขณะที่คลาส 4 (การเชื่อมต่อ Fitting แบบหลวมเล็กน้อย, วาล์วปิด) 1D-CNN model ไม่ได้ทำนายผิดเลยจากคอนฟิวชันเมตริก (confusion matrix) หรือรูปที่ 6B นี้ นั่นถือเป็นผลลัพธ์ที่ดี เนื่องจากแสดงให้เห็นว่าโมเดลสามารถระบุคลาสนี้ได้ถูกต้องทั้งหมด ซึ่งเป็นข้อมูลที่มีความคุ้มค่าโดยเฉพาะในบริบทของการตรวจจบบรรยากาศการเชื่อมต่อที่หลวมเล็กน้อย เนื่องจากการทำนายผิดอาจทำให้เกิดการประเมินสถานะของระบบที่ไม่ถูกต้องได้ แต่โมเดลมีการทำนายผิด 4 ครั้งสำหรับคลาส 5 โดยค่าทำนายผิดทั้งหมดนี้เป็น 'Class 4' แทนที่จะเป็น 'Class 5' อย่างไรก็ตาม ทั้งคลาส 4 และคลาส 5 จะเกี่ยวกับสถานการณ์การเชื่อมต่อที่หลวมอยู่ ดังนั้น โมเดลยังรู้ค่าทำนายที่ผิดเหล่านี้ทั้ง 4 ข้อมูล/สัญญาณ ยังคงอยู่ในสถานการณ์การเชื่อมต่อที่หลวม ไม่ใช่สถานการณ์การเชื่อมต่อที่แน่นหนาหรือแน่นตึง

ผู้วิจัยได้พัฒนาโมเดลการเรียนรู้เชิงลึกแบบ Convolutional Neural Network (CNN) สำหรับการตรวจจบบรรยากาศจากการเชื่อมต่อในระบบท่อลมป้อนโดยใช้ระบบเซ็นเซอร์แรงเร่งได้ โมเดลนี้เป็นการเสนอสร้างที่สำคัญสำหรับวงการ เนื่องจากมันใช้เทคนิคการเรียนรู้เชิงลึกโดยเฉพาะโครงสร้าง 1D-CNN เพื่อวิเคราะห์ข้อมูลเซ็นเซอร์แรงเร่งและจำแนกสภาพต่าง ๆ ของการเชื่อมต่อ

โมเดล 1D-CNN ถูกออกแบบเพื่อประมวลผลลำดับของข้อมูลแรงเร่งเพื่อระบุรูปแบบและลักษณะที่บ่งชี้ถึงคลาสที่แตกต่างกัน เช่น การเชื่อมต่อแน่น, การเชื่อมต่อระหว่างการเชื่อมต่อแน่นกลาง, และการเชื่อมต่อแบบหลวมความสามารถของโมเดลในการแยกแยะระหว่างคลาสเหล่านี้มีความหมายสำคัญในการประเมินสภาพของระบบลมป้อน

กระบวนการฝึกโมเดลมีการกำหนดค่าเริ่มต้นของข้อมูล, การสร้างภายใน, และการพิจารณาประสิทธิภาพของโมเดลเช่นกัน ที่สำคัญมีการพัฒนาโมเดล CNN ในการประมวลผลข้อมูลที่ถูกเตรียมการล่วงหน้า, การสร้างภายใน และการประเมินประสิทธิภาพในทางความถูกต้อง โดยรูปแบบ 1D-CNN ได้แสดงความแม่นยำสูงโดยเฉพาะในการจำแนก Class 0 จาก Class 2 และในการแยกแยะระหว่างการเชื่อมต่อแบบหลวม (Class 4 และ Class 5) การพบความสัมพันธ์ครั้งแรกคือความแม่นยำที่ไม่ยอมรั้งที่พบในโมเดลทั้งหมดที่ได้รับการฝึกซ้อมก่อนการขยายข้อมูลที่ถูก

เก็บมานั้น เป็นผลมาจากพฤติกรรมที่ไม่ธรรมชาติของข้อมูลที่ผู้วิจัยเก็บมาเมื่อเทียบกับข้อมูลจริงจากเส้นการผลิต ผู้วิจัยได้ทำการขยายข้อมูลที่เก็บมาโดยการจำลองข้อมูลจริง กระบวนการขยายข้อมูลยังเพิ่มประชากรของข้อมูลที่ถูกขยายเพิ่มขึ้น ซึ่งทำให้กระบวนการฝึกซ้อมและทดสอบที่ดีขึ้นเมื่อใช้ข้อมูลที่ถูกขยาย เพื่อสร้างโมเดลที่ดีขึ้น การทดสอบที่ส่งจากการทดสอบของผู้วิจัยคือว่าโมเดลของผู้วิจัยมีความแม่นยำ 99.2% หลังจากกระบวนการขยายข้อมูล เทียบกับความแม่นยำ 86.30% ด้วยข้อมูลก่อนการขยายข้อมูลของผู้วิจัย การค้นพบที่สามคือการประเมินความแม่นยำและการอธิบายของขั้นตอนวิธีที่เรียบง่ายช่วยให้สามารถสร้างแอปพลิเคชันขึ้นไปได้โดยขึ้นอยู่กับโมเดลและผลการประเมิน ดังนั้นผู้วิจัยสามารถแนะนำได้ว่าแอปพลิเคชันนี้สามารถแจ้งเตือนผู้ประกอบการในเส้นการผลิตให้ทราบเกี่ยวกับสถานการณ์ 'ใกล้ต่อจากความล้มเหลว' หรือ 'ความล้มเหลว' ของอุปกรณ์ในระบบท่อได้อย่างเชื่อถือได้ เทคนิคการตรวจจ็บรอยรั่วสามารถเอาชนะบางปัญหาที่เกี่ยวข้องกับการรั่วของอากาศในระบบลมแบบนิวแมติก ซึ่งรวมถึงการหยุดเครื่องจักรในระบบการผลิต ความไม่แม่นยำของกระบวนการ และเสียงรบกวน [25] การวิจัยของผู้วิจัยในอนาคตจะพัฒนาแอปพลิเคชันเพิ่มเติมและเปิดโอกาสให้นำไปใช้จริงในเส้นการผลิตเพื่อรับมือกับปัญหาที่เกี่ยวข้อง [25] ผู้วิจัยยังจะดำเนินการพัฒนาโมเดลเพื่อให้สามารถระบุตำแหน่งของรอยรั่ว เพื่อแก้ไขปัญหที่สำคัญที่ทำให้เกิดความล่าช้าในการค้นหาและแก้ไขข้อผิดพลาดเมื่อต้องการผู้ดูแลเส้นผลิตต้องตรวจสอบสายพันธุ์ตามเส้นจนกว่าจะพบว่าสายพันธุ์ที่เสียหาย ซึ่งอาจทำให้เกิดความล่าช้าในการค้นหาและการแก้ไขข้อผิดพลาดได้ significant

สุดท้ายนี้วิทยานิพนธ์ฉบับนี้ได้รับทุนสนับสนุนจาก สถาบันบัณฑิตวิทยาศาสตร์และเทคโนโลยีไทย (TGIST) สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ สัญญารับทุนเลขที่ SCA-CO-2563-12152-TH



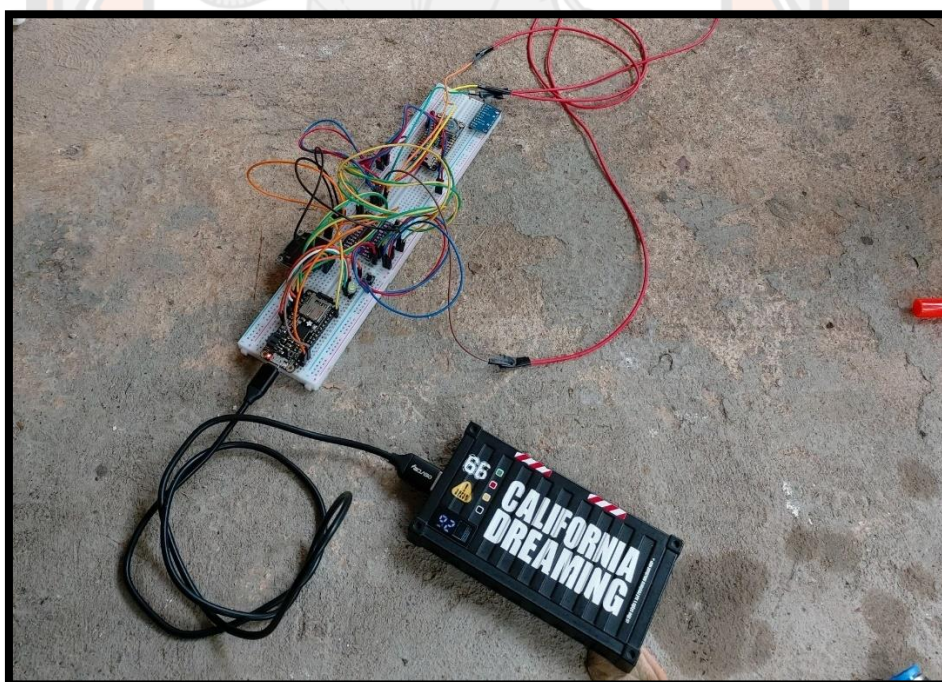




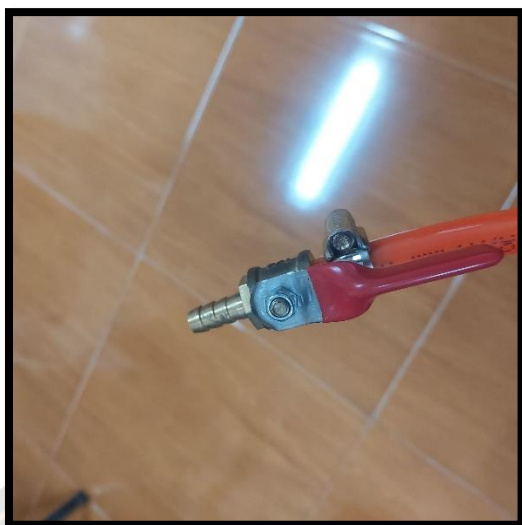
ภาคผนวก



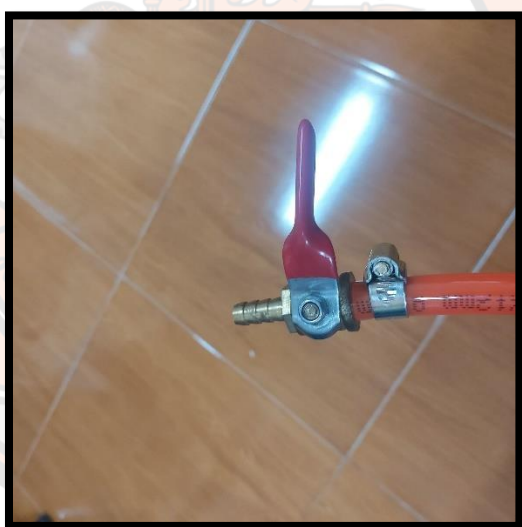
รูป 17 การติดตั้งอุปกรณ์ทหารทดลองเก็บข้อมูล



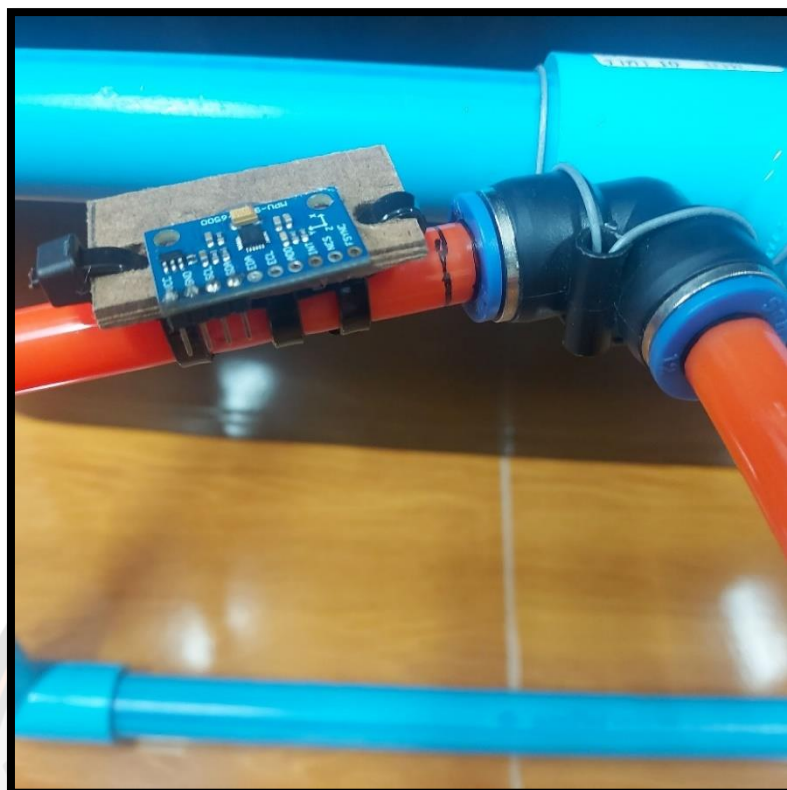
รูป 18 การติดตั้งอุปกรณ์รับข้อมูลจากเซนเซอร์และเขียนลงเมมโมรี่การ์ด



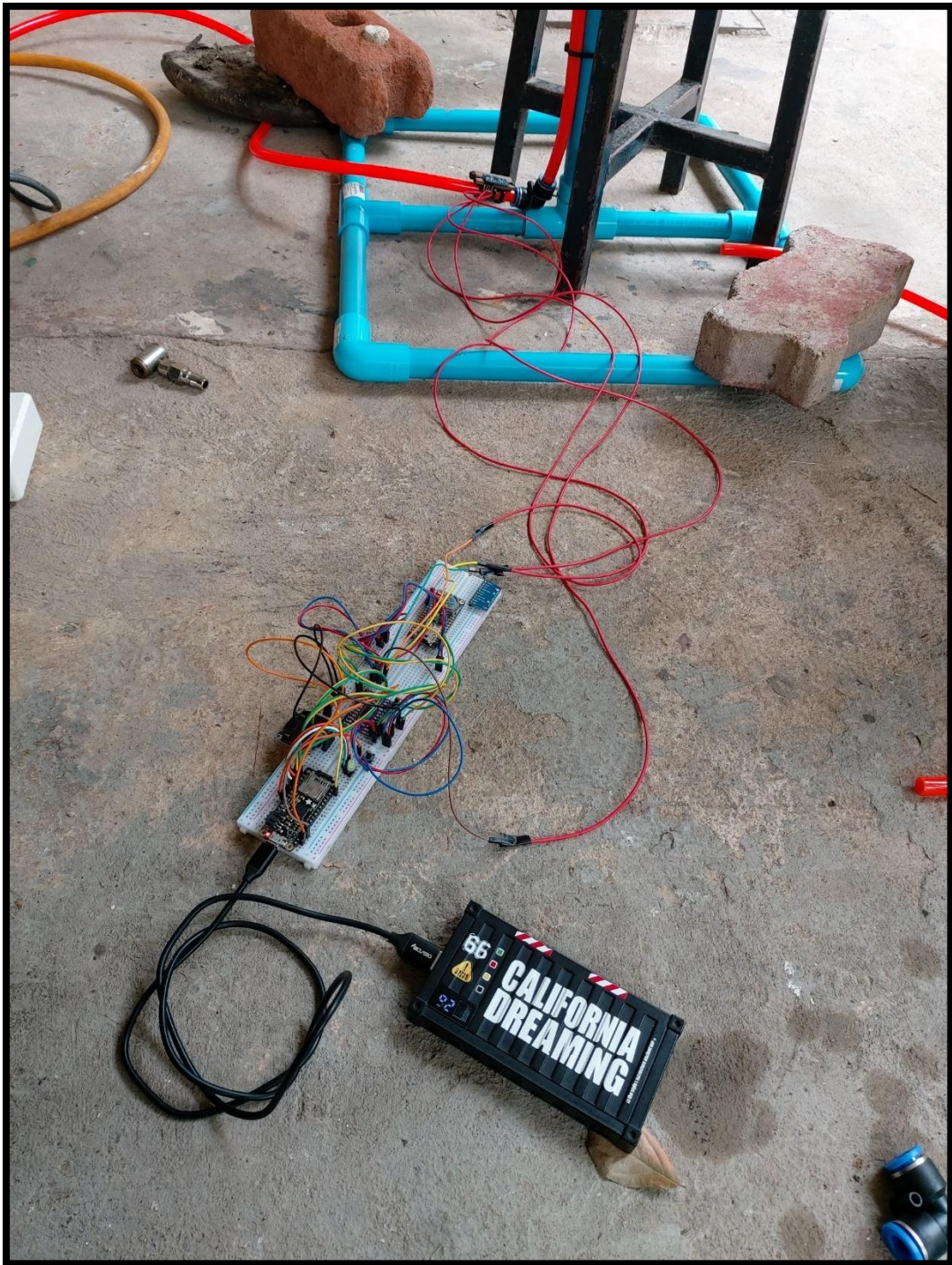
รูป 19 วาล์วเปิด



รูป 20 วาล์วปิด



รูป 21 การติดตั้งเซนเซอร์กับฟิตติ้ง



รูป 22 การติดตั้งภาพรวมของการทดลองทั้งหมด



อภิธานศัพท์

ลำดับ	คำศัพท์	คำที่ใช้
1	1D-CNN	1D-CNN
2	Accelerometer sensor	เซนเซอร์วัดความเร่งสามแกน
3	accuracy score	ค่าความถูกต้อง
4	augment data	ข้อมูลที่ถูกรับแต่ง
5	Class	คลาส
6	Classification	การจำแนกแบ่งกลุ่ม
7	Closed	วาล์วปิด
8	Compressed air leak	การรั่วซึมของการอัดอากาศ
9	Compressor	เครื่องอัดอากาศ
10	Confusion Matrix	คอนฟิวชันเมตริก
11	Convolutional layer	ชั้นการคอนโวลูชัน
12	Data	ข้อมูล
13	Data cleansing	การทำความสะอาดชุดข้อมูล
14	Data exploration	การสำรวจข้อมูล
15	data type	ลักษณะข้อมูล
16	Dataset	ชุดข้อมูล
17	Decision Tree, DT	เทคนิคต้นไม้ตัดสินใจ
18	Deep Learning, DL	เทคนิคการเรียนรู้เชิงลึก
19	fail situation	สถานการณ์ล้มเหลว
20	Feature importance	ความสำคัญของคุณลักษณะ
21	Fitting	ฟิตติ้ง
22	float	ตัวเลขทศนิยม
23	forward fill	ค่าที่อยู่ก่อนหน้าที่ไม่ขาดหายไป
24	Global average pooling layer	ชั้นกระจายค่าเฉลี่ย
25	High Density Polyethylene หรือ HDPE	ท่อพอลิเอทิลีนความหนาแน่นต่ำ

ลำดับ	คำศัพท์	คำที่ใช้
26	Hyperparameter	ไฮเปอร์พารามิเตอร์
27	Input features	คุณลักษณะสำหรับนำเข้า
28	Line	ไลน์
29	Low density polyethylene หรือ LDPE	ท่อพอลิเอทิลีนความหนาแน่นสูงท่อ
30	Machine Learning, ML	การเรียนรู้ของเครื่อง
31	Model	โมเดล
32	Model fitting	การหาตัวแบบที่เหมาะสม
33	near-failure situation	สถานการณ์ใกล้ล้มเหลว
34	Neural Network, NN	โครงข่ายประสาทเทียม
35	normal situation	สถานการณ์ปกติ
36	Not A Number (NaN)	ค่าว่าง
37	Opened	วาล์วเปิด
38	outlier	ค่าผิดปกติ
39	Pipe	ท่อ
40	Pneumatic line leakage detection	การตรวจจับการรั่วซึมในระบบท่อลม
41	Pneumatic pipe system	ระบบท่อลม
42	Point	จุด
43	Python	ภาษาไพธอน
44	Random Forest, RF	เทคนิคป่าไม้สุ่ม
45	Range of value	ช่วงของข้อมูล
46	raw data	ข้อมูลดิบ
47	Receiver	ถังรับอากาศ
48	Regression	การถดถอย
49	Scikit-learn	ไซคิทเลอร์น
50	Steel pipe	ท่อเหล็ก
51	String	ข้อความ
52	sub-sequences	ชุดข้อมูลย่อยที่ถูกแบ่งออกมาจากชุดข้อมูลหลัก
53	Support Vector Machine, SVM	เครื่องมือเวกเตอร์สนับสนุน

ลำดับ	คำศัพท์	คำที่ใช้
54	Test set	ข้อมูลทดสอบ
55	Train model	ฝึกสอนโมเดล
56	Training set	ข้อมูลฝึกสอน
57	Ultrasonic	อัลตราโซนิก
58	x axis	ค่าตามแนวแกน x
59	y axis	ค่าตามแนวแกน y
60	z axis	ค่าตามแนวแกน z
61	zero mean	การทำข้อมูลให้อยู่ศูนย์กลาง
62	คลาส 0	พิตตังตอกันแน่น วาล์วเปิด
63	คลาส 1	พิตตังตอกันแน่น วาล์วเปิด
64	คลาส 2	พิตตังตอกันเกือบแน่น วาล์วเปิด
65	คลาส 3	พิตตังตอกันเกือบแน่น วาล์วเปิด
66	คลาส 4	พิตตังตอกันหลวม วาล์วเปิด
67	คลาส 5	พิตตังตอกันหลวม วาล์วเปิด