

การตรวจจับและระบุรูปแบบความผิดปกติของเครื่องอัดอากาศแบบหมุนเหวี่ยงโดยใช้วิธีการเรียนรู้  
แบบหมู่



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชาระบบกายภาพที่เชื่อมประสานด้วยเครือข่ายไซเบอร์ ภาควิชาวิศวกรรมเครื่องกล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2564

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Fault detection and identification for Centrifugal Compressor by Ensemble model



A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Engineering in Cyber-Physical System

Department of Mechanical Engineering

FACULTY OF ENGINEERING

Chulalongkorn University

Academic Year 2021

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การตรวจจับและระบุรูปแบบความผิดปกติของเครื่องอัดอากาศแบบ
	หมุนเหวี่ยงโดยใช้วิธีการเรียนรู้แบบหมู่
โดย	นายธิตพัทธ์ สนธิโพธิ์
สาขาวิชา	ระบบกายภาพที่เชื่อมประสานด้วยเครือข่ายไซเบอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.รัชทิน จันทร์เจริญ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษา  
ตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

.....	คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)	
คณะกรรมการสอบวิทยานิพนธ์	
.....	ประธานกรรมการ
(ศาสตราจารย์ ดร.ไพโรจน์ สิงห์ถนัดกิจ)	
.....	อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.รัชทิน จันทร์เจริญ)	
.....	กรรมการ
(รองศาสตราจารย์ ดร.กฤษฎา พนมแข็ง)	
.....	กรรมการ
(อาจารย์ ดร.รีนา เซย์)	
.....	กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.พิเชษฐ พิณีจ)	

ชิติพัทธ์ สนธิโพธิ์ : การตรวจจับและระบุรูปแบบความผิดปกติของเครื่องอัดอากาศแบบหมุนเหวี่ยง โดยใช้วิธีการเรียนรู้แบบหมู่. ( Fault detection and identification for Centrifugal Compressor by Ensemble model) อ.ที่ปรึกษาหลัก : รศ. ดร.รัชทิน จันทร์เจริญ

ในวิทยานิพนธ์นี้ได้เสนอระบบควบคุมสุขภาพเครื่องจักรที่สามารถตรวจจับการทำงานที่ผิดปกติของเครื่องอัดอากาศแบบหมุนเหวี่ยงที่ติดตั้งในอุตสาหกรรมจริง ซึ่งเป็นเครื่องจักรที่มีความสำคัญต่อระบบโดยเมื่อมีการใช้งานเป็นระยะเวลานานทำให้ประสิทธิภาพลดลงและพบความผิดปกติที่เพิ่มขึ้น เครื่องจักรดังกล่าวจึงถูกติดตั้งเซนเซอร์เพื่อติดตามพฤติกรรมและบันทึกข้อมูล ซึ่งข้อมูลที่ได้รับมาจะยังไม่พบความเสียหายที่เกิดขึ้นอย่างชัดเจน วิธีการที่นำเสนอจะใช้การจำลองความผิดปกติให้กับข้อมูลที่ได้รับ จากนั้นจึงวิเคราะห์ประสิทธิภาพของเครื่องอัดอากาศและเปรียบเทียบกับการเรียนรู้ของเครื่องและเพอร์เซ็ปตรอนหลายชั้น ในการศึกษาพบว่าข้อมูลที่จะสะสมนั้นเป็นข้อมูลจริงที่มีความท้าทาย ซึ่งเทคนิคที่เสนอมามีประสิทธิภาพเพียงพอที่ใช้ในการตรวจจับความผิดปกติของเครื่องอัดอากาศแบบหมุนเหวี่ยงที่ติดตั้งในโรงงานจริง



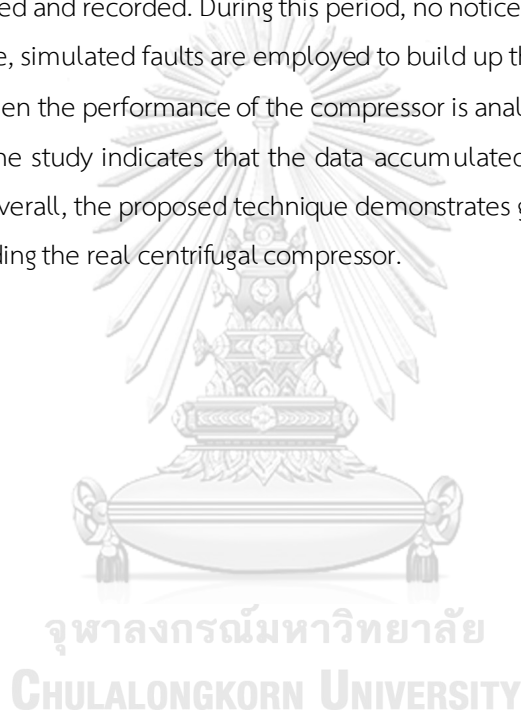
สาขาวิชา ระบบกายภาพที่เชื่อมประสานด้วย ลายมือชื่อนิสิต .....  
 เครือข่ายไซเบอร์  
 ปีการศึกษา 2564 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

# # 6370421521 : MAJOR CYBER-PHYSICAL SYSTEM

KEYWORD: fault detection, fault identification, ensemble model, centrifugal compressor

Titipat Sonthipo : Fault detection and identification for Centrifugal Compressor by Ensemble model. Advisor: Assoc. Prof. RATCHATIN CHANCHAROEN, Ph.D.

In this work, a machine's health monitoring system (MHMS) that can detect anomalies for a centrifugal compressor, operating in a real plant, is proposed. A compressor is a crucial machine but over time its performance can decline, and faults can develop. In the compressor under investigation, several sensors have been installed. Subsequently, its behavior is monitored and recorded. During this period, no noticeable fault is detected. In the proposed technique, simulated faults are employed to build up the data that can be used in the investigation, then the performance of the compressor is analyzed and compared with ML and MLP. Herein, the study indicates that the data accumulated is a good candidate for this challenging case. Overall, the proposed technique demonstrates good potential for detection of anomalies regarding the real centrifugal compressor.



Field of Study: Cyber-Physical System

Student's Signature .....

Academic Year: 2021

Advisor's Signature .....

## กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ รองศาสตราจารย์ ดร. รัชทิน จันทร์เจริญ อาจารย์ที่ปรึกษา และ นาย ทัดพล อัจ  
สำอังก์ และภาควิชาเครื่องกลที่ได้ให้คำแนะนำ แนวคิด สถานที่ปฏิบัติงาน ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ มาโดย  
ตลอดจนโครงการเล่มนี้สำเร็จสมบูรณ์ คณะผู้จัดทำจึงกราบขอบพระคุณทุกท่านเป็นอย่างสูง

นอกจากนี้ขอกราบขอบพระคุณ คุณพ่อและคุณแม่ ผู้ให้กำลังใจและให้โอกาสให้การศึกษาอันมีค่ายิ่ง ขอให้  
คุณความดีส่งผลถึงผู้มีพระคุณทุกท่าน หากโครงการนี้มีข้อบกพร่องประการใด ผู้จัดทำขอน้อมรับด้วยความเคารพยิ่ง  
เพื่อนำไปปรับปรุงแก้ไขในโอกาสต่อไป และหวังเป็นอย่างยิ่งว่าโครงการนี้จะอำนวยความสะดวกและเป็นประโยชน์ต่อผู้  
ที่สนใจได้เป็นอย่างดี

ธิตัพพัธน์ สนธิโพธิ์



## สารบัญ

	หน้า
.....	ก
บทคัดย่อภาษาไทย.....	ก
.....	ง
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญภาพ.....	1
สารบัญตาราง.....	3
บทที่ 1.....	4
บทนำ.....	4
1.1 ที่มาและความสำคัญของวิทยานิพนธ์.....	4
1.2 วัตถุประสงค์ของวิทยานิพนธ์.....	7
1.3 ขอบเขตของวิทยานิพนธ์.....	7
1.4 ขั้นตอนการดำเนินงานโดยสรุป.....	8
บทที่ 2.....	9
ปริทัศน์วรรณกรรม.....	9
2.1 ความผิดปกติในเครื่องอัดอากาศแบบหมุนเหวี่ยง.....	9
2.2 การตรวจจับและระบุรูปแบบความผิดปกติ.....	12
2.3 การตรวจจับความผิดปกติโดยใช้การเรียนรู้แบบหมู่.....	17
บทที่ 3.....	20
ทฤษฎีพื้นฐาน.....	20

3.1 เครื่องอัดอากาศแบบหมุนเหวี่ยง (Centrifugal Compressor).....	20
3.2 ระบบการวินิจฉัยความผิดปกติ (fault diagnosis system).....	22
3.2.1 การตรวจจับความผิดปกติ (Fault detection).....	23
การตรวจจับตามแบบจำลอง (Model based).....	24
การตรวจจับโดยใช้ฐานความรู้ (Knowledge based).....	25
การตรวจจับโดยใช้สัญญาณ (Signal based).....	25
3.2.2 การวินิจฉัยความผิดปกติ (Fault diagnosis).....	26
การวินิจฉัยความผิดปกติด้วยวิธีการแบ่งหมวดหมู่.....	27
วินิจฉัยความผิดปกติด้วยวิธีการอนุมาน.....	27
3.3 การเรียนรู้ของเครื่องจักร (Machine Learning ).....	28
3.4 การเตรียมข้อมูล (Data preprocess).....	29
3.5 การเรียนรู้แบบหมู่.....	32
3.6 การประเมินแบบจำลอง.....	37
บทที่ 4.....	40
ผลการทดลอง.....	40
4.1 เครื่องอัดอากาศแบบหมุนเหวี่ยง.....	40
4.2 ข้อมูลที่รวบรวม.....	41
4.4 การจำลองความผิดปกติ.....	45
การจำลองความผิดปกติกับชุดข้อมูล MaFaultDa.....	45
การจำลองความผิดปกติกับชุดข้อมูลเครื่องอัดอากาศแบบหมุนเหวี่ยง.....	48
4.5 แบบจำลอง.....	50
4.6 ผลการทดลอง.....	52
บทที่ 5.....	54



สรุปและวิจารณ์ผลการทดสอบ.....	54
5.1 สรุปและวิจารณ์ผลการทดสอบ.....	54
5.2 ปัญหาที่พบในการจัดทำวิทยานิพนธ์ .....	54
5.3 งานวิจัยในอนาคต.....	55
บรรณานุกรม.....	56
ประวัติผู้เขียน .....	60



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

## สารบัญภาพ

หน้า

รูป 1 ความสัมพันธ์ระหว่างอัตราการไหลกับความดันที่ทางออกของเครื่องอัดอากาศแต่ละชนิด.....	6
รูป 2 การเกิดความดันย้อนกลับ .....	10
รูป 3 เส้น Surge line .....	12
รูป 4สเปกตรัมระหว่างอ้างอิงกับข้อมูลที่ต้องการจำแนก .....	13
รูป 5 การแบ่งข้อมูลโดยใช้ระยะทางแบบยุคลิด .....	13
รูป 6 การจดจำรูปแบบโดยใช้สถิติ.....	14
รูป 7 ตำแหน่งความผิดปกติที่เกิด .....	15
รูป 8 เครื่องอัดอากาศแบบลูกสูบ.....	21
รูป 9 เครื่องอัดอากาศแบบหมุนเหวี่ยง.....	22
รูป 10 รูปแบบของความผิดปกติแต่ละชนิด.....	24
รูป 11 ภาพแบบจำลองทางคณิตศาสตร์.....	24
รูป 12 การตรวจสอบโดยใช้ค่าจริงกับใช้แนวโน้มโดยค่า $Y_{max}$ และ $Y_{min}$ จะอยู่ในรูปของ $Y'$ .....	26
รูป 13 วิธีการที่ใช้การวินิจฉัยความผิดปกติ.....	27
รูป 14 แผนภาพ Fault tree .....	28
รูป 15 การสร้างข้อมูลสังเคราะห์ .....	31
รูป 16 ต้นไม้ตัดสินใจ.....	33
รูป 17 ค่าฟังก์ชัน Entropy และ Gini สำหรับข้อมูล 2 รูปแบบ.....	34
รูป 18 การทำงานของ Bagging .....	35
รูป 19 การทำงานของ Boosting.....	36
รูป 20 ตารางเมทริกซ์ความสับสน.....	38
รูป 21 กระบวนการ Polymer Flashing .....	40

รูป 22 เครื่องอัดอากาศแบบหมุนเหวี่ยงยี่ห้อ Ebara model: 200TC.....	41
รูป 23 ตำแหน่งเซนเซอร์ที่ติดตั้งบนเครื่องอัดอากาศ.....	42
รูป 24 อัตราการไหลในรูปเวลาและความถี่.....	43
รูป 25 การกระจายตัวของข้อมูลเครื่องอัดอากาศ.....	44
รูป 26 เปรียบเทียบความสัมพันธ์ระหว่างข้อมูลปกติกับข้อมูลเบี่ยงศูนย์ในแนวตั้ง .....	47
รูป 27 เปรียบเทียบลักษณะของข้อมูลระหว่างความเสียหายจริง ความเสียหายจำลอง และค่าปกติ .....	47
รูป 28 การแบ่งช่วงของข้อมูลและค่าที่เพิ่ม/ลด.....	49



## สารบัญตาราง

	หน้า
ตาราง 1 ข้อมูลที่วัดได้จากเซนเซอร์.....	46
ตาราง 2 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความเสียหายของเครื่องอัดอากาศ โดยใช้แบบจำลองป่าต้นไม้แบบสุ่ม.....	50
ตาราง 3 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความเสียหายของเครื่องอัดอากาศ โดยใช้แบบจำลองต้นไม้ตัดสินใจเอ็กซ์ทรีมเกรเดียนต์บูตดิ้ง .....	51
ตาราง 4 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความเสียหายของเครื่องอัดอากาศ โดยใช้แบบจำลองซัพพอร์ตเวกเตอร์แมชชีน .....	51
ตาราง 5 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความเสียหายของเครื่องอัดอากาศ โดยใช้แบบจำลอง MLP .....	51
ตาราง 6 ประสิทธิภาพแต่ละแบบจำลองที่ผ่านกระบวนการ Normalization.....	52
ตาราง 7 ประสิทธิภาพแต่ละแบบจำลองที่ไม่ผ่านกระบวนการ normalize .....	53
ตาราง 8 ค่า F1-score แต่ละรูปแบบ .....	53

## บทที่ 1

### บทนำ

เนื้อหาในบทนี้จะนำเสนอที่มาและความสำคัญของปัญหา วัตถุประสงค์ ขอบเขตของวิทยานิพนธ์ ขั้นตอนและวิธีการดำเนินงาน งานวิจัยในอดีตที่เกี่ยวข้องกับวิทยานิพนธ์และเนื้อหาโดยสังเขปของวิทยานิพนธ์

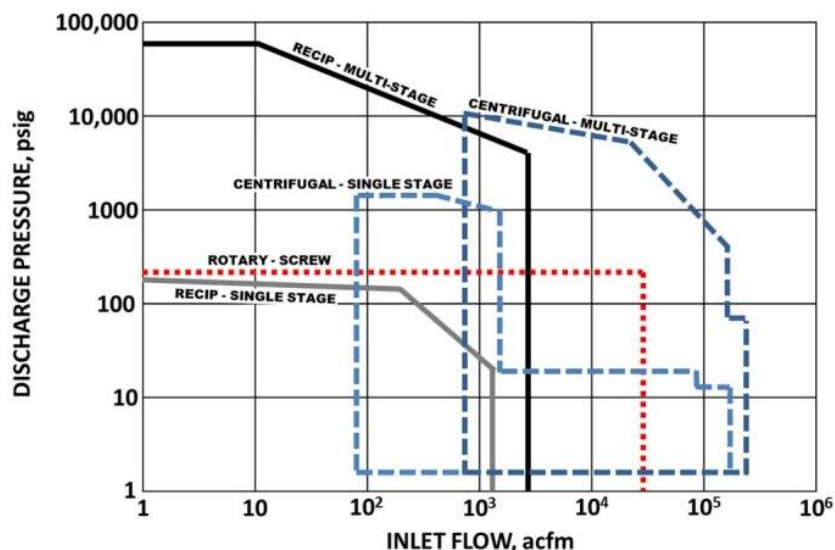
#### 1.1 ที่มาและความสำคัญของวิทยานิพนธ์

ในปัจจุบันอุตสาหกรรมเปลี่ยนแปลงจากยุคอุตสาหกรรม 3.0 เป็นยุคอุตสาหกรรม 4.0 เป็นการเชื่อมโลกของการผลิตเข้ากับเครือข่าย ซึ่งจะนำไปสู่นวัตกรรมใหม่อีกมากมาย โดยจะเน้นในการสื่อสารของเครื่องจักรกับเครื่องจักร (machine-to-machine หรือ M2M) และเทคโนโลยีอินเทอร์เน็ตของสรรพสิ่ง (Internet of things หรือ IoT) [1] เพื่อมาช่วยเพิ่มประสิทธิภาพการผลิตของระบบอัตโนมัติ ทำให้เกิดการตัดสินใจที่รวดเร็ว ถูกต้อง และทันเวลา ช่วยให้มีมนุษย์และเครื่องจักรทำงานร่วมกันได้ง่ายขึ้น เกิดโรงงานอัจฉริยะ (Smart Factory) แต่ในกระบวนการผลิตแบบอัตโนมัติจะต้องคำนึงถึงความต่อเนื่องของระบบเป็นสำคัญ เนื่องจากจะส่งผลกระทบต่อเป็นวงกว้างดังนั้นความเสียหายจึงเป็นสิ่งที่ไม่ยอมให้เกิดในกระบวนการผลิต จึงมีการทำนายและวินิจฉัยข้อบกพร่องรวมไปถึงการตรวจจับสถานะผิดปกติแบบทันทีทันใด เพื่อลดความสูญเสียในการผลิต เพิ่มประสิทธิภาพของระบบและเพิ่มความปลอดภัยของมนุษย์ เมื่อเครื่องจักรมีปัญหาหรือต้องการการซ่อมบำรุง การตรวจจับความผิดปกติจะช่วยประเมินความเป็นไปได้ที่จะเกิดเหตุการณ์ความเสียหายโดยใช้การบำรุงรักษาแบบคาดการณ์ล่วงหน้า (Predictive maintenance) การบำรุงรักษาแบบคาดการณ์ล่วงหน้าถูกนำมาใช้อย่างแพร่หลายเนื่องจากสามารถทำนายการเกิดความผิดปกติก่อนที่จะเกิดขึ้น [2] เพื่อนำมาวางแผนการหยุดการทำงานของเครื่องจักรในการซ่อมบำรุงก่อนที่จะแสดงความเสียหาย ทำให้ช่วยลดความเสียหายต่อสายการผลิตโดยใช้เพียงข้อมูลที่เก็บจากเครื่องจักรอย่างถูกต้องและเพียงพอ แต่เมื่อนำมาวิเคราะห์โดยผู้เชี่ยวชาญจะทำให้มีความแม่นยำที่สูงมากขึ้น

เครื่องอัดอากาศ (Compressor) เป็นอุปกรณ์ที่เพิ่มความดันของอากาศจากความดันบรรยากาศปกติให้สูงขึ้นตามที่ต้องการ โดยอาศัยหลักการลดปริมาตรของอากาศลงส่งผลให้ความดันเพิ่มขึ้น ซึ่งกระบวนการผลิตหลายกระบวนการจำเป็นต้องมีการเพิ่มความดันให้กับระบบ [2] จึงทำให้เครื่องอัดอากาศเป็นอุปกรณ์หนึ่งที่สามารถเห็นได้ในอุตสาหกรรมทั้งอุตสาหกรรมขนาดเล็กและอุตสาหกรรมขนาดใหญ่ มีประวัติการใช้งานอย่างยาวนานโดยในสมัยก่อนจะเป็นเครื่องสูบลม (Bellows) โดยเป็น

การใช้ทำเหยียบเพื่อป้อนอากาศเป็นเชื้อเพลิง แต่เนื่องจากผู้ใช้ต้องออกแรงตลอดเวลาที่ใช้งานทำให้เกิดความไม่ต่อเนื่องของอากาศที่นำเข้าไป จึงมีการพัฒนาตลอดจนมาถึงยุคปัจจุบันที่เครื่องอัดอากาศมีการใช้พลังงานไฟฟ้าทำให้สามารถทำงานได้อย่างอัตโนมัติ แม้จะผ่านเวลามายาวนานแต่เครื่องอัดอากาศก็ถูกนำมาใช้จนถึงปัจจุบัน ในยุคอุตสาหกรรม 4.0 จึงได้นำเทคโนโลยีมาประยุกต์ใช้กับเครื่องอัดอากาศ ทำให้สามารถสื่อสารกันได้แบบทันทีซึ่งเป็นหลักสำคัญของอุตสาหกรรม 4.0 โดยมีการรวบรวมข้อมูลของเครื่องจักรตลอดเวลาเป็นข้อมูลมหัต (Big data) โดยข้อมูลเครื่องจักรที่นำมาใช้ในงานวิจัยนี้คือ เครื่องอัดอากาศแบบหมุนเหวี่ยง (Centrifugal Compressor) เป็นเครื่องอัดอากาศที่ทำงานด้วยการเปลี่ยนพลังงานจลน์เป็นความดันของอากาศ โดยลมจะถูกดูดเข้าไปยังแกนกลางเพลลาใบพัดและถูกเหวี่ยงตัวไปตามแนวรัศมีของใบพัดสู่ผนังของเครื่องอัดและถูกส่งต่อตามท่อ อากาศอัดจะมีความดันที่เพิ่มขึ้นแต่ความเร็วยังคงที่ส่งผลให้มีอุณหภูมิที่สูงขึ้น เครื่องอัดอากาศแบบหมุนเหวี่ยงเป็นที่นิยมอย่างมากในอุตสาหกรรมเคมี เมื่อเทียบกับเครื่องอัดอากาศชนิดอื่น เครื่องอัดอากาศแบบหมุนเหวี่ยงสามารถใช้ได้กับแก๊สที่มีน้ำหนักโมเลกุลได้มากกว่าเครื่องอัดอากาศชนิดอื่น ได้อัตราการไหลที่มากกว่าหากมีความดันทางเข้าน้อย แสดงดังภาพที่ 1[3] มีการสั่นสะเทือนน้อยเนื่องจากมีความสมดุลในการหมุนของมอเตอร์ทำให้มีการสั่นที่น้อยและมีต้นทุนการใช้งานที่ต่ำ เนื่องจากมีจำนวนชิ้นส่วนที่เคลื่อนที่น้อยรับแรงคงที่ทำให้มีการใช้งานที่ยาวนานมากกว่าเครื่องอัดอากาศชนิดอื่น โดยส่วนใหญ่เครื่องอัดอากาศชนิดนี้จะใช้ในการขนส่งแก๊ส เนื่องจากปริมาตรที่สูงเมื่อเทียบกับอัตราการเพิ่มแรงดันรวมไปถึงต้นทุนที่ใช้ในการผลิตที่ต่ำกว่า ตัวอย่างตัวแปรที่นำมาใช้

เมื่อเครื่องจักรมีการใช้งานเป็นเวลานานจะเกิดความล้าขึ้น โดยจะแสดงอาการผิดปกติขึ้นมี 3 รูปแบบหลัก คือ 1. ความผิดปกติแบบทันทีทันใด 2. ความผิดปกติแบบค่อยเป็นค่อยไป และ 3. ความผิดปกติแบบไม่เป็นรูปแบบ โดยความผิดปกติที่พบเห็นมากที่สุดจะเกิดในรูปแบบค่อยเป็นค่อยไป เป็นความผิดปกติเนื่องมาจากการเสื่อมสภาพของเครื่องจักร การใช้งานที่ไม่เหมาะสมหรือสภาพแวดล้อมที่เปลี่ยนไป ส่งผลกระทบต่อแบบจำลองที่ใช้ในการตรวจจับความผิดปกติทำให้ประสิทธิภาพลดลงเมื่อเวลาผ่านไป เนื่องจากแบบจำลองส่วนใหญ่จะใช้กับข้อมูลแบบคงที่ (Static) การใช้แบบจำลองทั่วไปเมื่อใช้ไประยะเวลาหนึ่งประสิทธิภาพจะลดลงเนื่องจากข้อมูลที่มีการเปลี่ยนแปลง



รูป 1 ความสัมพันธ์ระหว่างอัตราการไหลกับความดันที่ทางออกของเครื่องอัดอากาศแต่ละชนิด

ในอุตสาหกรรม 4.0 เป็นยุคที่มีการนำเทคโนโลยีเข้ามาใช้ภาคอุตสาหกรรมมากขึ้น โดยจะมีการใช้ข้อมูลร่วมกับปัญญาประดิษฐ์ (Artificial Intelligence) นำมาวิเคราะห์การทำงานของระบบเพื่อตรวจจับความเสียหายที่เกิดขึ้นพร้อมทั้งระบุชนิดของความเสียหายโดยใช้การเรียนรู้ของเครื่อง (Machine Learning) สร้างแบบจำลองการเรียนรู้ให้กับระบบโดยหากมีข้อมูลมากความสามารถในการเรียนรู้ของเครื่องก็จะมากตาม แต่ในความเป็นจริงการเก็บรวบรวมข้อมูลให้มีคุณภาพนั้นทำได้ยาก เนื่องจากการรวบรวมข้อมูลที่มีความสมบูรณ์มีต้นทุนที่สูง จากปัจจัยดังกล่าวทำให้ข้อมูลที่บันทึกได้ส่วนใหญ่ไม่มีคุณภาพส่งผลให้เมื่อนำไปใช้ความแม่นยำจะน้อยลงจึงจำเป็นต้องนำไปผ่านกระบวนการเตรียมข้อมูล เพื่อให้ข้อมูลพร้อมใช้งานก่อนนำไปใช้งานได้ จึงได้เสนอการใช้ข้อมูลสังเคราะห์ (Synthetic data) เป็นข้อมูลที่สร้างขึ้นมาเพื่อจำลองเหตุการณ์ที่ต้องการเพื่อเรียนรู้เหตุการณ์ดังกล่าว โดยจะแสดงถึงค่าทางสถิติที่อ้างอิงจากความเป็นจริง อีกทั้งยังช่วยลดต้นทุนในการรวบรวมข้อมูลและปัญหาจากการรวบรวมที่ไม่มีคุณภาพ เช่น ข้อมูลไม่สมบูรณ์ เนื่องจากบางเหตุการณ์ไม่สามารถทำให้เกิดเพื่อเก็บข้อมูลหรือมีต้นทุนสูงในการสร้างเพื่อบันทึกข้อมูล จึงมีการสร้างข้อมูลสังเคราะห์ขึ้นเพื่อเข้ามาใช้แทนเหตุการณ์ดังกล่าว โดยสามารถสร้างเหตุการณ์ได้ไม่จำกัดตามที่ต้องการทดลองอีกทั้งช่วยแก้ปัญหาทางสถิติที่เกิดขึ้นจริงในธรรมชาติ เช่น ข้อมูลไม่ต่อเนื่อง ข้อมูลไม่มีการตอบสนอง โดยในการสร้างข้อมูลสังเคราะห์สามารถแบ่งออกเป็น 3 วิธีการหลักคือ 1. สร้างจากค่าทางสถิติที่เกิดขึ้นจริง 2. เพิ่มสัญญาณรบกวนไปยังข้อมูลจริง และ 3. ใช้การเรียนรู้ของเครื่องในการสร้างข้อมูลสังเคราะห์ โดยเหตุการณ์ที่จำลองขึ้นมาเป็นการจำลองความเสื่อมสภาพของ

อุปกรณ์ส่งผลให้ประสิทธิภาพการทำงานที่ลดลงเมื่อเวลาผ่านไป ปัญหาดังกล่าวสามารถแก้ไขโดยการเรียนรู้แบบหมู่ เนื่องจากมีการปรับปรุงค่าน้ำหนักตลอดทำให้มีความแม่นยำที่สูงกว่าเมื่อเวลาผ่านไป เมื่อเทียบการเรียนรู้ของเครื่องทั่วไปที่ประสิทธิภาพจะลดลงเมื่อเจอเหตุการณ์ดังกล่าว โดยการเรียนรู้แบบหมู่จะเป็นการนำหลายแบบจำลองมารวมผลลัพธ์เข้าด้วยกัน เพื่อให้ได้ผลลัพธ์ที่ดีกว่าผลลัพธ์ของแบบจำลองตัวเดียว เนื่องจากมีประสิทธิภาพที่ดีกว่าการใช้แบบจำลองเดี่ยว อีกทั้งยังสามารถจัดการปัญหาการพอกติเกินและการกระจายตัวของผลลัพธ์ ซึ่งเกิดได้ในแบบจำลองส่วนใหญ่ หลักจากนั้นจะความสามารถในการตรวจจับความผิดปกติที่เกิดขึ้นโดยใช้เมทริกซ์ความสับสน วัดค่าอัตราการตรวจจับผิด (False alarm rate) และอัตราการพลาดการตรวจจับ (Miss alarm rate) เป็นค่าที่แสดงถึงอัตราการทำนายผิดของแบบจำลอง เพื่อประเมินความสามารถของแบบจำลอง

ในวิทยานิพนธ์ฉบับนี้ได้นำเสนอวิธีการตรวจจับและวินิจฉัยความผิดปกติในชุดข้อมูลที่ไม่มีสถานะผิดปกติด้วยการเรียนรู้แบบหมู่ โดยใช้ชุดข้อมูลเครื่องอัดอากาศแบบหมุนเหวี่ยงที่มีแต่ข้อมูลสถานะปกติ จึงได้มีการสร้างชุดข้อมูลความเสียหายโดยอ้างอิงจากหลักการทำงานของเครื่องจักรนั้น และใช้ข้อมูลสังเคราะห์เป็นตัวแทนของข้อมูลความเสียหาย หลักจากนั้นให้แบบจำลองเรียนรู้พฤติกรรมความสัมพันธ์ของแต่ละตัวแปร และนำไปใช้ในการตรวจจับรูปแบบความเสียหายที่ได้ให้แบบจำลองเรียนรู้ เพื่อนำไปใช้ตรวจจับความผิดปกติที่คาดว่าจะเกิดในสถานการณ์จริง

## 1.2 วัตถุประสงค์ของวิทยานิพนธ์

1. เพื่อตรวจจับและวินิจฉัยความผิดปกติ โดยใช้การเรียนรู้ของเครื่องจักรกับข้อมูลที่มีมากกว่า 2 รูปแบบ
2. ประยุกต์ใช้การเรียนรู้แบบหมู่กับข้อมูลเครื่องอัดอากาศแบบหมุนเหวี่ยงจากโรงงานจริง โดยใช้ข้อมูลสังเคราะห์เพื่อเรียนรู้รูปแบบความผิดปกติ
3. เปรียบเทียบความสามารถในการตรวจจับสถานะผิดปกติโดยใช้ การเรียนรู้ของเครื่องกับการเรียนรู้แบบหมู่ โดยใช้เมทริกซ์ความสับสน (confusion matrix)

## 1.3 ขอบเขตของวิทยานิพนธ์

1. ใช้ชุดข้อมูลเครื่องอัดอากาศแบบหมุนเหวี่ยง (Ebara Model:200TC G-M) ชุดข้อมูลมีข้อมูลทั้งหมด 105,709 แถว 11 หลัก โดยมีการเก็บข้อมูลตั้งแต่วันที่ 24 กุมภาพันธ์ 2020 เวลา 13.00 น. ถึงวันที่ 25 กุมภาพันธ์ 2021 เวลา 13.55 น. โดยจะเก็บข้อมูลทุก 5 นาที และมีตัวแปร



ทั้งหมด 11 ตัว คือ กระแสไฟฟ้าของเครื่องอัดอากาศ, ความเร็วรอบการหมุนของเครื่องอัดอากาศ, กำลังไฟฟ้าของเครื่องอัดอากาศ, ความเร็วรอบใบพัด, อุณหภูมิแก๊สทางออก, อัตราการไหลแก๊สทางออก, ผลต่างความดันระหว่างทางออกและทางเข้า, พื้นที่การเปิดปิดทางเข้า, ความหนาแน่นแก๊ส, ปริมาณของเครื่องปฏิกรณ์, ผลต่างความดันระหว่างใบพัด

2. ตรวจจ็บบรูปแบบความผิดปกติ 3 รูปแบบที่ต้องการตรวจจ็บบ คือ อุตันที่ทางเข้า, ประสิทธิภาพลดลง, อุตันที่ทางออก

3. ใช้แบบจำลองการเรียนรู้แบบหมู่ในการตรวจจ็บบและวินิจฉัย โดยจะแบ่งเป็นเทคนิค Boosting และเทคนิค Bagging โดยเลือกอัลกอริทึมออกมาใช้ดังนี้ เทคนิค Boosting จะเลือกใช้วิธีการ XGBoost ในส่วนของ Bagging จะเลือกวิธีการ Random Forest ในการตรวจจ็บบความผิดปกติของเครื่องจักรพร้อมทั้งระบุรูปแบบความผิดปกติที่เกิดขึ้น

4. ข้อมูลเครื่องจักรไม่มีสถานการณ์ทำงานผิดปกติ จึงใช้ข้อมูลสังเคราะห์ในการสร้างข้อมูลที่ทำงานผิดปกติ เพื่อนำมาใช้ในการเรียนรู้ให้แบบจำลอง

#### 1.4 ขั้นตอนการดำเนินงานโดยสรุป

1. เก็บรวบรวมข้อมูลที่ได้จากเซนเซอร์ทั้งหมด 11 ตัว นำมาเข้าขั้นตอนการเตรียมข้อมูล (Data Preprocess) กำจัดค่าที่ไม่ต้องการ เพื่อให้ข้อมูลพร้อมใช้งาน

2. สร้างรูปแบบความผิดปกติให้กับข้อมูล เพื่อนำไปให้แบบจำลองเรียนรู้ และปรับค่าตัวแปร เพื่อให้มีความแม่นยำที่สูงขึ้น และนำไปทดสอบตรวจจ็บบความผิดปกติ

3. นำแบบจำลองไปทดสอบในชุดข้อมูลทดสอบ และประเมินความถูกต้อง

4. เปรียบเทียบความถูกต้องที่ได้จากแบบจำลองการเรียนรู้ของเครื่องทั่วไปกับการแบบจำลองการเรียนรู้แบบหมู่

## บทที่ 2

### ปรีทัศน์วรรณกรรม

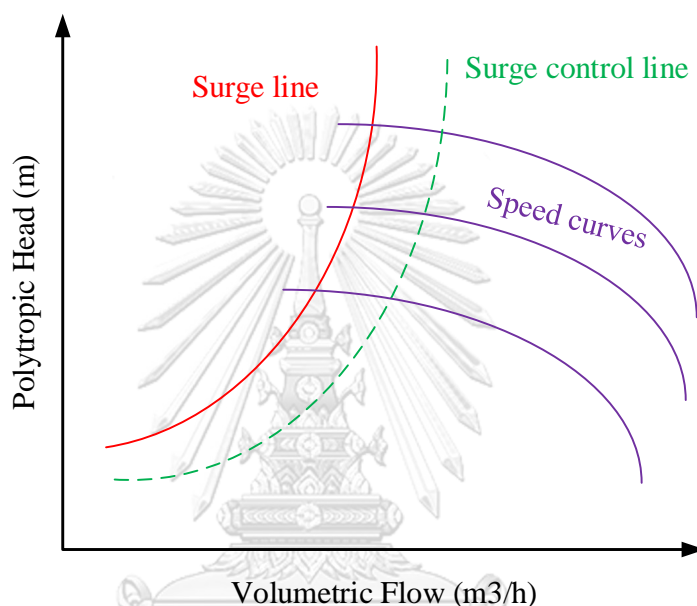
เนื้อหาในบทนี้จะรวบรวมผลงานวิจัยที่เกี่ยวข้องกับวิทยานิพนธ์ โดยใช้การเรียนรู้ของเครื่อง (machine learning) และการเรียนรู้แบบหมู่ นำมาประยุกต์ใช้เพื่อระบุรูปแบบการทำงานที่ผิดปกติของเครื่องอัดอากาศแบบหมุนเหวี่ยงที่ติดตั้งในอุตสาหกรรมเคมีแห่งหนึ่ง ในกระบวนการคัดแยกสารเคมี โดยบทนี้จะกล่าวถึง ลักษณะทางกายภาพของเครื่องอัดอากาศแบบหมุนเหวี่ยง (บทที่ 2.1) การตรวจจับและระบุรูปแบบความผิดปกติ (บทที่ 2.2) การตรวจจับความผิดปกติโดยใช้การเรียนรู้แบบหมู่ (บทที่ 2.3) และบทสุดท้ายที่ 2.4 เป็นการสรุปเนื้อหาจากการค้นหาเอกสารอ้างอิง

#### 2.1 ความผิดปกติในเครื่องอัดอากาศแบบหมุนเหวี่ยง

ในการขนส่งแก๊สในระยะทางไกลเครื่องอัดอากาศมักจะเป็นทางเลือกแรก เนื่องจากความสามารถในการเพิ่มความดันให้กับระบบและการเพิ่มอัตราการไหล ซึ่งสามารถทำได้มากกว่าเครื่องจักรชนิดอื่นๆ แต่เนื่องจากความต่อเนื่องในการขนส่งทำให้เครื่องอัดอากาศแบบหมุนเหวี่ยงเป็นเครื่องจักรที่มีการทำงานอย่างต่อเนื่องภายใต้สภาวะความดันและอุณหภูมิสูง รวมไปถึงสภาพแวดล้อมภายนอกที่ส่งผลให้มีโอกาสสูงที่จะเกิดความผิด จึงมีการวิเคราะห์ความผิดปกติที่เกิดขึ้นกับเครื่องอัดอากาศแบบหมุนเหวี่ยงขึ้นซึ่งสามารถเกิดได้จากหลายสาเหตุทั้งปัจจัยภายในตัวเครื่องจักรและปัจจัยภายนอก

หลักการทำงานของเครื่องอัดอากาศคือ ดึงแก๊สจากภายนอกเข้าเครื่องผ่านใบพัดเพิ่มความดันแล้วส่งออกไปยังทางออกไปยังกระบวนการถัดไปในระบบ โดยถ้าความต้านทานที่ทางออกของแก๊สมีค่าสูงขึ้น อัตราการไหลของแก๊สที่ทางเข้าจะลดลง ทำให้ความดันของแก๊สภายในจะเพิ่มขึ้นและถูกอัดตัวที่ทางออกของเครื่องอัดอากาศแต่ไม่สามารถไหลออกไปได้ โดยเครื่องจักรอาจจะยังสามารถใช้งานได้แต่หากเกินค่าที่กำหนด (Surge line limit) แสดงดังรูปที่ 2 [4] [5] [6] ส่งผลให้แก๊สที่ทางออกที่อัดด้วยความดันที่สูงเกิดการไหลย้อนกลับเข้าสู่ตัวเครื่องอัดอากาศ ปรากฏการณ์นี้เรียกว่า Surge หรือ Surging ซึ่งสามารถสร้างความเสียหายให้กับชิ้นส่วนภายในตัวเครื่องจักรได้ [7] [8] หากสารที่เข้ามาในเครื่องจักรเป็นของเหลวที่ไม่สามารถอัดตัวได้ ปริมาตรที่ไหลเข้ามาจะมีเท่ากับปริมาตรที่ไหลออก แต่แก๊สสามารถอัดตัวได้ทำให้ ปริมาตรแก๊สที่ไหลเข้าเครื่องจักรจะสูงกว่าปริมาตรที่ไหลออก ดังนั้นถ้าเป็นกรณีของเหลวที่ไม่สามารถไหลออกได้ เครื่องอัดอากาศจะไม่สามารถปล่อยของเหลวออกไปได้ อัตราการไหลจึงจะมีค่าเป็นศูนย์ ในกรณีที่สารภายในเป็นแก๊สที่สามารถอัดตัวได้

พบว่าแก๊สที่ทางออกจะไม่สามารถเคลื่อนออกไปได้ แต่ยังสามารถถึงถึงแก๊สเข้าไปในเครื่องจักรได้อยู่ เมื่อแก๊สขาออกเกิดการสะสมมากขึ้น จะเกิดการไหลย้อนสวนทางกับแก๊สที่ไหลเข้ามา ขณะที่เกิด Surge การไหลย้อนกลับของแก๊สด้านทางออกจะมีแรงดันที่มากกว่าด้านทางเข้ามาก ทำให้เอาชนะแก๊สที่ถูกดูดเขาไปได้ โดยปกติอัตราการไหลที่ต่ำที่สุดจะมีค่าเป็นศูนย์ แต่ในกรณีนี้อัตราการไหลต่ำสุดจะมีค่าน้อยกว่าศูนย์



รูป 2 การเกิดความดันย้อนกลับ

เพื่อป้องกันไม่ให้เกิดการไหลย้อนกลับจึงมีการศึกษาวิธีการคิดหาค่าที่สูงที่สุดที่สามารถใช้งานได้โดยการคำนวณหาค่าของการเกิดการไหลย้อนกลับ (Surge line) โดยคำนวณจากค่าสัมประสิทธิ์ของความดันและอัตราการไหล สมการที่ 1,2 ตามลำดับ

$$\psi = \frac{\Delta P}{((1/2) \rho U_m^2)} \quad (1)$$

$$\phi = \frac{V_z}{U_m} \quad (2)$$

โดย  $\psi$  คือ ค่าสัมประสิทธิ์ของความดัน

$\Delta P$  คือ ผลต่างความดันที่ผ่านเครื่องอัดอากาศ

$U_m$  คือ ความเร็วรอบของแกนหมุน

$\varphi$  คือ ค่าสัมประสิทธิ์ของอัตราการไหล

$V_z$  คือ ความเร็วตามแนวแกน

ความสัมพันธ์ระหว่างขนาดของพื้นที่ที่แก๊สไหลเข้ากับอัตราการไหลและความดันที่เปลี่ยนแปลงสามารถหาได้ดังนี้

$$\frac{Q_s}{\varphi_s \cdot S} = \rho \cdot U_m \quad (3)$$

$$\left(\frac{P_2}{P_1}\right) = \left(\frac{P_1 + \Delta P_s}{P_1}\right) \quad (4)$$

โดย  $P_1$  และ  $P_2$  คือ ความดันที่ทางเข้า และ ความดันที่ทางออกของเครื่องอัดอากาศ

$Q$  คือ อัตราการไหลที่ทางเข้า

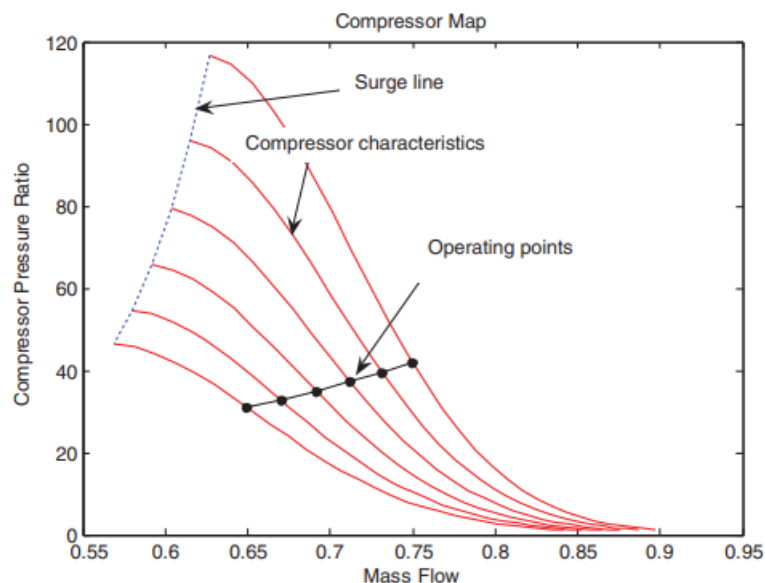
$S$  คือ ขนาดทางเข้าของเครื่องอัดอากาศ

จากนั้นแทนค่า  $\Delta P_s$  ในสมการที่ 4 ลงในสมการที่ 1 และ 2 จะได้เป็น

$$\left(\frac{P_2}{P_1}\right)_s - 1 = BQ_{s,s}^2 \quad (5)$$

$$B = \frac{1}{2S^2 \cdot \rho P_1} \left(\frac{\Psi_s}{\varphi_s^2}\right) \quad (6)$$

จากสมการที่ 5 จะได้ความสัมพันธ์ระหว่างอัตราส่วนความดัน (P) กับ อัตราการไหล (Q) ซึ่งมีค่าผันตามกัน โดยนำ B ตามสมการที่ 6 ไปแทนลงในสมการที่ 5 หลังจากนั้นนำไปพล็อตจะได้ Surge line ในรูปแบบพาราโบลา (เส้นป๊ะ) ดังรูปที่ 3 โดยจะแสดงความสัมพันธ์ระหว่างอัตราส่วนความดันกับ ปริมาตรการไหลของสาร[6]



รูป 3 เส้น Surge line

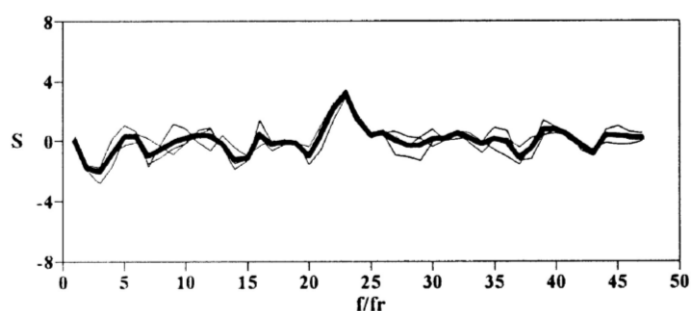
โดยสาเหตุการเกิดการไหลย้อนกลับ จะส่วนใหญ่จะเกิดจากการอุดตันของทางเดินแก๊สภายในตัวเครื่องที่ทางเข้าและทางออก สามารถเกิดได้จากหลายสาเหตุ เช่น อุณหภูมิในเครื่องอัดอากาศมีค่ามากกว่าจุดเดือดของแก๊สทำให้เกิดการควบแน่นกลายเป็นของเหลว, อุณหภูมิของแก๊สที่มาจากระบบก่อนหน้าที่ค่าสูงเมื่อลำเลียงผ่านท่อที่มีความยาวเกินไปทำให้แก๊สไม่ควบแน่นอย่างสมบูรณ์ ทำให้มีของเหลวหลังจากกลั่นตัวภายในท่อ, ความสกปรกภายในตัวอุปกรณ์ เมื่อมีสิ่งแปลกปลอมเข้าไปทำให้ขึ้นสว่นภายในสีกกร่อน หรือ ความดันภายในท่อที่สูงเกินไปทำให้ขึ้นสว่นภายในแตกร้าว ทำให้มีสิ่งเจือปนจากภายนอกเข้ามาอุดตัน รวมไปถึงการควบแน่นของสารภายในท่อ ที่สามารถทำให้เกิดการอุดตันของท่อได้ในอนาคต

## 2.2 การตรวจจับและระบุรูปแบบความผิดปกติ

วิธีการตรวจจับความผิดปกติของเครื่องจักรสามารถทำได้โดยใช้แบบจำลอง (model) เป็นการวิเคราะห์ข้อมูลที่ได้มาแต่ละตัวแปรเพื่อประมาณผลลัพธ์ที่ต้องการจากความสัมพันธ์ของแต่ละตัวแปรกับผลลัพธ์ โดยการตรวจจับสามารถแบ่งออกได้เป็น 3 วิธี คือ ตรวจจับโดยใช้ฐานความรู้ (Knowledge-based) ตรวจจับโดยใช้แบบจำลอง (Model-based) และ ตรวจจับโดยใช้ข้อมูล (Data-based)

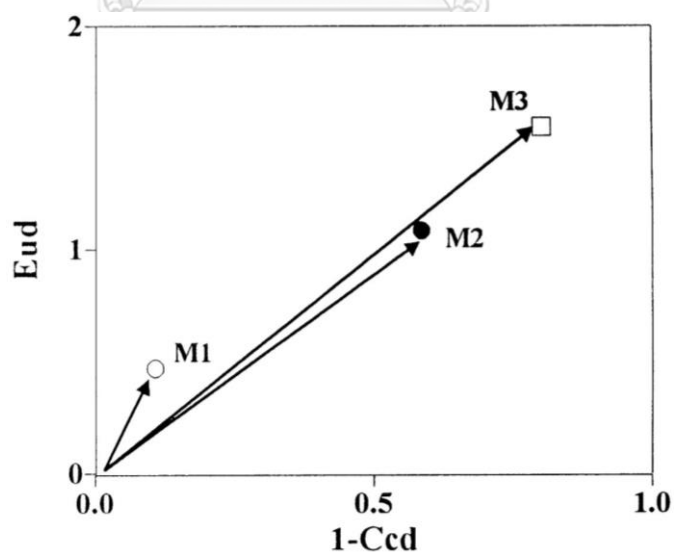
การจดจำรูปแบบ ( Pattern recognition ) เป็นสาขาวิชาหนึ่งในทางวิทยาศาสตร์ที่มีวัตถุประสงค์เพื่อแบ่งแยกวัตถุตามหมวดหมู่หรือตัวเลข โดยวัตถุดังกล่าวอาจจะเป็นรูปภาพ, สัญญาณ

หรือ สิ่งอื่นที่ต้องการจะแบ่งแยก โดยจะใช้รูปแบบในการแบ่งสิ่งนั้นๆ โดยมีการนำมาใช้ในปี 1998 โดย Aretakis [9] เพื่อแบ่งแยกชนิดของความเสียหายโดยการบันทึกว่า การสั่น และ เสียง เพื่อนำมาวิเคราะห์โดยทำได้หลายวิธี เช่น 1. การจดจำรูปแบบจากลักษณะเด่นที่เกิดขึ้น เป็นการกำหนดรูปแบบความเสียหายเด่นเพื่อนำมาเป็นจุดที่นำมาอ้างอิงในการจดจำรูปแบบ แสดงดังรูปที่ 4 โดยเส้นสีเข้มจะเป็นรูปแบบที่นำมาใช้อ้างอิงและเส้นอื่นจะเป็นรูปแบบที่จะนำมาเปรียบเทียบเพื่อวิเคราะห์



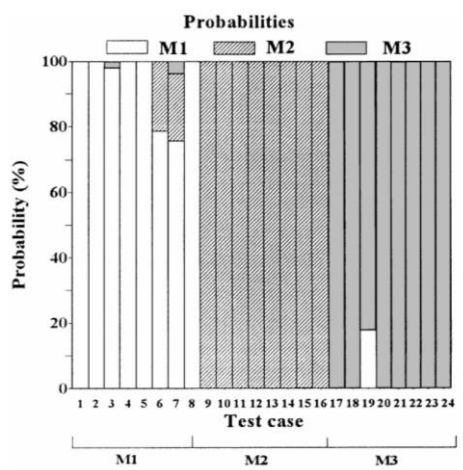
รูป 4 สเปกตรัมระหว่างอ้างอิงกับข้อมูลที่ต้องการจำแนก

2. การจดจำรูปแบบโดยเรขาคณิต เป็นการคัดแยกรูปแบบโดยใช้การวัดระยะทางแบบยูคลิด (Euclidean Distance) เพื่อแยกความเหมือนความต่างของข้อมูล แสดงดังรูปที่ 5 จะพบว่าความเสียหายของ M1 มีค่าระยะทางที่ใกล้เคียงกับรูปแบบอ้างอิงมากที่สุด



รูป 5 การแบ่งข้อมูลโดยใช้ระยะทางแบบยูคลิด

3. การจำจํารูปแบบโดยใช้สถิติ ใช้ตัวแปรที่คล้ายกันนำมาคำนวณทางสถิติเพื่อหาความเหมือนความต่างของข้อมูลในรูปของความน่าจะเป็น แสดงดังรูปที่ 6 พบว่าชุดข้อมูลที่ 7 มีความถูกต้องน้อยที่สุด เนื่องจากจำแนกถูกต้องว่าเป็น M1 75% แต่จำแนกว่าเป็น M2 20%, M3 5%



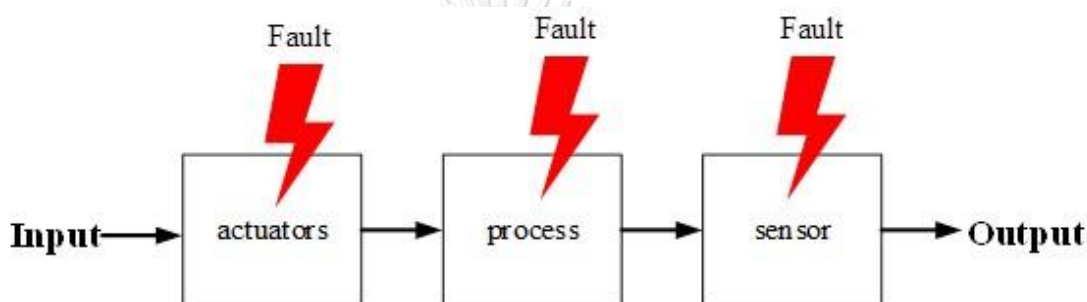
รูป 6 การจดจํารูปแบบโดยใช้สถิติ

จะเห็นได้ว่าเทคนิคการคัดแยกแบบนี้มีประสิทธิภาพที่ต่างกัน จากการทดลองพบว่าวิธีการคัดแยกด้วยการจดจํารูปแบบให้ผลลัพธ์ในทางบวกมากกว่าทางลบมากกว่า แต่อย่างไรก็ตามจำนวนของคำตอบที่ถูกก็มีปัจจัยที่ต่างกันออกไปขึ้นกับชนิด, รูปแบบ รวมไปถึงความซับซ้อนของข้อมูล Clement J.D. [10] , ทำการจำลองข้อมูลสถานะความเสียหายเพื่อใช้ตรวจจับความเสียหายเนื่องจากข้อมูลอยู่ในสถานะปกติทั้งหมด จึงจำลองสถานะความเสียหายโดยเพิ่มค่าของข้อมูลสามเท่าของส่วนเบี่ยงเบนมาตรฐาน เพื่อสร้างข้อมูลที่มีสถานะความเสียหายให้ระบบทำจดจํารูปแบบแล้วนำข้อมูลอีกชุดที่มีการเพิ่มค่าสองเท่าของความแปรปรวน เนื่องจากมีค่าที่เพิ่มไปไม่เท่ากันเพื่อว่าระบบสามารถตรวจจับได้หรือไม่

ตรวจจับความเสียหายโดยใช้ฐานความรู้ (Knowledge-based) เป็นวิธีการวิเคราะห์เชิงคุณภาพ ( Quantitative analytical ) ที่มีพื้นฐานจากความรู้ที่มี ศึกษาและเข้าใจถึงหลักการทำงานของอุปกรณ์ เพื่อทำความเข้าใจตัวแปรที่ส่งผลให้เกิดความเสียหายต่อ Paul M [11] และทำการตั้งเงื่อนไขเป็นกฎโดยใช้หลักการทางสถิติในการคำนวณค่าระบุสถานะเตือนระดับความอันตรายที่จะเกิดขึ้น Mohammadreza Tahan [12] ได้ใช้ การซ่อมบำรุงโดยใช้ Condition-based maintenance (CBM) เป็นการตั้งเงื่อนไขในการดูแลรักษาโดยจะมีการใช้ข้อมูลประกอบการวิเคราะห์ร่วมกันผู้เชี่ยวชาญเพื่อให้ได้ความแม่นยำที่มากยิ่งขึ้น อีกทั้งยังสามารถอธิบายลักษณะของ

ความเสียหายที่เกิดขึ้นที่มนุษย์ไม่สามารถสังเกตได้ด้วยตาเปล่า แต่จะแสดงจากตัวแปรที่มีผลต่อความเสียหายรูปแบบนั้นๆ

การตรวจจับโดยใช้ข้อมูล (Data-based) เป็นวิธีการตรวจจับโดยใช้ข้อมูลที่เก็บรวบรวมวิเคราะห์ถึงลักษณะเฉพาะของข้อมูลหรือพฤติกรรมที่เกิดขึ้น โดยส่วนประกอบของระบบ ส่วนใหญ่จะประกอบด้วย อุปกรณ์ทำงาน (Actuator) ตัวรับรู้ (Sensor) และ ระบบ (Process) ดังนั้นความผิดปกติจึงสามารถเกิดได้ 3 ตำแหน่ง คือ อุปกรณ์ทำงาน (Actuator fault) ตัวรับรู้ (Sensor fault) ระบบ (Parameter or structure fault) แสดงดังรูปที่ 7



รูป 7 ตำแหน่งความผิดปกติที่เกิด

หลายอุปกรณ์ที่วัดค่าสัญญาณหรือข้อมูลจะแสดงถึงลักษณะของข้อมูล ซึ่งถ้าหากเกิดการเปลี่ยนแปลงของค่าจากเดิมจะแสดงถึงความผิดปกติของตัว อุปกรณ์ ตัวรับรู้ และ ตัวระบบ การตรวจจับความผิดปกติของข้อมูลโดยใช้ข้อมูลจึงถูกนำมาใช้ โดยอาจจะใช้กับการสัน ความเร็ว ความเร่งของเครื่องจักรเพื่อหาความผิดปกติเช่น ปัญหาการไม่สมดุล (Imbalance) การกระทบกันของเครื่องจักร โดยในการตรวจจับจะใช้การเปรียบเทียบพฤติกรรมของข้อมูลในสถานะปกติและนำมาให้ปัญญาประดิษฐ์เรียนรู้เพื่อจดจำรูปแบบข้อมูล เพื่อให้ในการคัดแยกครั้งถัดไป โดยวิธีการตรวจจับโดยข้อมูลเป็นที่นิยมอย่างมากรวมถึงมีการพัฒนาและวิจัยอย่างต่อเนื่องทำให้มีวิธีการในการตรวจจับที่มากขึ้น Wimalin L. [13] ใช้โครงข่ายประสาทเทียม (Neural Network), กฎการอุปนัย (Rule Induction), วิธีการเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors) ในการจดจำรูปแบบในลักษณะของแนวโน้ม วัฏจักร และการเปลี่ยนไปที่ละนิด พบว่ากฎการอุปนัยได้ผลลัพธ์ที่ดีที่สุดอีกทั้งสามารถอธิบายข้อมูลที่ซับซ้อนด้วยความสัมพันธ์ของกฎพื้นฐานทำให้เข้าใจง่ายแต่ข้อเสียคือให้ผลลัพธ์ไม่ดี หากข้อมูลสัญญาณรบกวนมาก Yang Z. [14] ใช้วิธี Support vector data description (SVDD) และนำมาเปรียบเทียบกับ วิธีอื่นเช่น Support Vector Machine (SVM) และการวิเคราะห์องค์ประกอบหลัก (Principle Component Analysis (PCA)) พบว่า SVDD สามารถระบุได้ 90%



แต่วิธีการนี้จะใช้โดยวิธีการเดียวไม่ได้จำเป็นต้องมีการใช้ร่วมกันกับวิธีการอื่นเพื่อให้ความแม่นยำที่มากขึ้น แต่ในปี 2019 Yang Y. [15] ได้ใช้ PCA แล้วพบว่าในการคำนวณตำแหน่งของแต่ละค่าใกล้เคียงกันแต่มีค่าสูญเสีย (loses) ที่ไม่เท่ากันซึ่งการวิเคราะห์องค์ประกอบหลักสัมพันธ์ (Relative Principle Component Analysis (RPCA)) สามารถแก้ไขปัญหาดังกล่าวได้แต่มีการคำนวณค่าน้ำหนักของแต่ละข้อมูลที่เยอะจึงได้พัฒนาต่อไปเป็น Mutual Separately Relative Principle Component Analysis (MI-RPCA) สามารถจัดการหาค่าของแต่ละข้อมูลได้โดยมีความแม่นยำที่มากกว่า

นอกจากนี้ยังได้มีการนำเสนอวิธีต่าง ๆ เพื่อใช้ในการตรวจจับความผิดปกติของเครื่องจักร เช่น ในปี 2019 Xiaochuan Li [16] ได้ตรวจจับความเสียหายในช่วงเริ่มต้น (Incipient fault) และทำนายอายุการใช้ที่เหลือของเครื่องอัดอากาศในอุตสาหกรรม โดยใช้การวิเคราะห์ตัวแปรคานอนนิคอล (Canonical variate analysis (CVA)) เพื่อจำแนกความเสียหายแบบเริ่มต้น ซึ่งเป็นการรวมกันระหว่างวิธี Exponentially Weighted Moving Average (EWMA) และ Pearson correlation ในการตรวจจับความผิดปกติ และใช้วิธี metabolism grey forecasting model (MGFM) ในการทำนายอายุการใช้งานที่เหลือ (Remaining Useful life) มีการใช้การเรียนรู้ของเครื่องจักร (Machine Learning) และ การเรียนรู้เชิงลึก (Deep Learning) มากขึ้นในปี 2020 Mohamed [17] ต้องการสร้างเครื่องมือที่ตรวจจับและวินิจฉัยการทำงานที่ผิดปกติโดยใช้ Supervised และ Semi-supervised ทำงานร่วมกันแบบคู่ขนาน โดยจะใช้ Supervised ในการวินิจฉัย เนื่องจากมีความแม่นยำมากกว่าโดยใช้ Random forest (RF) แล้วใช้ Unsupervised ในส่วนของการตรวจจับโดยใช้ PCA และ M-RBC ซึ่งได้ผลลัพธ์ที่ดีกว่าเมื่อทำงานแบบคู่ขนาน

จากงานวิจัยที่กล่าวไปทั้งหมดเป็นที่มาของวิทยานิพนธ์นี้ ในการตรวจจับความเสียหายด้วยวิธีการผสมระหว่างการใช้ฐานความรู้ที่มีเฉพาะด้านในการกำหนดตัวแปรที่ก่อให้เกิดความเสียหายในรูปแบบที่ต้องการตรวจจับแล้วใช้การเรียนรู้ของเครื่องจักรทำการเรียนรู้ จัดจำรูปแบบของข้อมูลเพื่อนำไปแยกแยะข้อมูลต่อไป โดยวิธีการเรียนรู้ที่จะใช้คือการเรียนรู้แบบรวมหมู่ในการแยกแยะรูปแบบข้อมูลที่ได้เรียนรู้เนื่องจากมีประสิทธิภาพมากกว่าการแยกแยะโดยใช้วิธีการเดียวแยกแยะข้อมูล

### 2.3 การตรวจจับความผิดปกติโดยใช้การเรียนรู้แบบหมู่

การเรียนรู้ของเครื่องจักรในการตรวจจับความเสียหายมีความสำคัญอย่างมากเนื่องจากมนุษย์ไม่มีศักยภาพพอที่จะสามารถวิเคราะห์ได้ การเรียนรู้ของเครื่องจักรจึงมีความสำคัญและมีการศึกษาพัฒนาอยู่ตลอด โดยการเพิ่มประสิทธิภาพวิธีหนึ่งคือ การรวมกันของหลายวิธีการเข้าด้วยกัน หรือเรียกว่า “การเรียนรู้แบบหมู่” เป็นวิธีการที่รวมการเรียนรู้ของเครื่องจักรหลากหลายวิธีเข้าด้วยกัน ซึ่งสามารถกำจัดปัญหาที่ทำให้ประสิทธิภาพลดลง เช่น ความไม่สมดุลของชนิดข้อมูล (Class Imbalance), การเบนออกของข้อมูล (Concept drift) และ ปัญหาจากขนาดของข้อมูล (Curse of dimensionality) โดยสามารถแบ่งได้ 2 เทคนิคหลัก คือ Bagging และ Boosting

เทคนิค Bagging เป็นเทคนิคที่ลดความแปรปรวนของคำตอบ โดยสร้างชุดข้อมูลใหม่โดยสุ่มจากข้อมูลที่มีเพื่อที่จะนำมาใช้แทนข้อมูลที่มี เพื่อเลี่ยงปัญหาการใช้ข้อมูลเดิมในการเรียนรู้ ทำให้สามารถแก้ปัญหาระบบจดจำข้อมูลในการแบ่งแยกชนิด แทนที่ระบบจะเรียนรู้เพื่อที่จะวิเคราะห์ความสัมพันธ์ของข้อมูล เพื่อที่จะแบ่งแยกชนิดของข้อมูล ช่วยลดปัญหาข้อมูลไม่สมดุล วิธีการที่ใช้ Bagging ที่พบได้ทั่วไปคือ Random Forest (RF) และ Bagging โดยในปี 2010 Aldrich[18] ได้ตรวจจับและวินิจฉัยความผิดปกติโดยใช้การสกัดคุณลักษณะ (Feature extraction) โดยใช้ Random forest (RF) และใช้ตัวแปรที่สำคัญ (Variable importance) เพื่อวัดและนำมาใช้ในการวินิจฉัยความผิดปกติ เนื่องจาก RF มีอัตราการแจ้งเตือนผิดพลาด (Missing Alarm rate) ที่ต่ำ Luca [19] จึงได้ใช้ Random Forest เพื่อระบุรูปแบบความผิดปกติของแผ่นเวเฟอร์ของสารกึ่งตัวนำโดยใช้กับชุดข้อมูลจากการจำลองและชุดข้อมูลจากอุตสาหกรรมจริง โดยจะวัดระยะของข้อมูลโดยคำนวณจากค่าเฉลี่ย ของส่วนเบี่ยงเบนมาตรฐานของชุดข้อมูลปกตินำมาคำนวณค่าขอบเขตโดยจะกำหนดไว้ว่าถ้า ค่าเฉลี่ยบวกด้วยสามเท่าของส่วนเบี่ยงเบนมาตรฐาน หากข้อมูลมีค่าเกินกว่าที่กำหนดจะถือว่าเป็นชุดข้อมูลที่มีความผิดปกติ โดย Radhia[19] เห็นว่า RF เมื่อนำมาใช้กับชุดข้อมูลดิบจะให้ประสิทธิภาพที่ไม่ดี เนื่องจากข้อมูลมีจำนวนมากเกินไปและมีค่ารบกวน จึงได้ปรับปรุง RF เพื่อเอาชนะปัญหาที่กล่าวมาโดยการลดจำนวนข้อมูลที่เข้ามาในการเรียนรู้ (Training) โดยเลือกองค์ประกอบหลักของเคอร์เนล (Kernel principal components (KPCs)) โดยใช้การวิเคราะห์องค์ประกอบหลัก (PCA) เพื่อลดขนาดมิติของข้อมูล จากนั้น KPCs จะถูกนำไปแบ่งแยกโดยใช้ RF เพื่อวินิจฉัยความผิดปกติ ในปี 2020 Goksu [20] ได้ศึกษาวิธีการ Bagging พบว่ามีปัญหาในการสุ่มเลือกในการสร้างชุดข้อมูลใหม่ (Bootstrap Sample) เนื่องจากประสิทธิภาพเนื่องบางชุดข้อมูลมีความเหมาะสมสำหรับการเรียนรู้ บางชุดไม่เหมาะสม จึงได้ทำการปรับปรุง Bagging ใหม่และให้ชื่อ

ว่า Enhanced Bagging (eBagging) เนื่องจากการสุ่มข้อมูลเพื่อใช้เป็นชุดข้อมูลในการเรียนรู้มีความน่าจะเป็น 36.8% ที่ข้อมูลจะไม่ถูกเลือกไปใช้ทำให้ไม่สมดุลระหว่างข้อมูล ดังนั้นจึงแก้ไขโดยใช้ข้อมูลทั้ง 2 ประเภทลงไปในชุดข้อมูลสุ่ม เมื่อเทียบกับวิธีการ Bagging, Random Forest (RF), Boosting, Support Vector Machines (SVM), Decision Tree (C4.5), k-Nearest Neighbor (kNN) และ Naïve Bayes (NB) โดยใช้ชุดข้อมูลทั้งหมด 33 ชุดข้อมูลในการทดสอบ

เทคนิค Boosting เป็นเทคนิคที่ช่วยลดความผิดพลาดของค่าคำตอบ ทำได้โดยปรับค่าน้ำหนัก (Weight) หลังจากเรียนรู้เสร็จโดยหากแยกแยะผิดค่าน้ำหนักจะถูกปรับให้น้อยลงเป็นการเรียนรู้ข้อผิดพลาดจากการกระทำก่อนหน้าทำให้มีความแม่นยำที่สูงขึ้นให้กับระบบ โดยในปี 2018 Dahai[21] ได้ตรวจจับความผิดปกติของกังหันลมโดยใช้ RF ร่วมกับ XGBoost โดย RF ใช้ในการจัดลำดับความสำคัญของตัวแปร (Feature Importance) ที่เก็บมาจากเซนเซอร์หรือเก็บจากตัวแปรที่ได้จากความรู้ แล้วนำตัวแปรที่มีความสำคัญมากที่สุดสามอันดับแรกมาให้ตัวระบบเรียนรู้โดยใช้วิธี XGBoost เมื่อเทียบกับวิธีการ SVM พบว่า วิธีการใช้ RF ร่วมกับ XGBoost ผลลัพธ์ที่ดีกว่า โดยปัญหาอย่างหนึ่งที่สำคัญคือการเรียนรู้ไม่สมดุล (Imbalance Learning) Uma R. [22] ซึ่งเกิดจากมีชุดลักษณะของข้อมูลไม่เท่ากัน ทำให้ประสิทธิภาพในการแบ่งประเภทมีค่าลดลง การสุ่มข้อมูลใหม่ (Re-sampling) ของข้อมูลที่ไม่สมดุลจะช่วยควบคุมการกระจายตัวของข้อมูลได้ทำให้ข้อมูลมีความสมดุลเพิ่มมากขึ้น แต่บางครั้งก็ข้อมูลที่จำเป็นอาจไม่ถูกเลือกทำให้เกิดปัญหาเกินพอดี (Overfitted) เป็นปัญหาที่ข้อมูลจดจำแทนการเรียนรู้ทำให้เมื่อได้รับชุดข้อมูลใหม่ที่ไม่เหมือนข้อมูลที่จดจำมาจะไม่สามารถแบ่งประเภทข้อมูลได้ โดยการจัดการปัญหาความไม่สมดุลจะอยู่ในขั้นตอนก่อนกระบวนการ (Pre-Processing) การเลือกคุณลักษณะ (Feature selection) เป็นอีกวิธีการหนึ่งที่ใช้เพื่อลดข้อมูลที่ไม่จำเป็นก่อนเข้ากระบวนการ Deleplace [23] ได้ใช้การเลือกคุณลักษณะในการเลือกตัวแปรที่นำมาใช้ก่อนนำเข้ากระบวนการวิเคราะห์เพื่อช่วยเพิ่มประสิทธิภาพของแบบจำลองและลดเวลาในการเรียนรู้เนื่องจากมีข้อมูลที่ลดลง ในปี 2020 Priyanka[24] ได้นำวิธีการของ Ensemble มาโหวตผลลัพธ์สุดท้ายในการวินิจฉัยความผิดปกติของลูกปืน โดยวิธีการที่เลือกใช้คือ RF, AdaBoost, GB, XGBoost และ Extra tree นำมาปรับค่าตัวแปรที่ดีที่สุดโดยใช้วิธีการค้นหาตาราง (Grid search) จากนั้นนำมาโหวตเสียงส่วนมาก

จากงานวิจัยที่กล่าวไปทั้งหมดเป็นที่มาของวิทยานิพนธ์นี้ ในการตรวจจับความเสียหายโดยใช้วิธีการเรียนรู้แบบกลุ่ม ซึ่งจะได้ผลลัพธ์ที่ดีกว่าการใช้แบบจำลองเพียงแบบจำลองเดียว อีกทั้งช่วยลด

ปัญหาเกินความพอดี เพิ่มประสิทธิภาพในการตรวจจับความผิดปกติ และ มีความทนทาน (Robustness) เป็นการลดการกระจายตัวของค่าผลลัพธ์ ทำให้ค่าผลลัพธ์ที่ได้มีความแม่นยำมากขึ้น



### บทที่ 3

#### ทฤษฎีพื้นฐาน

#### 3.1 เครื่องอัดอากาศแบบหมุนเหวี่ยง (Centrifugal Compressor)

เครื่องอัดอากาศเป็นอุปกรณ์ที่ทำหน้าที่เพิ่มความดัน[25] [26] ซึ่งมีการนำไปใช้กันอย่างแพร่หลายในหลายกลุ่มอุตสาหกรรมเนื่องจากกระบวนการผลิตจำเป็นจะต้องมีการเพิ่มความดันในระบบโดยการเปลี่ยนพลังงานจลน์หรือให้การบีบอัดตัวของอากาศเพื่อเพิ่มความดันตามสมการ

$$-Q + m_i \left( u_i + \frac{V_i^2}{2} + gz_i \right) = -w + m \left( u_e + \frac{V_e^2}{2} + gz_e \right) \quad (3.1.1)$$

โดย  $Q$  คือ อัตราการถ่ายเทความร้อนของเครื่องอัดอากาศ

$w$  คือ อัตราการถ่ายเทงานที่เกิดจากเครื่องอัดอากาศ

$m$  คือ อัตราการถ่ายเทมวลของสารที่ใช้

$V_i, V_e$  คือ ความเร็วของสารที่ใช้ของทางเข้ากับทางออก

$Z_i, Z_e$  คือ ความสูงของทางเข้ากับทางออก

$U_i, U_e$  คือ พลังงานภายในของทางเข้ากับทางออก

$$\text{หรือ} \quad \frac{dE_k}{dt} + \frac{dE_p}{dt} + \frac{dU}{dt} = \sum_{input} m_i \left( H_i + \frac{u_i^2}{2} + g_i \right) - \sum_{output} m_j \left( H_j + \frac{u_j^2}{2} + g_j \right) + Q - W_s \quad (3.1.2)$$

โดย  $E_k$  คือ พลังงานจลน์ของระบบ

$E_p$  คือ พลังงานศักย์ของระบบ

$U$  คือ พลังงานภายในระบบ

โดยจะมีงานที่เกิดจากการไหลของสาร  $j$  ดังสมการ  $W_{f,j} = P_j V_j$  (3.1.3) ทำให้เกิดความดันขึ้น ซึ่งเป็นพลังงานภายในของสาร  $j$  ดังสมการ  $H_j = U_j + P_j V_j$  (3.1.4) เมื่อนำสมการ 3.1.4 มาใส่เพิ่มในสมการ 3.1.2 จะได้เป็น

$$-Q + m \left( U_i + P_i V_i + \frac{V_i^2}{2} + gZ_i \right) = -W_f + m \left( U_e + P_e V_e + \frac{V_e^2}{2} + gZ_e \right) \quad (3.1.5)$$

จะได้เป็นสมการพลังงานของระบบระหว่างทางเข้าจนถึงทางออก ซึ่งตัวอย่างการใช้งานเช่น ในอุตสาหกรรมเคมีการใช้เครื่องอัดอากาศในการขนส่งแก๊สเพื่อเพิ่มความดันในระบบต่าง ๆ จะส่งผลให้ค่า P มีค่าเพิ่มขึ้น ดังนั้นตัวแปร V ก็อาจจะมีแนวโน้มที่ลดลงตามสมการ โดยทั่วไปเครื่องอัดอากาศแบ่งได้ 2 ประเภท คือ ประเภทปริมาตรแทนที่เชิงบวก และประเภทไดนามิก



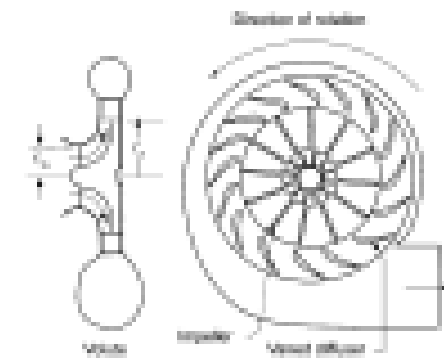
รูป 8 เครื่องอัดอากาศแบบลูกสูบ

เครื่องอัดอากาศประเภทปริมาตรแทนที่เชิงบวก(Positive Displacement compressors) มีหลักการทำงานคือ เมื่อลูกสูบเคลื่อนที่ลงจะดูดอากาศเข้าไปภายในช่องว่างและเกิดแรงเหวี่ยงเพื่อที่จะอัดอากาศเป็นการนำอากาศมาแทนที่ปริมาตรช่องว่างโดยการเคลื่อนของลูกสูบจึงเป็นที่มาของชื่อ โดยการอัดอากาศจะเกิดในช่วงที่ลูกสูบเคลื่อนที่ขึ้น โดยอากาศจะถูกอัดด้วยลูกสูบทำให้ปริมาตรของอากาศลดลงและความดันสูงขึ้น โดยเครื่องอัดอากาศแบบนี้ที่นิยมคือ เครื่องอัดอากาศแบบลูกสูบ (Reciprocating Compressors) ดังแสดงในรูปที่ 8 เครื่องอัดอากาศชนิดนี้ใช้กับอัตราการไหลต่ำ ไม่สม่ำเสมอ เหมาะกับการใช้งานที่โหลดไม่สม่ำเสมอ

เครื่องอัดอากาศแบบไดนามิก (Dynamic Compressors) มีหลักการทำงานคือให้พลังงานกลแก่อากาศเพื่อให้อากาศมีความเร็วที่สูงขึ้น จากนั้นอากาศจะถูกดูดเข้าไปในตัวเครื่องทำให้เกิดการเปลี่ยนแปลงจากพลังงานจลน์เป็นแรงดันของอากาศ เครื่องอัดอากาศประเภทนี้ได้แก่ เครื่องอัดอากาศแบบหมุนเหวี่ยง (Centrifugal Compressor) ลักษณะการใช้งานของเครื่องอัดอากาศชนิดนี้คือ ใช้กับงานที่มีอัตราการไหลสูง เหมาะกับงานที่มีโหลดคงที่ เนื่องจากการที่มีอัตราการไหลที่สูงทำให้นิยมนำมาใช้กับอุตสาหกรรมเคมีในการขนส่งแก๊ส

เครื่องอัดอากาศแบบหมุนเหวี่ยง (Centrifugal Compressor) เป็นเครื่องอัดอากาศที่เปลี่ยนพลังงานจลน์เป็นแรงดัน ส่วนใหญ่จะถูกนำไปใช้ในการขนถ่ายแก๊สธรรมชาติ โดยอากาศจะถูกดูดเข้าไปยังแกนกลางเพลลาใบพัดซึ่งมีครีบนำน้ำ (Guide Vane Position) เป็นตัวควบคุมปริมาณอากาศที่

จะเข้าไปยังแกนกลางของเพลา และจะถูกเหวี่ยงออกตามแนวรัศมีใบพัด (Impeller) ดังแสดงในรูปที่ 9 ทำให้อากาศเกิดความเร่งจากการเหวี่ยงของใบพัดไปโดยในขั้นตอนนี้จะทำให้อุณหภูมิและความดันของอากาศมีค่าสูงขึ้นแต่ปริมาตรมีค่าลดลง เนื่องจากการที่อุณหภูมิสูงขึ้นทำให้ต้องมีการระบายความร้อนโดยการใช้ น้ำ หรือ อากาศ มาช่วยในการระบายความร้อนเพื่อนำไปใช้ต่อที่ขั้นตอนต่อไป



รูป 9 เครื่องอัดอากาศแบบหมุนเหวี่ยง

โดยในงานวิทยานิพนธ์นี้จะใช้ข้อมูลของเครื่องอัดอากาศแบบหมุนเหวี่ยงในการทดสอบการตรวจจับความผิดปกติในรูปแบบความผิดปกติที่ต่างกัันตามที่กำหนด มี 3 รูปแบบ คือ ปัญหา อุดตันที่ทางเข้า ปัญหาประสิทธิภาพพลดลง ปัญหาการอุดตันที่ทางออก ซึ่งสามารถทำให้เกิดความผิดปกติต่อเครื่องจักรได้หากปล่อยให้เครื่องจักรทำงานในสถานะผิดปกติดังกล่าวเป็นเวลานาน ปัญหาดังกล่าวเป็นปัญหาที่เกิดขึ้นภายในตัวเครื่องจักรทำให้ไม่สามารถสังเกตความผิดปกติได้จากภายนอก ส่งผลให้ต้องใช้ค่าที่ได้จากเซนเซอร์มาทำการวิเคราะห์ผ่านตัวแปรที่ต่างกััน แต่เนื่องจากแต่ละปัญหามีลักษณะเฉพาะที่ต่างกัันทำให้สามารถสังเกตพฤติกรรมที่ต่างกัันของปัญหาแต่ละรูปแบบโดยจะสังเกตได้จากพฤติกรรมของตัวแปรของเครื่องจักรที่วัดได้จากเซนเซอร์

### 3.2 ระบบการวินิจฉัยความผิดปกติ (fault diagnosis system) [29]

ในการทำงานของเครื่องจักรมีโอกาสที่จะเกิดความผิดพลาดได้ตลอดมีหลายปัจจัยอาจเกิดมาจาก ปัจจัยภายใน (การขาดสารหล่อลื่น, การเสียดสีที่มากเกินไป, อุณหภูมิภายในที่มากเกินไป, การรั่ว, การสึกกร่อน) ปัจจัยภายนอก(ความชื้น, ฝุ่น, ปฏิกริยาเคมี) เป็นสาเหตุที่ทำให้เกิดความผิดปกติโดยมีการเปลี่ยนแปลงของพฤติกรรมของเครื่องจักรที่เล็กน้อย[27] ซึ่งจำเป็นต้องมีการควบคุมโดยการตั้งจุดที่เป็นค่ามากที่สุดที่เป็นไปได้ โดยจะให้ข้อมูลอยู่ในช่วงที่กำหนด  $Y(t)$  โดย  $Y_{min} < Y(t) < Y_{max}$  หากเกินค่าที่กำหนดจะถือว่าเป็น ความผิดปกติ (Fault) ซึ่งเป็นค่าที่เกินกว่าที่กำหนดไว้ หรือ เกินกว่าปกติ

แบบผิดปกติจากมาตรฐาน โดยความผิดปกติที่เกิดขึ้นอาจทำให้เกิดการทำงานที่ผิดปกติเป็นสาเหตุให้ไม่สามารถทำงานได้เต็มประสิทธิภาพ โดยสามารถเกิดได้จากหลายปัจจัย เช่น การออกแบบที่ผิดพลาด การประกอบที่ไม่สมบูรณ์ การสึกกร่อน การใช้เกินกว่าที่จะรับได้ รวมไปถึงความผิดปกติจากคนควบคุมซึ่งเป็นสาเหตุมาจากมนุษย์ที่ทำงานผิดพลาด (Human error) โดยจะมีการจัดการความเสี่ยงตามความผิดปกติดังนี้

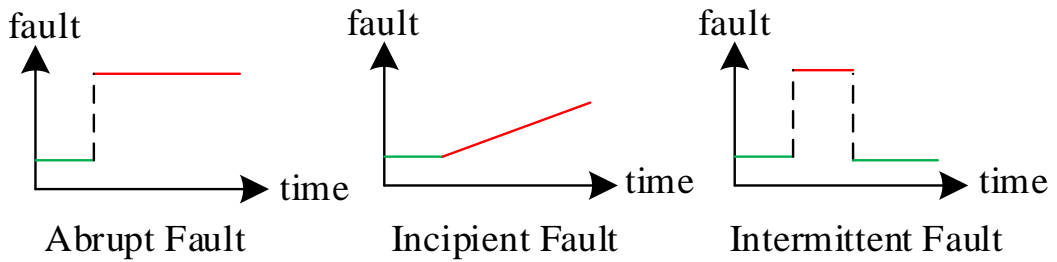
- stop operation หยุดการทำงานเมื่ออยู่ภายใต้การทำงานที่อาจทำให้เกิดอันตราย
- change operation หยุดการขยายความผิดปกติเพื่อไม่ให้กระทบไปยังขั้นตอนการทำงานถัดไป ทำได้โดยการลดการทำงาน ลดความเร็ว ความดัน หรือ อุณหภูมิ
- reconfiguration ทำการกำหนดค่าใหม่โดยระหว่างนั้นใช้เซนเซอร์ตัวอื่นที่ใกล้เคียงกันทำงานแทนส่วนที่กำลังกำหนดค่าใหม่เพื่อให้ระบบยังดำเนินต่อไป
- maintenance ทำการปรับตัวแปรทันทีเพื่อให้สอดคล้องกับการทำงานและเปลี่ยนอุปกรณ์ที่มีการเสื่อมสภาพ
- repair ซ่อมปรับปรุงแก้ไขใหม่ทั้งหมดที่เสียหาย

โดยกระบวนการที่กล่าวมาจะส่งผลทำให้สูญเสียหากไม่ทำการแก้ไขอาจทำให้เกิดการทำงานล้มเหลว (Failure) เป็นเหตุการณ์ที่การทำงานหยุดชะงักโดยไม่สามารถทำตามหน้าที่ได้ มีสาเหตุมาจากมีความผิดปกติเกิดขึ้นหลายจุด การบิดงอ การแตกหรือความล้าที่เกิดขึ้นทำให้ไม่สามารถใช้งานต่อได้ เพื่อหลีกเลี่ยงเหตุการณ์ดังกล่าวจึงมีการศึกษาเกี่ยวกับการตรวจจับการเสียและแยกแยะลักษณะของความผิดปกติของเครื่องจักรมากขึ้น

### 3.2.1 การตรวจจับความผิดปกติ (Fault detection)

เป็นการใช้ความสัมพันธ์ระหว่างตัวแปรที่เกี่ยวข้องกับการเปลี่ยนแปลงที่อาจเกิดมาจากความผิดปกติ[27] [28] [29] รูปแบบความผิดปกติอาจจะอยู่ในรูปแบบทันทีทันใด (abrupt fault) มีลักษณะเป็นการเปลี่ยนแปลงอย่างทันทีทันใด หากเกิดความผิดปกติในรูปแบบนี้ควรจะรีบแก้ไขอย่างทันที รูปแบบเริ่มเกิดขึ้น (Incipient fault) มีลักษณะเป็นการเปลี่ยนแปลงทีละน้อย ทำให้มีเวลาในการปรับตัวเพื่อปรับตัวและรับมือทำให้ผลกระทบจะไม่รุนแรงเท่ารูปแบบก่อนหน้า รูปแบบไม่ต่อเนื่อง (Intermittent fault) มีลักษณะเป็นการเกิดที่ไม่สามารถคาดเดาได้อาจจะเกิดเป็นรูปแบบเป็นรอบ หรือ เกิดแบบไม่เป็นรอบก็ได้ ดังแสดงในรูปที่ 10





รูป 10 รูปแบบของความผิดปกติแต่ละชนิด

โดยความผิดปกติส่วนใหญ่เกิดจากความเปลี่ยนแปลงของตัวแปร ดังสมการที่ 3.1.1

$$Y(t) = Y_u(t) + f(t) \tag{3.1.1}$$

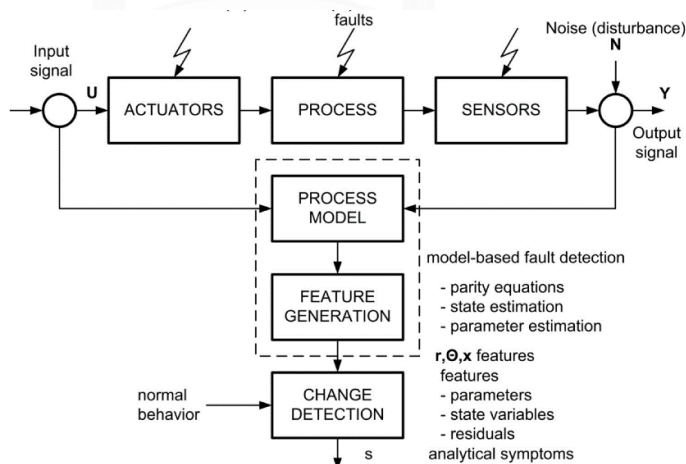
โดย  $Y(t)$  คือ ค่าของตัวแปรที่สนใจ

$f(t)$  ค่าของความเปลี่ยนแปลงที่เพิ่มเข้ามา

โดยการตรวจจับเสียหายสามารถแบ่งออกได้ 3 วิธีการ คือ การตรวจจับตามแบบจำลอง การตรวจจับโดยใช้ฐานความรู้ และการตรวจจับโดยใช้สัญญาณ

การตรวจจับตามแบบจำลอง (Model based)

เป็นตรวจจับความผิดโดยใช้แนวคิดในการวิเคราะห์ซ้ำ โดยจะทำการเปรียบเทียบค่าผลลัพธ์ที่ได้จริงกับค่าผลลัพธ์ที่ได้จากแบบจำลองทางคณิตศาสตร์ [29]แสดงดังรูปที่ 11



รูป 11 ภาพแบบจำลองทางคณิตศาสตร์

เป็นการอธิบายโดยใช้ภาพประกอบเพื่อให้เข้าใจต่อความเข้าใจ จะแสดงถึงระบบในรับค่าเข้ามาแล้วนำไปให้ตัวกระตุ้นให้ทำงานต่อมานำไปเข้าสู่กระบวนการซึ่งมีความผิดปกติเกิดขึ้น เมื่อเซนเซอร์วัดค่าออกมาได้ ค่าที่วัดได้ออกมานั้นอาจจะมีค่าเป็นค่าที่วัดได้จริงที่ได้จากการทำงานที่มีความผิดปกติร่วมกับค่าสัญญาณรบกวนซึ่งอาจจะมาจากสภาพแวดล้อมออกมาเป็นผลลัพธ์ ซึ่งเรานำค่าที่ได้เหล่านั้นมาทำการเข้าแบบจำลองของกระบวนการอีกครั้ง เพื่อทำการวิเคราะห์โดยใช้สมการต่างๆ เพื่อดูถึงสัญญาณความผิดปกติและนำมาปรับใช้ให้สามารถตรวจจับความผิดพลาดได้ในอนาคต โดยแบบจำลองที่เข้าง่ายคือ แบบจำลองโดยใช้สมการความเท่าเทียม (Parity Equation) เป็นแบบจำลองที่มีการตรวจสอบความสอดคล้องของสมการทางคณิตศาสตร์ของระบบกับค่าที่ได้รับ แล้วเอาผลลัพธ์ของสมการมาเปรียบเทียบความ

#### การตรวจจับโดยใช้ฐานความรู้ (Knowledge based)

เป็นวิธีการตรวจจับที่ใช้ผู้มีความรู้เฉพาะด้านมาหาคำแนะนำหรือควบคุมดูแล โดยการตรวจจับความผิดปกติในวิธีการนี้จะใช้การประเมินจากค่าที่ได้รับมานำมาให้ผู้เชี่ยวชาญเฉพาะทางด้านนั้นๆ วิเคราะห์เพื่อวางแผนการป้องกันดูแลต่อไป โดยเมื่อผู้เชี่ยวชาญได้รับข้อมูลมาวิเคราะห์ เพื่อหาค่าลักษณะเฉพาะที่บอกถึงพฤติกรรมที่ผิดปกติสามารถทำได้หลากหลายวิธีเช่น การตั้งค่าขอบเขตเพื่อตรวจหาค่าที่เกินค่าที่ตั้งไว้โดยจะเป็นค่าที่ได้จากผู้เชี่ยวชาญเป็นผู้กำหนด

#### การตรวจจับโดยใช้สัญญาณ (Signal based)

เป็นวิธีการที่ทำโดยใช้ข้อมูลเก่ามาทำการวิเคราะห์ ตัวอย่างของวิธีการนี้ได้แก่ การตรวจสอบขอบเขต ตรวจสอบแนวโน้ม การทำแบบจำลองสัญญาณ เป็นต้น โดยการตรวจสอบขอบเขตมีการกำหนดค่ามากที่สุด ( $Y_{max}$ ) และค่าน้อยสุด ( $Y_{min}$ ) โดยสถานะปกติเพื่อแสดงถึงช่วงที่สามารถเป็นไปได้ หากนอกเหนือจากช่วงจะถือว่าเป็นความผิดปกติ สามารถแบ่งช่วงได้โดย

$$Y_{min} < Y(t) < Y_{max} \quad (3.1.2)$$

โดยใช้การคำนวณแนวโน้มของตัวแปรโดยการหาอนุพันธ์อันดับที่หนึ่ง ดังสมการที่ 3.1.3

$$\dot{Y} = \frac{dY(t)}{dt} \quad (3.1.3)$$

โดยสามารถตรวจสอบสถานะปกติได้โดย

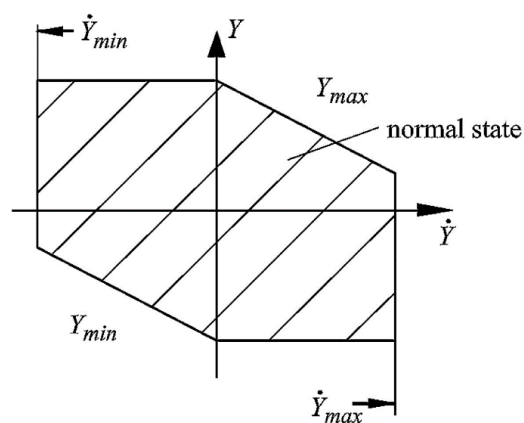
$$\dot{Y}_{min} < \dot{Y}(t) < \dot{Y}_{max} \quad (3.1.4)$$

นำวิธีการก่อนหน้าใช้กับวิธีการก่อนหน้าได้โดยใช้แนวโน้มในการกำหนดค่าขอบเขต ดังสมการ 3.1.5 และ 3.1.6

$$Y_{max} = f(\dot{Y} > 0) \quad (3.1.5)$$

$$Y_{min} = f(\dot{Y} < 0) \quad (3.1.6)$$

เพื่อให้มีความไวในการตรวจจับ และหลีกเลี่ยงการแจ้งเตือนหลอกของแนวโน้มเล็กๆ จึงมีค่าความชันที่ทำให้มีความไวในการตรวจจับที่เร็วขึ้น จะเป็นดังรูปที่ 12



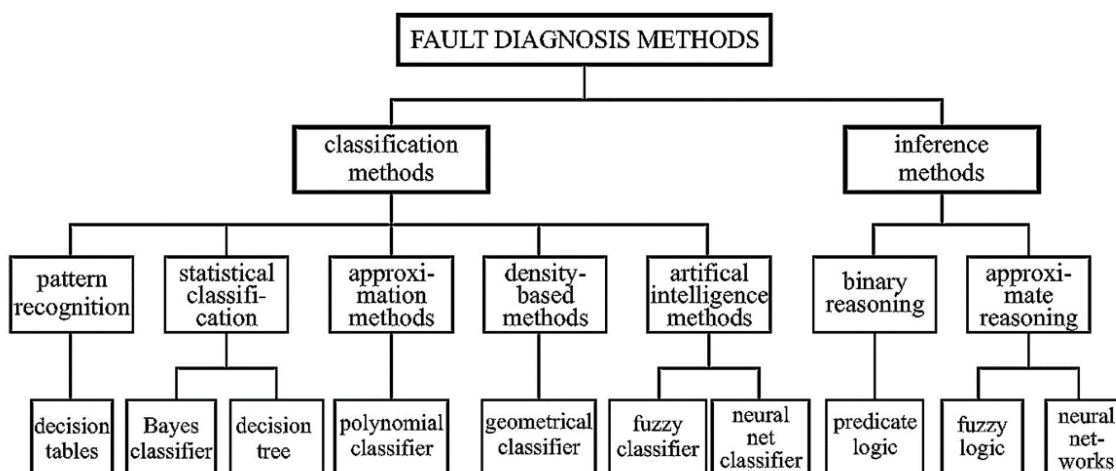
รูป 12 การตรวจสอบโดยใช้ค่าจริงกับใช้แนวโน้มโดยค่า  $Y_{max}$  และ  $Y_{min}$  จะอยู่ในรูปของ  $\dot{Y}$

Pattern recognition เป็นแบบจำลองสัญญาณแบบหนึ่งใช้งานโดยให้แบบจำลองจดจำรูปแบบของข้อมูลที่ถูกใส่เข้ามาเพื่อแสดงผลให้ใกล้เคียงกับค่าที่ใส่เข้ามา หากเป็นการเรียนรู้แบบมีผู้สอนโดยเมื่อได้รับค่าเข้ามาก็ทำการจดจำข้อมูลไว้ในขั้นตอนของการฝึกชุดข้อมูล สถานะของข้อมูลจะถูกเก็บไว้และจะแสดงข้อมูลออกมาตามที่ได้จดจำไว้ในตอนแรก

### 3.2.2 การวินิจฉัยความผิดปกติ (Fault diagnosis)

การวินิจฉัยความผิดปกติมีเป้าหมายเพื่อระบุรายละเอียดของความผิดปกติเท่าที่เป็นไปได้ เช่น ขนาดของความผิดปกติ ตำแหน่งของความผิดปกติ และเวลาที่ตรงจับความผิดปกติได้ โดยกระบวนการดังกล่าวมีพื้นฐานมาจากการวิเคราะห์ การสังเกต รวมถึงการเรียนรู้จากลักษณะอาการ (Symptoms) ต่างๆที่เกิดขึ้น โดยการวินิจฉัยจะเป็นผลจากการสังเกตลักษณะอาการของความผิดปกติแล้วนำมาสรุปเป็นเหตุการณ์ที่ทำให้เกิดและกล่าวได้ว่าเหตุการณ์นี้เป็นลักษณะพฤติกรรมที่ทำให้เกิดความผิดปกติในรูปแบบนั้นๆ

วิธีการวินิจฉัยความผิดปกติแบ่งออกเป็น 2 วิธีหลัก แสดงดังรูปที่ 13 โดยจะแบ่งออกเป็น การแบ่งหมวดหมู่ (Classification methods) และการอนุมาน (Inference methods)



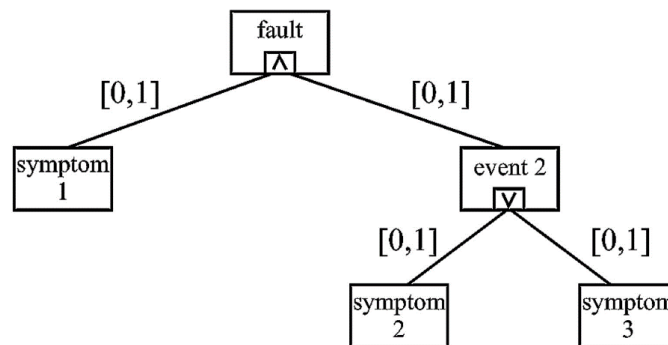
รูป 13 วิธีการที่ใช้การวินิจฉัยความผิดปกติ

การวินิจฉัยความผิดปกติด้วยวิธีการแบ่งหมวดหมู่

เป็นวิธีการวินิจฉัย [30] ที่เข้าใจง่าย เนื่องจากใช้ความสัมพันธ์ระหว่างอาการที่เกิดขึ้น ความผิดปกติในการวินิจฉัยโดยการวินิจฉัยด้วยการจดจำรูปแบบ (Pattern Recognition) จะมีลักษณะอาการต้นแบบ (Reference Symptoms) ให้จำรูปแบบกับความเสียหายที่เกิดขึ้น แล้วนำข้อมูลที่ต้องการวินิจฉัยมาเปรียบเทียบกับรูปแบบที่ได้จดจำไว้เพื่อระบุความผิดปกติที่เกิดจากการจดจำรูปแบบการ

วินิจฉัยความผิดปกติด้วยวิธีการอนุมาน

เป็นการวินิจฉัยที่ใช้ความสัมพันธ์ระหว่างความผิดปกติกับลักษณะอาการเพื่อหาสาเหตุของความผิดปกติ จากความผิดปกติไปยังเหตุการณ์ที่เป็นจุดสังเกตและไปยังลักษณะอาการที่เกิดเป็นการหาต้นเหตุของการเกิดความผิดปกติ ยกตัวอย่างเช่น Fault tree เป็นการใช้รูปแสดงความสัมพันธ์ของความน่าเชื่อถือกับความผิดปกติ โดยจะพิจารณาในระบบเลขฐานสองคือถูกหรือผิด โดยจะบอกถึงอันตรายที่จะเกิดขึ้นหรือความผิดปกติ แล้วไล่ลงมาเป็นเหตุการณ์ก่อนที่จะเกิด ความผิดปกติไปยังสาเหตุของความผิดปกติ โดยจะมีความเชื่อมโยงกันเป็นเหมือนต้นไม้ระหว่างความผิดปกติ เหตุการณ์ และอาการ แสดงดังรูปที่ 14



รูป 14 แผนภาพ Fault tree

แสดงถึงความสัมพันธ์ของเลขฐานสองที่นำไปสู่ความผิดปกติที่จะเกิด โดย fault tree สามารถนำไปประยุกต์ใช้ได้หลายแบบ เช่น การใช้ fault tree ในช่วงการออกแบบเพื่อระบุสถานะวิกฤตของระบบย่อย แล้วนำไปคำนวณหาความน่าจะเป็นที่ระบบจะทำงานล้มเหลว

โดยในงานวิทยานิพนธ์นี้จะใช้การวินิจฉัยแบบแบ่งหมวดหมู่ (Classification method) ในรูปแบบการจดจำรูปแบบ (Pattern recognition) โดยจะใช้การเรียนรู้ของเครื่องจักรช่วยในการจดจำรูปแบบของข้อมูลเพื่อตรวจจับสถานะที่ผิดปกติและระบุชนิดของความผิดปกติที่เกิดขึ้น

### 3.3 การเรียนรู้ของเครื่องจักร (Machine Learning )

การเรียนรู้ของเครื่องจักร คือ ระบบการเรียนรู้ของคอมพิวเตอร์ที่มีความสามารถเรียนรู้ได้ด้วยตนเองโดยใช้ข้อมูลเป็นตัวช่วยในการเรียนรู้ การเรียนรู้ของเครื่องจักรเป็นส่วนหนึ่งของปัญญาประดิษฐ์ ซึ่งจะใช้กับข้อมูลที่มีความซับซ้อนมากมนุษย์ไม่สามารถประมวลผลได้ หลักการทำงานคือป้อนชุดข้อมูลมาให้แบบจำลองศึกษาแล้วทำการจำแนกข้อมูลที่ได้เรียนรู้มา โดยการเรียนรู้นั้นจะใช้กลไกการทำงาน หรือเรียกว่าอัลกอริทึม (Algorithm) ในการวิเคราะห์ข้อมูลซึ่งมีการพัฒนาอยู่ตลอดเวลาเพื่อให้มีความแม่นยำ สะดวกในการใช้งานมากขึ้น ในปัจจุบันมีการใช้การเรียนรู้ของเครื่องอย่างแพร่หลาย เช่น การวิเคราะห์การตลาดเพื่อแนะนำสินค้าในช่วงเวลาที่เหมาะสม การเรียนรู้ของเครื่องจักรแบ่งเป็น 2 ชนิดตามการเรียนรู้ คือ การเรียนแบบมีผู้สอน (Supervised Learning) และการเรียนแบบไม่มีผู้สอน (Unsupervised Learning)

การเรียนแบบมีผู้สอน คือ การเรียนรู้ประเภทหนึ่งที่ต้องมีผู้สอน ดังนั้นจึงมีป้ายกำกับ (Label) เพื่อที่จะระบุชนิดข้อมูลและนำไปเรียนรู้ต่อไป ตัวอย่างการใช้งานการเรียนรู้แบบมีผู้สอน เช่น การจดจำรูปแบบ (pattern Recognition) เป็นการจดจำรูปแบบของข้อมูล อาจจะเป็นรูปแบบการ

เสียหายของเครื่องจักรเพื่อระบุลักษณะความผิดปกติของเครื่องจักรได้ โดยส่วนใหญ่จะนำมาประยุกต์กับการแก้ปัญหา 2 ชนิดคือ การแยกประเภท (Classification) และ การถดถอย (Regression) โดยการแยกประเภทเป็นการทำนายค่าของการเรียนรู้โดยมีลักษณะเป็นการแบ่งข้อมูลออกตามประเภทของข้อมูลนั้น การถดถอยเป็นการทำนายข้อมูลมีลักษณะเป็นแนวโน้มของข้อมูล โดยการหาความสัมพันธ์ระหว่างข้อมูลตัวแปรภายในเพื่อให้ได้แนวโน้มความสัมพันธ์ของตัวแปรนั้นๆ

การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) คือ การเรียนรู้ประเภทหนึ่งที่ไม่จำเป็นต้องมีผู้สอน ไม่จำเป็นต้องมีป้ายกำกับ อาศัยความสัมพันธ์กับระหว่างตัวแปรเพื่อนำมาจัดกลุ่มของข้อมูลแยกแยะชนิดของข้อมูล ตัวอย่างการใช้งาน การเรียนรู้แบบไม่มีผู้สอนเช่น การจัดกลุ่มลูกค้าจากข้อมูลมหาศาล เพื่อที่จะออกแบบนโยบายการขายให้ตรงกลุ่มเป้าหมาย

### 3.4 การเตรียมข้อมูล (Data preprocess)

การวิเคราะห์ข้อมูล[31] รวมถึงการสร้างแบบจำลองสิ่งที่สำคัญที่สุดคือ ชุดข้อมูลที่นำมาใช้ โดยทั่วไปข้อมูลที่ได้จริงส่วนใหญ่จะมีลักษณะไม่พร้อมใช้งาน เช่น ข้อมูลไม่สมบูรณ์ ข้อมูลมีสัญญาณรบกวน หรือ ข้อมูลไม่สอดคล้องกัน หากนำมาใช้จะทำให้ระบบมีการเรียนรู้ที่ผิดหรือความถูกต้องแม่นยำมีค่าน้อยกว่าความเป็นจริง ดังนั้นจึงมีขั้นตอนการเตรียมข้อมูลเกิดขึ้นเพื่อให้ข้อมูลนั้นเหมาะสมกับการใช้งานมากยิ่งขึ้น โดยการเตรียมข้อมูล คือ เทคนิคที่ทำการเพิ่ม ลด หรือ เปลี่ยนข้อมูล โดยขึ้นกับลักษณะของข้อมูลที่ได้รับมาในตอนแรก

ข้อมูลไม่สมบูรณ์ (Missing data) คือ การที่ข้อมูลมีการสูญหาย มีไม่ครบ ขาดไปบางส่วน หรือ ขาดรายละเอียดของข้อมูล ซึ่งข้อมูลที่สูญหายจะส่งผลให้แบบจำลองมีปัญหา ดังนั้นควรจะจัดการกับข้อมูลที่สูญหายก่อนนำไปใช้งาน สามารถเกิดได้จากหลายสาเหตุ เช่น อุปกรณ์รับค่ามีปัญหา สัญญาณขาดหายทำให้ไม่ได้รับข้อมูล สามารถส่งผลให้การวิเคราะห์ข้อมูลไม่ถูกต้องเท่าที่ควร ข้อมูลสูญหายมีหลายลักษณะเช่น ข้อมูลหายแบบสุ่มสมบูรณ์ หมายถึง ข้อมูลสูญหายไปอย่างสุ่มไม่ขึ้นกับค่าใดค่าหนึ่ง และข้อมูลที่หายเป็นอิสระต่อกัน, ข้อมูลหายแบบสุ่ม หมายถึง ข้อมูลสูญหายไปอย่างสุ่มภายในบางส่วนหรือเฉพาะบางกลุ่มตัวแปร และ ข้อมูลหายแบบไม่สุ่ม หมายถึง ข้อมูลสูญหายไปขึ้นกับค่าของข้อมูลในกลุ่มหรือตัวแปรเดียวกันหรือตัวแปรอื่น โดยการจัดการปัญหาดังกล่าวที่ถูกต้องและดีที่สุดคือ การนำข้อมูลที่แท้จริงมาใส่ แต่ในความเป็นจริงมีความเป็นไปได้น้อยที่จะสามารถทำได้ เพื่อให้การวิเคราะห์ทำได้ใกล้เคียงและแม่นยำ จึงมีการจัดการข้อมูลดังนี้ การลบทิ้ง

เป็นการลบข้อมูลที่ไม่ครบถ้วนหรือไม่สมบูรณ์ออก การแทนที่ด้วยข้อมูลตัวหนึ่ง เป็นการหาข้อมูลมาเติมเช่น ค่าทางสถิติ โดยทำได้ง่ายแต่ข้อมูลอาจจะไม่ตรงกับความจริง

ค่าผิดปกติ (Outlier) เป็นค่าที่มีค่าแตกต่างไปจากกลุ่มข้อมูลส่วนใหญ่อย่างผิดปกติ คือ มีค่าโดดเด่นข้อมูลส่วนใหญ่ผิดปกติ หากนำข้อมูลไปใช้อาจทำให้ข้อมูลวิเคราะห์ข้อมูลคลาดเคลื่อนได้ จึงจำเป็นต้องตรวจสอบว่าชุดข้อมูลที่นำไปใช้มีข้อมูลที่มีค่าผิดปกติอยู่หรือไม่และจัดการค่าผิดปกติออกจากชุดข้อมูล โดยสามารถจัดการค่าได้โดยใช้วิธีการทางสถิติเช่น IQR หรือ Interquatile Range เป็นเทคนิคที่จัดการค่าที่ผิดปกติโดยจะกำหนดค่าบน ค่าล่าง คิดมาจากค่าที่ 25<sup>th</sup> 75<sup>th</sup> ของเปอร์เซ็นต์ไทด์ หากข้อมูลมีค่าอยู่นอกช่วงที่กำหนดถือว่าเป็นค่าผิดปกติและถูกกำจัดออกไป โดยค่าบน ค่าล่างสามารถหาได้จาก

ค่าบน (Upper)

$$\text{Upper} = Q_3 + 1.5 * \text{IQR} \quad (3.1)$$

ค่าล่าง (Lower)

$$\text{Lower} = Q_1 - 1.5 * \text{IQR} \quad (3.2)$$

IQR

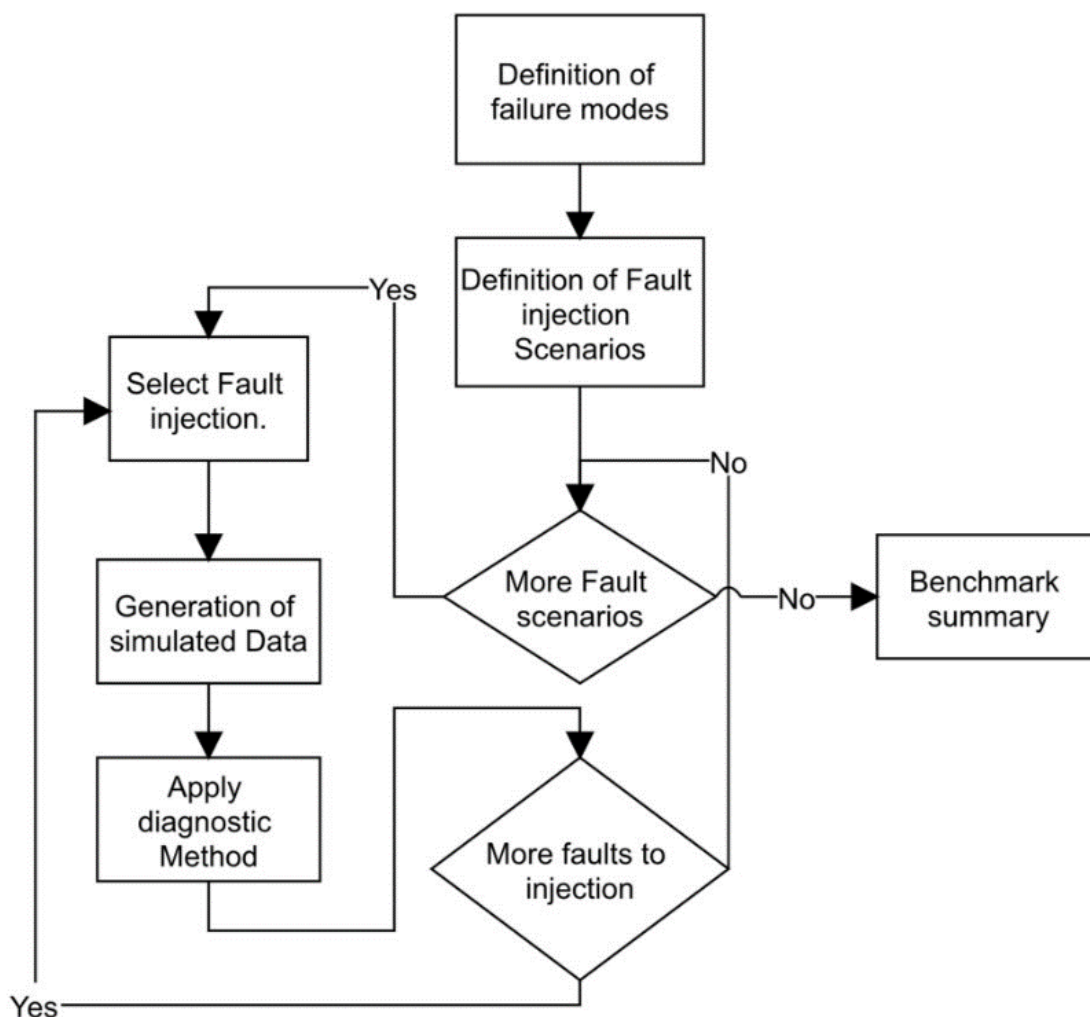
$$\text{IQR} = Q_3 - Q_1 \quad (3.3)$$

โดย  $Q_3$  คือ ค่าที่ตำแหน่งเปอร์เซ็นต์ไทด์ที่ 75

$Q_1$  คือ ค่าที่ตำแหน่งเปอร์เซ็นต์ไทด์ที่ 25

การเพิ่มความผิดปกติให้แบบจำลองเรียนรู้โดยใช้ข้อมูลสังเคราะห์ (Synthetic fault injection) ในการเรียนรู้แบบมีผู้สอนจำเป็นต้องมีตัวกำกับเพื่อบอกสถานะของข้อมูล เนื่องจากข้อมูลจริงไม่มีข้อมูลผิดปกติ จึงนำข้อมูลสังเคราะห์มาใช้แทนข้อมูลผิดปกติ เพื่อจำลองเหตุการณ์ ความผิดปกติที่ต้องการให้แบบจำลองเรียนรู้ โดยจะมีแผนการทำงานแสดงดังรูปที่ 15 ทำได้โดยระบุรูปแบบความเสียหายที่ต้องการจำลองและตัวแปรที่มีผลให้เกิดความเสียหายรูปแบบนั้นๆ หลังจากนั้นปรับแต่งข้อมูลให้มีลักษณะคล้ายการทำงานผิดปกติตามที่กำหนด โดยความผิดปกติที่ให้แบบจำลองเรียนรู้จะเป็นความผิดปกติในรูปแบบของความผิดปกติแบบค่อยๆเป็นค่อยไป (Gradual fault) ที่

อ้างอิงมาจากพฤติกรรมการเสื่อมของเครื่องจักร ที่จะทำให้ค่ามีค่าปรับตัวไปในทิศทางใดทางหนึ่ง โดยในการจำลองจะเพิ่มค่าความเสียหายให้กับชุดข้อมูลที่ 50% ของข้อมูล โดยจะเป็นการเพิ่มความเสียหายขึ้นไปอย่างต่อเนื่อง โดยเพิ่มความเสียหายจาก 0 ไปถึง 3 เท่าของส่วนเบี่ยงเบนมาตรฐาน



รูป 15 การสร้างข้อมูลสังเคราะห์

(3SD) โดยอ้างอิงจาก Drift detection method (DDM) เป็นการสังเกตพฤติกรรมของข้อมูล โดยเมื่อค่าความน่าจะเป็นที่จะคัดแยกข้อมูลผิดมีมากขึ้นเมื่อเทียบกับค่าส่วนเบี่ยงเบนมาตรฐาน DDM จะตรวจสอบเงื่อนไขในการตรวจจับแล้วแจ้งเตือนสถานะผิดปกติ กับ แจ้งเตือนถึง สถานการณ์เปลี่ยนข้อมูล (Drift) การแจ้งเตือนสถานะผิดปกติสามารถตรวจจับโดยใช้กฎสามซิกมา (three-sigma rule) เป็นการมีส่วนเบี่ยงเบนมาตรฐานเป็นตัวกำหนดสถานะ หากข้อมูลเกินค่า 3SD จะถือว่าข้อมูลนั้นเป็นข้อมูลผิดปกติ



### 3.5 การเรียนรู้แบบหมู่

การเรียนรู้แบบหมู่ (Ensemble learning) [30] [32] เป็นแบบจำลองที่เพิ่มประสิทธิภาพการทำงานโดยการนำหลายๆแบบจำลองเข้ามาทำงานรวมกันแล้วเพื่อให้ได้ผลลัพธ์ที่ดีกว่าการวิเคราะห์โดยใช้แบบจำลองเพียงตัวเดียว เนื่องจากแต่ละแบบจำลองมีตัวเรียนรู้ (Learner) ที่ต่างกัน ทำให้มีการปรับแต่งตัวแปร การใช้อัลกอริทึม รวมไปถึง ปัญหาการทำงานที่แตกต่างกัน แต่ละแบบจำลองจะมีคำตอบเป็นของตนเองทำให้มีคำตอบที่หลากหลาย Ensemble จึงมีการเลือกผลลัพธ์ที่เป็นเสียงส่วนมากเพื่อเป็นคำตอบ เพื่อลดความขัดแย้งและเพิ่มความแม่นยำ หากยังมีหลายตัวเรียนรู้อีกเท่าไรผลลัพธ์ที่ได้ก็จะมีแนวโน้มเชื่อถือได้โดยการใช้หลายแบบจำลองมาประมวลผลแล้วเอามาทำการโหวตเสียงส่วนมาก (Majority Voting) โดยจะนำผลลัพธ์ของหลายแบบจำลองมาโหวตเสียงข้างมากหากมีผลรวมเกิน 50% จะถือว่าเป็นผลลัพธ์ที่ดีที่สุด สามารถเขียนเป็นสมการได้ดังนี้

$$\sum_{t=1}^T d_{t,c^*} = \max_c \sum_{t=1}^T d_{t,c} \quad (3.4)$$

โดย  $T$  คือ จำนวนของแบบจำลอง โดยที่  $t = 1, \dots, T$

$C$  คือ จำนวนของผลลัพธ์ หรือ Label โดยที่  $c = 1, \dots, C$

$d_{t,c}$  คือ ผลลัพธ์ของแบบจำลอง  $t$

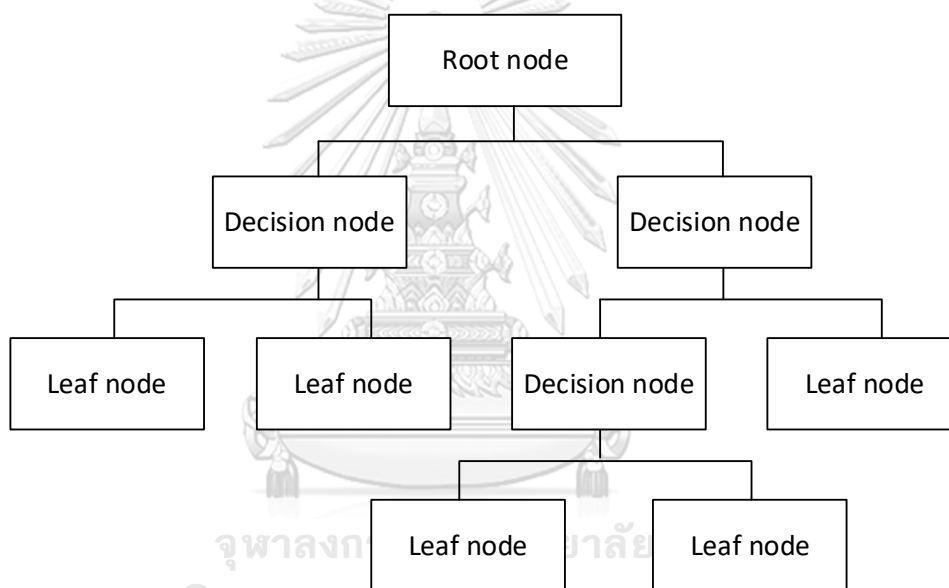
การใช้เพียงแบบจำลองเดียวอาจทำให้เกิดปัญหาหรือมีข้อจำกัดตามมา เช่น

- ความแม่นยำที่ต่ำ เนื่องจากแต่ละแบบจำลองมีข้อจำกัดที่ต่างทำให้ การใช้แบบจำลองหลายตัวมาช่วยวิเคราะห์จะช่วยลดผลของข้อจำกัดนั้นลงไป
- ความผิดพลาดที่มาก การใช้เพียงแบบจำลองเดียว การวิเคราะห์บางครั้งอาจผิดได้ ถ้ามีหลายแบบจำลองก็จะช่วยกันวิเคราะห์ข้อมูลได้ดีกว่า เป็นการช่วยลดความผิดพลาดได้

ซึ่งแบบจำลองที่นำมาใช้สามารถใช้ได้หลากหลายวิธีการของการเรียนรู้ของเครื่อง เช่น Linear regression, Naïve Bay, K-Nearest Neighbors และอื่นๆ แต่ที่นิยมเอามาใช้คือ ต้นไม้ตัดสินใจ (Decision tree) โดยนำมาประยุกต์กับเทคนิคของการเรียนรู้แบบหมู่มีอยู่ 3 เทคนิค คือ Bagging, Boosting และ Stacking โดยในที่นี้จะกล่าวถึงเพียง Bagging และ Boosting เท่านั้น

### ต้นไม้ตัดสินใจ

เป็นการเรียนรู้ที่ใช้การแทนความรู้ให้อยู่ในรูปต้นไม้ตัดสินใจ ใช้สำหรับการจำแนกประเภทของข้อมูล โดยแบ่งออกเป็นกลุ่มต่างๆ โดยใช้คุณลักษณะของข้อมูลที่ต่างกันในการจำแนก ต้นไม้ตัดสินใจ ประกอบไปด้วยบัพ (node) และกิ่ง (link or branch) ที่ต่อจากบัพ บัพที่ปลายสุดจะเรียกว่า บัพใบ (leaf node) มีวิธีคิดจากด้านบนไปยังด้านล่างโดยจะแตกกิ่งก้านสาขาเป็นเหมือนต้นไม้ [33] แสดงดังรูปที่ 16 ซึ่งแต่ละกิ่งจะแสดงถึงทางเลือกที่ต่างออกไป ในส่วนของใบแต่ละใบจะแสดงถึงผลลัพธ์การตัดสินใจ โดยต้นไม้ตัดสินใจจะมีจุดเริ่มต้นของต้นไม้ตัดสินใจจะเรียกว่า บัพราก (root node) หลังจากนั้นจะสร้างบัพที่ละบัพเพื่อตรวจสอบคุณลักษณะและแยกย่อยลงไปตามค่าของกิ่ง



รูป 16 ต้นไม้ตัดสินใจ

โดยในการสร้างบัพข้อมูลจะถูกสกัดโดยการวัดค่าข้อมูลที่ได้รับ (Information gain) เพื่อที่จะนำมา คัดแยกคุณลักษณะไปยังแต่ละบัพของต้นไม้ โดยค่า Gain [34] ใช้ในการหาคุณลักษณะที่ดีที่สุดที่สามารถ คำนวณได้จากค่า Entropy ดังสมการที่ 3.5

$$E(S) = - \sum_{i=1}^n p_i \log_2 p_i \quad (3.5)$$

โดยที่  $E(S)$  คือ ค่า Entropy ของเซตข้อมูลทั้งหมด ,

$P_i$  คือ ค่าความน่าจะเป็นของข้อมูลที่สนใจ

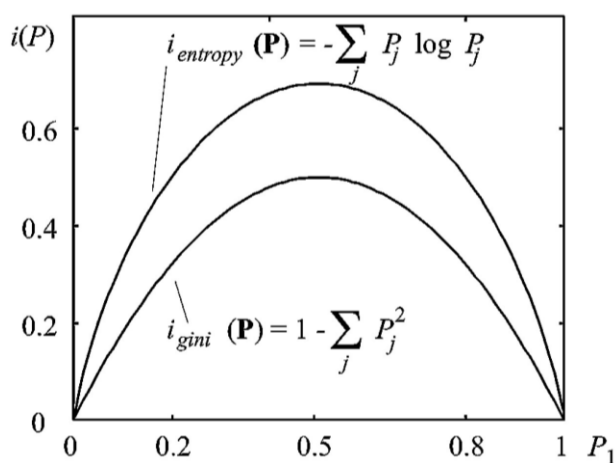
นอกจากใช้ฟังก์ชัน Entropy แล้วยังสามารถใช้ฟังก์ชัน Gini ซึ่งสามารถคำนวณได้ดังสมการที่ 3.6

$$G(S) = 1 - \sum_i p_i^2 \quad (3.6)$$

โดยที่  $G(S)$  คือ ค่า Gini impurity ของเซตข้อมูลทั้งหมด

$P_i$  คือ ค่าความน่าจะเป็นของข้อมูลที่สนใจ

การคำนวณโดยใช้ Gini ใช้เวลาน้อยกว่าเนื่องจากมีขนาดข้อมูลที่น้อยกว่าเมื่อเทียบกับ Entropy หากนำมาวาดกราฟเพื่อเปรียบเทียบค่าในการแยกแยะชุดข้อมูล 2 ชนิด แสดงดังรูปที่ 17



รูป 17 ค่าฟังก์ชัน Entropy และ Gini สำหรับข้อมูล 2 รูปแบบ เมื่อได้ค่า Entropy แล้วจะนำไปแทนใน สมการที่ 3.6 เพื่อหาค่า Gain

$$Gain(A) = E(S) - \sum_{i=1}^n \frac{|S_i|}{S} E(S_i) \quad (3.6)$$

โดยที่  $S_i = P(V_1), P(V_2), P(V_3), \dots, P(V_n)$  เป็นเซตของข้อมูลทั้งหมด

$n$  คือ จำนวนของคุณลักษณะ  $A$

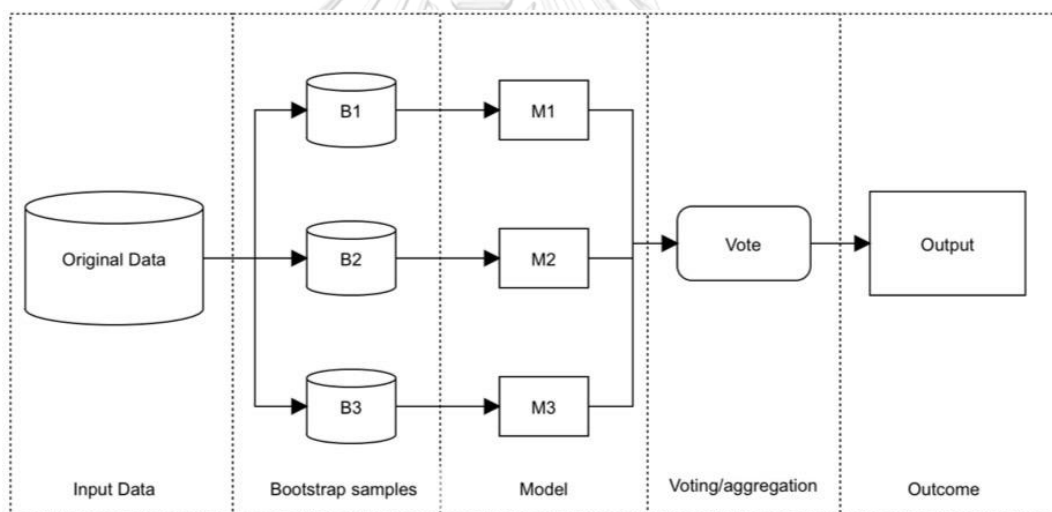
$|S_i|$  คือ จำนวนของข้อมูลที่แบ่งเป็นส่วนของ  $S_i$

$|S|$  คือ จำนวนทั้งหมดของ  $S$

โดยค่า  $Gain$  ที่ได้รับจะบ่งบอกความสามารถในการแยกแยะข้อมูล โดยทดสอบเลือกค่าจากแต่ละคุณลักษณะที่มีค่ามากที่สุดเพื่อที่จะทำหน้าที่เป็นบัพราค

### - เทคนิค Bagging

เทคนิค Bagging หรือ Bootstrap Aggregating คือ เป็นเทคนิคที่สร้างชุดข้อมูลใหม่ที่มีขนาดเท่ากับชุดข้อมูลเดิม แล้วสุ่มตัวอย่างจากข้อมูลเดิมมาใส่ในชุดข้อมูลที่สร้างขึ้นมาใหม่ แสดงดังรูปที่ 18 โดยการสุ่มเป็นการสุ่มแบบแทนที่ไม่สนลำดับ ชุดข้อมูลที่ได้จากการสุ่มนี้จะเรียกว่า bootstrap sample นำข้อมูลชุดใหม่มาทำการเรียนรู้แบบขนาน เนื่องจากการสุ่มข้อมูลทำให้ช่วยลดปัญหาพอดีเกินที่ทำให้แบบจำลองจดจำข้อมูลแทนการเรียนรู้ ซึ่งเป็นสิ่งที่ไม่ต้องการให้เกิดขึ้นกับแบบจำลองแล้วนำผลลัพธ์ทั้งหมดมาเฉลี่ยหรือโหวตเสียงส่วนมากแล้วแต่ประเภทของข้อมูลเพื่อให้ได้ผลลัพธ์สุดท้ายที่มีความแม่นยำมากกว่าการใช้วิธีการเดี่ยววิเคราะห์ข้อมูล ตัวอย่างวิธีการของเทคนิคนี้คือ Random forest เป็นอัลกอริทึมที่มีพื้นฐานจากต้นไม้ตัดสินใจ เป็นการนำต้นไม้ตัดสินใจหลายๆต้นมาทำการวิเคราะห์ และทำการโหวตค่าผลลัพธ์ทำให้มีความแม่นยำที่มากกว่า เมื่อเทียบกับต้นไม้ตัดสินใจต้นเดียว

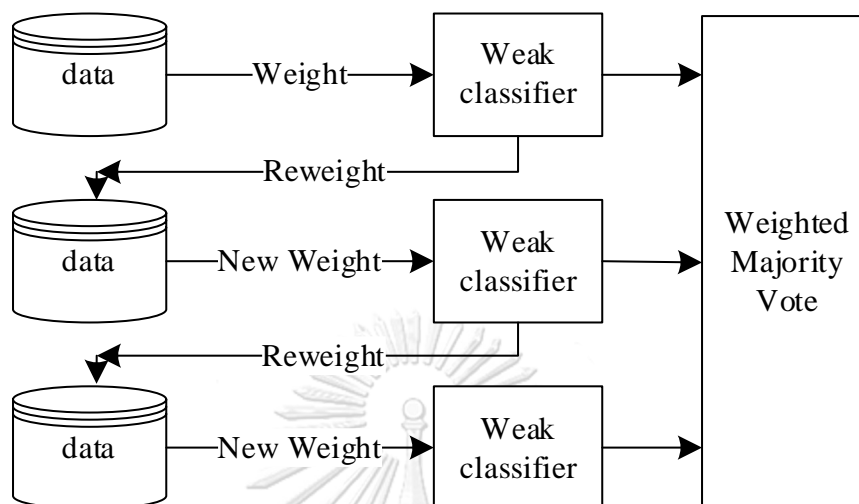


รูป 18 การทำงานของ Bagging

### - เทคนิค Boosting

เป็นเทคนิคที่นำตัวแบ่งประเภทที่มีความแม่นยำน้อย (Weak Classifier) มาทำนายข้อมูลจากนั้นจะเอาผลลัพธ์มาแก้ไขค่าน้ำหนัก (Weight) ค่าน้ำหนักใหม่จะถูกใช้กับตัวแบ่งแยกถัดไปทำให้ผลลัพธ์ที่ได้มีความแม่นยำมากกว่าครั้งก่อนหน้า โดยจะทำการแบบจำลองได้ผลลัพธ์ที่พอใจ เป็นการเปลี่ยนจากตัวแบ่งประเภทความแม่นยำน้อย เป็นตัวแบ่งประเภทความแม่นยำสูง (Strong Classifier) โดย

นำชุดข้อมูลมาทำการเรียนรู้ซ้ำๆ แล้วปรับค่าน้ำหนักโดยขึ้นกับผลลัพธ์ก่อนหน้า ทำให้ค่าน้ำหนักถัดไปมีความแม่นยำมากขึ้น รูปแผนการทำงานแสดงดังรูปที่ 19



รูป 19 การทำงานของ Boosting

และสามารถคำนวณผลลัพธ์การทำนายดังสมการที่ 3.5

$$H(x) = \text{sign} \left( \sum_{m=1}^M \alpha_m H_m(x) \right) \quad (3.5)$$

โดย  $w_m$  คือ ชุดของค่าน้ำหนัก โดยที่  $m = 1, \dots, M$  โดยชุดของค่าน้ำหนักอย่างง่าย จะมีค่าน้ำหนักเท่ากันทุกตัว

$H_m$  คือ ผลลัพธ์จากการทำนาย

โดยวิธีการ Boosting เป็นวิธีการที่สร้างโดยการนำชุดข้อมูลเรียนรู้ซ้ำๆ ในตัวแบ่งประเภทตัวเดิม โดยเริ่มแรกค่าน้ำหนักแต่ละตัวแบ่งประเภทจะมีค่าน้ำเท่ากันตามสมการที่ 3.6

$$w_i = \frac{1}{n} \quad \text{โดย } i \in \{1, \dots, N\} \quad (3.6)$$

โดย  $N$  คือ จำนวนของข้อมูล

และเมื่อนำมาเทรนอีกครั้งจะมีการเปลี่ยนค่าน้ำหนักโดยคิดจากค่าผิดพลาดที่เกิดขึ้นจากผลลัพธ์ก่อนหน้า ดังสมการที่ 3.7

$$\text{err}_m = \sum_{i=1}^N w_i^{(m)} h(-y_i H_m(x_i)) \quad (3.7)$$

โดย  $x_i$  คือ ส่วนประกอบของชุดข้อมูลแกน  $x$  เช่น ความดัน อุณหภูมิ  
 $y_i$  คือ ส่วนประกอบของชุดข้อมูลแกน  $y$  เช่น label ระบุสถานะ  
 $H_m$  คือ ค่าผลลัพธ์ที่ได้จากการเทรนครั้งก่อนหน้า

ทำการปรับน้ำหนักของชุดข้อมูลใหม่โดยทำให้ดีขึ้นกว่าชุดข้อมูลเก่า โดยการเปลี่ยนค่าน้ำหนักจะขึ้นกับความแม่นยำของชุดข้อมูลที่ทำก่อนหน้า ดังสมการที่ 3.8 และ 3.9

$$\alpha_m = \frac{1}{2} \log \left( \frac{1 - \text{err}_m}{\text{err}_m} \right) \quad (3.8)$$

จะได้ค่าน้ำหนักใหม่เป็น สมการ 3.9

$$v_i^{(m)} = w_i^{(m)} \exp(-\alpha_m y_i H_m(x_i)) \quad (3.9)$$

ทำการปรับแต่งให้เหมาะก่อนนำไปใช้โดยการ Renormalize น้ำหนัก

$$S_m = \sum_{j=1}^N v_j \quad \text{โดย } i=1, \dots, N \quad (3.10)$$

จะได้ น้ำหนักที่จะใช้ในครั้งถัดไปเป็นดังสมการที่ 3.11

$$w_i^{(m+1)} = \frac{v_i^{(m)}}{S_m} \quad (3.11)$$

หรือ

$$w_i^{(m+1)} = \frac{w_i^{(m)} \exp(-\alpha_m y_i H_m(x_i))}{S_m}$$

โดยจะได้ผลลัพธ์สุดท้ายจะเป็นการคำนวณค่าน้ำหนัก ดังสมการที่ 3.5

### 3.6 การประเมินแบบจำลอง

การประเมินแบบจำลองเป็นสิ่งที่จำเป็นเนื่องจากเป็นขั้นตอนที่ทำให้รู้ว่าแบบจำลองที่ใช้ในการเรียนรู้มีความแม่นยำ ความถูกต้อง เท่าใด โดยในที่นี้จะพูดถึง Confusion Matrix , F1-score , Accuracy

Confusion Matrix เป็นการประเมินแบบจำลองโดยใช้เมทริกซ์โดยจะแบ่งเป็นการแสดงค่าระหว่างค่าจริงกับค่าที่คาดเดาขึ้นมาดัง แสดงในรูปที่ 20 เพื่อให้ทราบว่ามีค่าที่ถูก-ผิดเท่าใด โดยจะแบ่งออกเป็น

	ข้อมูลจริง บวก(1)	ข้อมูลจริง ลบ(0)
ข้อมูลทำนาย บวก(1)	True positives	False positives
ข้อมูลทำนาย ลบ(0)	False negatives	True negatives

รูป 20 ตารางเมทริกซ์ความสับสน

- TP (True Positive) คือ ข้อมูลมีค่าเป็น จริง ระบบทำนายว่า จริง
- TN (True Negative) คือ ข้อมูลมีค่าเป็น จริง ระบบทำนายว่า ไม่จริง
- FP (False Positive) คือ ข้อมูลมีค่าเป็น ไม่จริง ระบบทำนายว่า ไม่จริง
- FN (False Negative) คือ ข้อมูลมีค่าเป็น ไม่จริง ระบบทำนายว่า ไม่จริง

โดยค่าดังกล่าวจะถูกนำไปใช้ในการคำนวณ

- ความเที่ยงตรง (Precision) เป็นการวัดความแม่นยำของค่าทำนายโดยทำการวัดที่ละตัวกำกับ โดยคิดจากสัดส่วนของจำนวนการทำนายว่าถูกต้อง (True positive) เทียบกับผลรวมของจำนวนการทำนายว่าถูกต้อง และการทำนายว่าไม่ถูกต้อง (False positive) ดังสมการที่ 3.5.1

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3.12)$$

- การเรียกคืน (Recall) เป็นการวัดความถูกต้องของแบบจำลองจากค่าจริง โดยคิดจากสัดส่วนของจำนวนการทำนายว่าถูกต้อง เทียบกับผลรวมของจำนวนการทำนายว่าถูกต้อง และการทำนายถูกว่าผิด (False negative) ดังสมการที่ 3.5.2

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3.13)$$

- ความแม่นยำ (Accuracy) เป็นการวัดความของแบบจำลองโดยเป็นการคิดกับทุกค่า โดยคิดจากสัดส่วนของผลรวมจำนวนการทำนายว่าถูกต้องและการทำนายถูกเป็นผิด (True negative) เทียบกับผลรวมของจำนวนทั้งหมด ดังสมการที่ 3.5.3

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3.14)$$

- ความถูกต้องของ F1 (F1-score) เป็นการวัดความถูกต้องของแบบจำลองโดยคำนวณจากความแม่นยำและการเรียกคืน โดยจะนำทั้งสองหาค่าเฉลี่ยฮาร์โมนิกที่เป็นความจริงของสองค่า ซึ่งคำนวณได้ดังสมการที่ 3.5.4

$$\text{F1} = \frac{2 \cdot (\text{precision} \cdot \text{recall})}{\text{precision} + \text{recall}} \quad (3.15)$$

เมื่อเราทำการประเมินแบบจำลองแล้ว จะทำให้ทราบว่าแบบจำลองนี้พร้อมนำไปใช้งานจริงหรือไม่ เพราะจะเป็นขั้นตอนที่บอกถึงความแม่นยำเมื่อนำไปใช้งาน หรือ เพื่อนำไปเปรียบเทียบกับแบบจำลองอื่นและนำไปปรับปรุงแก้ไขเพื่อให้มีความแม่นยำเหมาะกับการใช้งานมากขึ้น



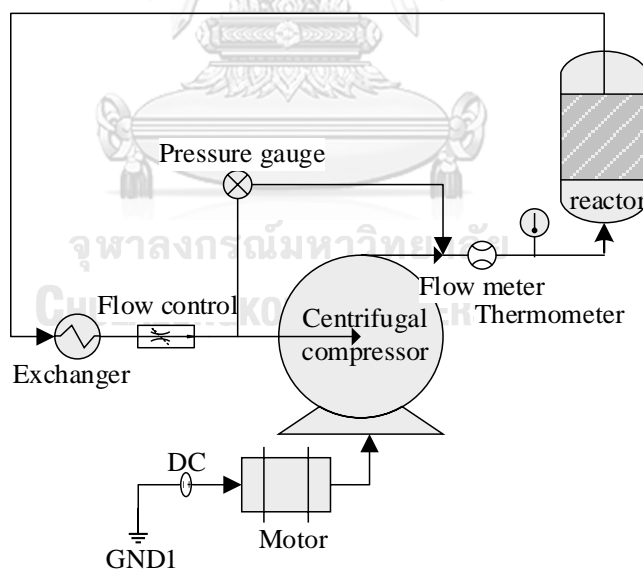
## บทที่ 4

### ผลการทดลอง

ในบทนี้จะกล่าวถึงการทดลองเพื่อให้ผลลัพธ์ในการตรวจจับความผิดปกติของเครื่องอัดอากาศแบบหมุนเหวี่ยง โดยจะอธิบายถึงลักษณะของข้อมูลที่ได้รับ แบบจำลองที่ใช้ในการหาผลลัพธ์ โดยมีหลายวิธีการรวมเข้าด้วยกัน ตัวแปรที่มีผลกับแบบจำลองแต่ละแบบ รวมไปถึงการประเมินผลลัพธ์ที่ได้จากแบบจำลอง

#### 4.1 เครื่องอัดอากาศแบบหมุนเหวี่ยง

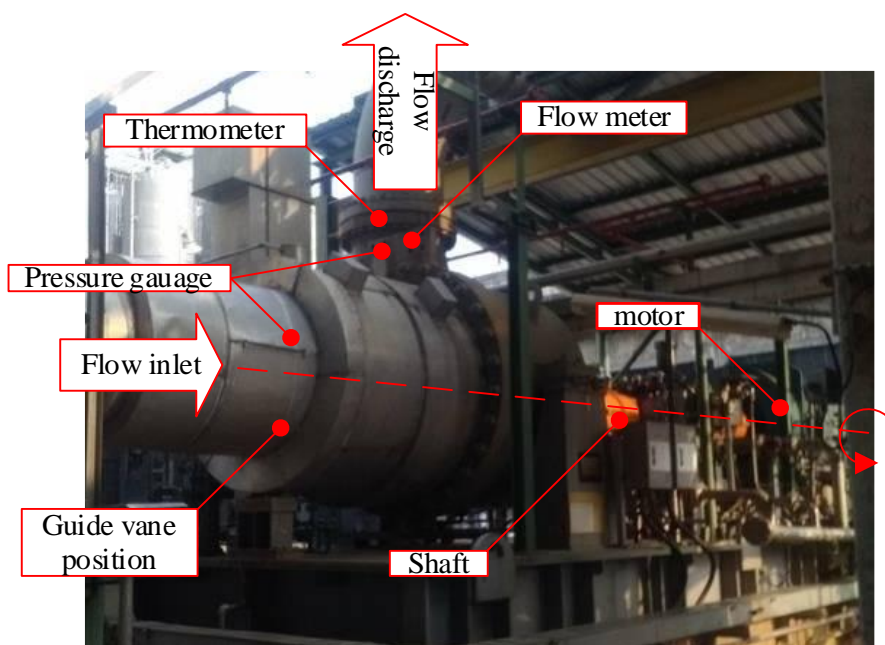
ระบบที่ทำการศึกษาเป็นกระบวนการ Polymer Flashing แสดงดังรูปที่ 21 เป็นกระบวนการระเหยบางส่วนเพื่อแยกสารออกจากกันในการเกิดกระบวนการดังกล่าวแก๊สที่ถูกวาบ (flashed) นั้นมีหลายปัจจัยที่ทำให้เกิดปฏิกิริยาดังกล่าว โดยทั่วไปการวาบไฟของสารประกอบไฮโดรคาร์บอนจะมีค่าพลังงานที่ต่างกันตามความเข้มข้นของสารประกอบ โดยปัจจัยสำคัญที่ทำให้เกิดการวาบไฟคือ อุณหภูมิและความดัน



รูป 21 กระบวนการ Polymer Flashing

โดยเครื่องอัดอากาศที่นำมาศึกษาเป็นเครื่องอัดอากาศแบบหมุนเหวี่ยงจากบริษัท Ebara รุ่น 200TC แสดงดังรูปที่ 22 ถูกติดตั้งเนื่องจากสามารถเพิ่มปริมาตรการไหล และความดันที่เข้าไปในเครื่องปฏิกรณ์ โดยหน้าที่หลักของเครื่องอัดอากาศคือขนส่งสารประกอบไฮโดรคาร์บอนโดยส่วนใหญ่จะเป็น โพรพิลีน แก๊สที่ได้รับจะมากจากยอดของเครื่องปฏิกรณ์ซึ่งเป็นแก๊สที่มีอุณหภูมิสูงตั้งนั้นก่อนนำเข้าสู่

เครื่องอัดอากาศจำเป็นต้องลดอุณหภูมิก่อน โดยจะผ่านเครื่องแลกเปลี่ยนความร้อนเพื่อลดอุณหภูมิลงก่อนเข้า เพื่อป้องกันการเกิดความเสียหายต่อเครื่องอัดอากาศ หลังจากนั้นแก๊สออกจากเครื่องอัดอากาศแล้วถูกส่งไปยังด้านล่างของเครื่องปฏิกรณ์อีกครั้ง โดยแก๊สที่ถูกส่งเข้าไป จะมีอุณหภูมิและอุณหภูมิที่สูงเพื่อช่วยเร่งปฏิกิริยาเคมีภายในเครื่องปฏิกรณ์



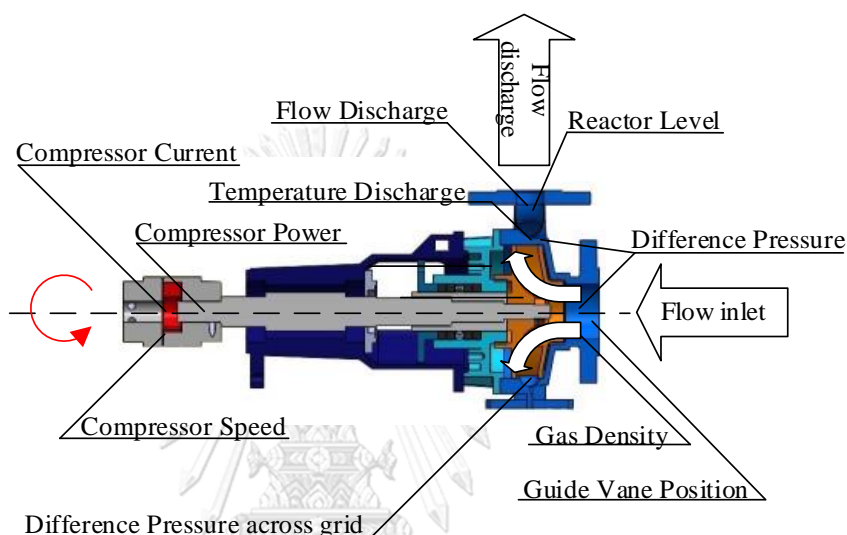
รูป 22 เครื่องอัดอากาศแบบหมุนเหวี่ยงยี่ห้อ Ebara model: 200TC

#### 4.2 ข้อมูลที่รวบรวม

ในการวิเคราะห์การทำงานของเครื่องอัดอากาศโดยใช้การเรียนรู้ของเครื่อง สิ่งสำคัญคือข้อมูลที่รวบรวมจนกลายเป็นข้อมูลมหัต (Big Data) โดยยิ่งข้อมูลมีความหลากหลาย ครอบคลุมก็จะส่งผลให้การเรียนรู้ของเครื่องมีประสิทธิภาพมากขึ้น ในการรวบรวมข้อมูลทำได้ติดตั้งเซนเซอร์ตามตำแหน่งต่างๆบนเครื่องอัดอากาศแบบหมุนเหวี่ยง แสดงดังรูป 23 เพื่อที่จะรวบรวมข้อมูลและนำมาวิเคราะห์โดยข้อมูลที่บันทึกจากเซนเซอร์มีดังนี้

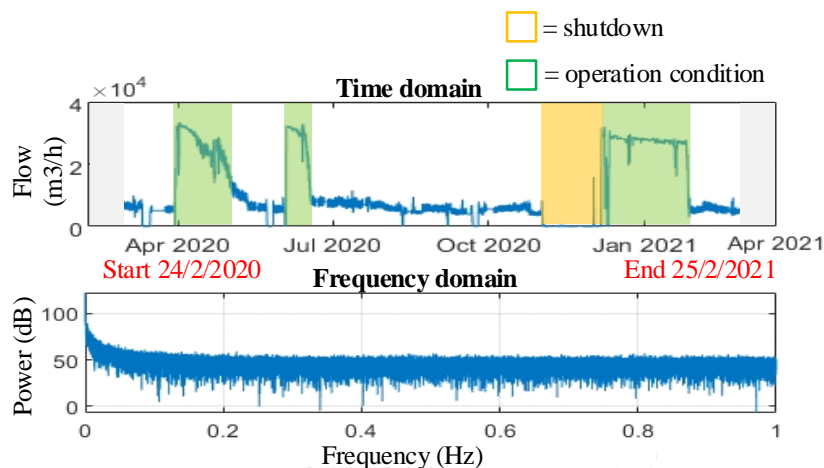
- กำลังไฟฟ้าของมอเตอร์ (กิโลวัตต์)
- กระแสไฟฟ้าของมอเตอร์ (แอมป์)
- ความเร็วรอบของมอเตอร์ (รอบ/นาที)
- ความหนาแน่นของแก๊ส (กิโลกรัม/ลูกบาศก์เมตร)
- ตัวควบคุมแก๊สที่ทางเข้า (%)

- เครื่องมือวัดอุณหภูมิที่ทางออก (องศาเซลเซียส)
- เครื่องมือวัดอัตราการไหลที่ทางออก (ลูกบาศก์เมตร/ชั่วโมง)
- ระดับสารในเครื่องปฏิกรณ์ (%)
- ผลต่างความดันระหว่างใบพัด (มิลลิเมตรน้ำ)
- ผลต่างความดันระหว่างทางเข้าและออก (กิโลกรัม/ลูกบาศก์เซนติเมตร)



รูป 23 ตำแหน่งเซนเซอร์ที่ติดตั้งบนเครื่องอัดอากาศ

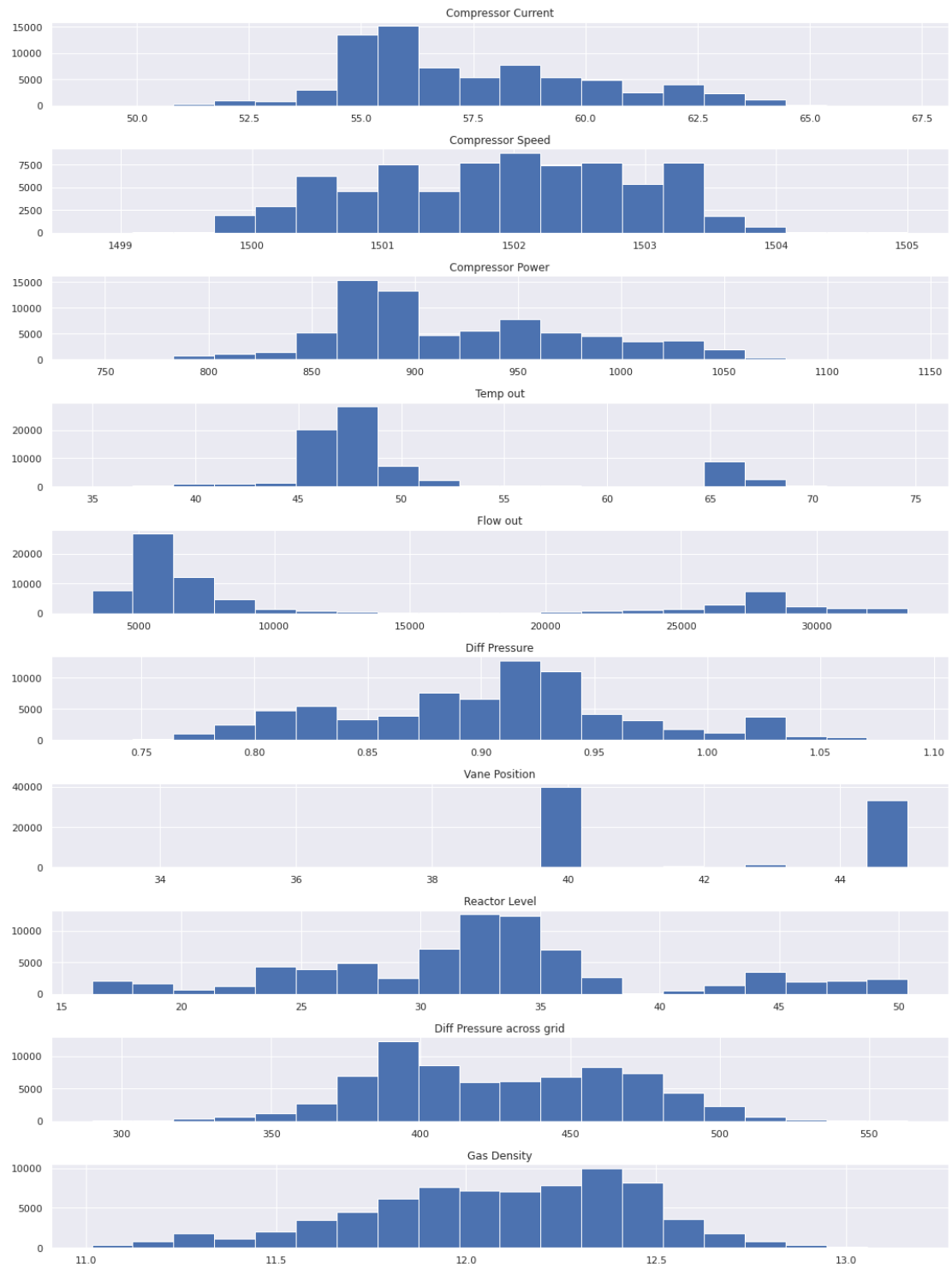
ชุดข้อมูลที่นำมาศึกษาถูกบันทึกอยู่ในช่วงวันที่ 24 กุมภาพันธ์ 2020 เวลา 1.00 น. ถึงวันที่ 25 กุมภาพันธ์ 2021 เวลา 13.55 น. ข้อมูลที่ได้รับจะถูกบันทึกทุก 5 นาที โดยมีข้อมูลรวมทั้งหมด 105,709 เพื่อความสะดวกในการวิเคราะห์ หลังจากนั้นนำมาวาดกราฟพบว่าตัวแปรอัตราการไหลออกพบว่าการพุ่งขึ้นของค่าอย่างผิดปกติ ซึ่งมีค่าต่างจากค่าปกติอย่างมากซึ่งเป็นเงื่อนไขการทำงานในรูปแบบสถานการณ์ทำงานพิเศษ (การเพิ่มกำลังการผลิต) ที่ทำให้ค่าของอัตราการไหลมีการพุ่งขึ้นจากค่าปกติ ซึ่งเป็นปัจจัยภายนอกที่ไม่สามารถควบคุมได้และยังมีค่าที่มีค่าเป็นศูนย์ เป็นช่วงที่มีช่วงหยุดการทำงาน อาจเกิดมาจากการปิดเพื่อบำรุงรักษา แสดงดังรูป 24



รูป 24 อัตราการไหลในรูปแบบเวลาและความถี่

#### 4.3 การทำเหมืองข้อมูล

เมื่อได้รับข้อมูลที่น่ามาวิเคราะห์ ต่อมาจะทำเหมืองข้อมูลเป็นการวิเคราะห์ข้อมูลจำนวนมาก เพื่อหาความสัมพันธ์ของข้อมูล เพื่อจำแนกประเภท รูปแบบ และหาความน่าจะเป็นที่จะเกิดขึ้น และนำไปประกอบการตัดสินใจในด้านต่างๆ เมื่อรวบรวมเสร็จ เนื่องจากข้อมูลที่ได้รับเป็นข้อมูลดิบที่ไม่สมบูรณ์ยังไม่พร้อมใช้งาน จำเป็นต้องมีการปรับปรุงข้อมูลให้พร้อมใช้งานโดยมีเรียกขั้นตอนนี้ว่า การเตรียมข้อมูล (Data preprocess) เป็นขั้นตอนที่สำคัญเนื่องจากจะทำให้ขั้นตอนอื่นๆมีประสิทธิภาพเพิ่มขึ้น โดยจะเริ่มจากการทำความสะอาดข้อมูล (Data Cleaning) เป็นการทำความสะอาดข้อมูลเพื่อตรวจสอบข้อผิดพลาดที่อาจเกิดขึ้นกับข้อมูล เช่น ข้อมูลซ้ำซ้อน ข้อมูลไม่ถูกต้อง หรือ ข้อมูลสูญหาย โดยในงานนี้จะทำการลบข้อมูลที่มีปัญหาออกไปทิ้งแล้ว ปัญหาค่าความผิดปกติของข้อมูล (Outlier) เป็นค่าที่มีการกระจายตัวเกินกว่าปกติ โดยจะใช้ค่าพิสัยระหว่างควอไทล์ (Interquartile range, IQR) เพื่อระบุค่าผิดปกติ เนื่องจากข้อมูลที่ได้รับมามีขนาดที่ต่างกันมาก ทำให้ต้องมีการปรับค่าข้อมูลโดยใช้เทคนิค Normalization เป็นการปรับช่วงขอบเขตของข้อมูลให้อยู่ในช่วงเดียวกันเพื่อที่เหมาะกับการนำไปคำนวณต่อไป โดยจะอยู่ในช่วง  $[0, 1]$  ถัดไปจะทำการเลือกคุณลักษณะ (Feature) ของข้อมูลเพื่อนำมาวิเคราะห์เนื่องจากข้อมูลที่ได้รับมีจำนวนมาก ขั้นตอนนี้ทำเพื่อลดจำนวนข้อมูลที่น่ามาวิเคราะห์เพื่อประหยัดเวลารวมถึงต้นทุนในการคำนวณ โดยจะตัดคุณลักษณะที่ไม่เกี่ยวข้องหรือมีความสัมพันธ์น้อยออก โดยดูการกระจายตัวของข้อมูลโดยแสดงในรูปของกราฟฮิสโตแกรม แสดงดังรูปที่ 25



รูป 25 การกระจายตัวของข้อมูลเครื่องอัดอากาศ

จากแผนภาพฮิสโตแกรมพบว่าข้อมูลมีการกระจายตัวแบบปกติ แต่มีบางส่วนที่มีการกระจายตัวแบบ 2 ยอด คือมีการกระจายตัวข้อมูล 2 กลุ่ม ที่มีจุดสูงสุด 2 ตำแหน่ง เกิดจากข้อมูลมีความแตกต่างกัน ซึ่งเป็นข้อมูลที่มีนัยสำคัญ

#### 4.4 การจำลองความผิดปกติ

เนื่องจากข้อมูลที่ได้รับเป็นข้อมูลจริงที่ไม่มีค่าผิดปกติ จึงทำการจำลองข้อมูลผิดปกติเข้าไปในข้อมูล โดยเป็นความผิดปกติสะสมภายในโดยจะมีลักษณะเพิ่มสูงขึ้น ในการจำลองข้อมูลผิดปกติ จะใช้การเพิ่ม/ลดค่า 1.5 เท่าของส่วนเบี่ยงเบนมาตรฐานโดยจะแบ่งเป็นช่วงตามที่กำหนด ซึ่งจะเพิ่มให้กับตัวแปรที่มีผลกับความผิดปกติในช่วงนั้น ๆ [35] โดยแต่ละช่วงจะมีสาเหตุ ลักษณะของข้อมูลที่แตกต่างกัน โดยได้ทำการทดลองกับชุดข้อมูล MaFaultDa [36] เพื่อเปรียบเทียบประสิทธิภาพในการตรวจจับความผิดปกติจากการเรียนรู้ความผิดปกติที่จำลองให้ข้อมูล

การจำลองความผิดปกติกับชุดข้อมูล MaFaultDa

MaFaultDa เป็นชุดข้อมูลจากเครื่องจักรหมุนเหวี่ยงในสภาวะการทำงานผิดปกติ และบันทึกค่าตามตำแหน่งต่าง ๆ ซึ่งเป็นชุดข้อมูลที่มีทั้งข้อมูลปกติและข้อมูลผิดปกติ ผิดปกติ เพื่อยืนยันความถูกต้องของวิธีการที่นำเสนอ เนื่องจากข้อมูลที่ได้รับไม่มีข้อมูลที่มีความผิดปกติ จึงจำเป็นต้องจำลองความผิดปกติในชุดข้อมูลที่มีโดยใช้วิธีการเพิ่มค่าให้กับตัวแปรที่ส่งผลต่อลักษณะการเกิดความเสียหายรูปแบบนั้น โดยข้อมูลแสดงถึงค่าที่ได้จากเซนเซอร์ดังนี้ 1. บันทึกค่าจากเครื่องมือวัดการหมุนรอบต่อ นาที (Tachometer) เพื่อวัดค่าความถี่ของการหมุน 2-4 เซนเซอร์ที่บันทึกค่าจากเครื่องมือวัดความเร่ง (Accelerometers) 3 แกนจากการ คือ ค่าความเร่งในแนวแกน (axial) ค่าความเร่งในแนวรัศมี (radial) และ ค่าความเร่งในแนวเส้นสัมผัสผิว (tangential) บันทึกค่าจากเครื่องบันทึกเสียง (Microphone) เพื่อบันทึกเสียงขณะที่ระบบกำลังทำงาน แสดงดังตารางที่ 1

ในการจำลองจะเริ่มจากการศึกษาข้อมูลและพฤติกรรมของข้อมูลในแต่ละรูปแบบความผิดปกติ โดยในที่นี้จะศึกษา รูปแบบการเยื้องศูนย์ในแนวตั้ง โดยจะนำข้อมูลมาเปรียบเทียบกับชุดข้อมูลการทำงานปกติจะพบว่า มีการเพิ่ม/ลดของข้อมูลที่ต่างกัน แสดงดังรูปที่ 26 โดยจะเห็นว่าข้อมูลบางตัวที่มีค่าเกินกว่าความปกติ จากนั้นเพิ่มค่าให้กับชุดข้อมูลเพื่อเป็นการจำลองชุดข้อมูลผิดปกติ โดยการนำข้อมูลปกติมาเพิ่มค่า 3 เท่าของส่วนเบี่ยงเบน แล้วนำไปให้แบบจำลองเรียนรู้ หลังจากนั้นจะใช้ข้อมูลความผิดปกติที่เกิดขึ้นจริงกับเป็นชุดทดสอบเพื่อหาความสามารถในการคัดแยกความผิดปกติของ

วิธีการที่เสนอ เมื่อทำการจำลองความผิดปกติให้กับข้อมูล หากนำมาเปรียบเทียบลักษณะของข้อมูล  
ปกติ ข้อมูลผิดปกติจริง และ ข้อมูลผิดปกติจำลอง แสดงดังรูปที่ 27

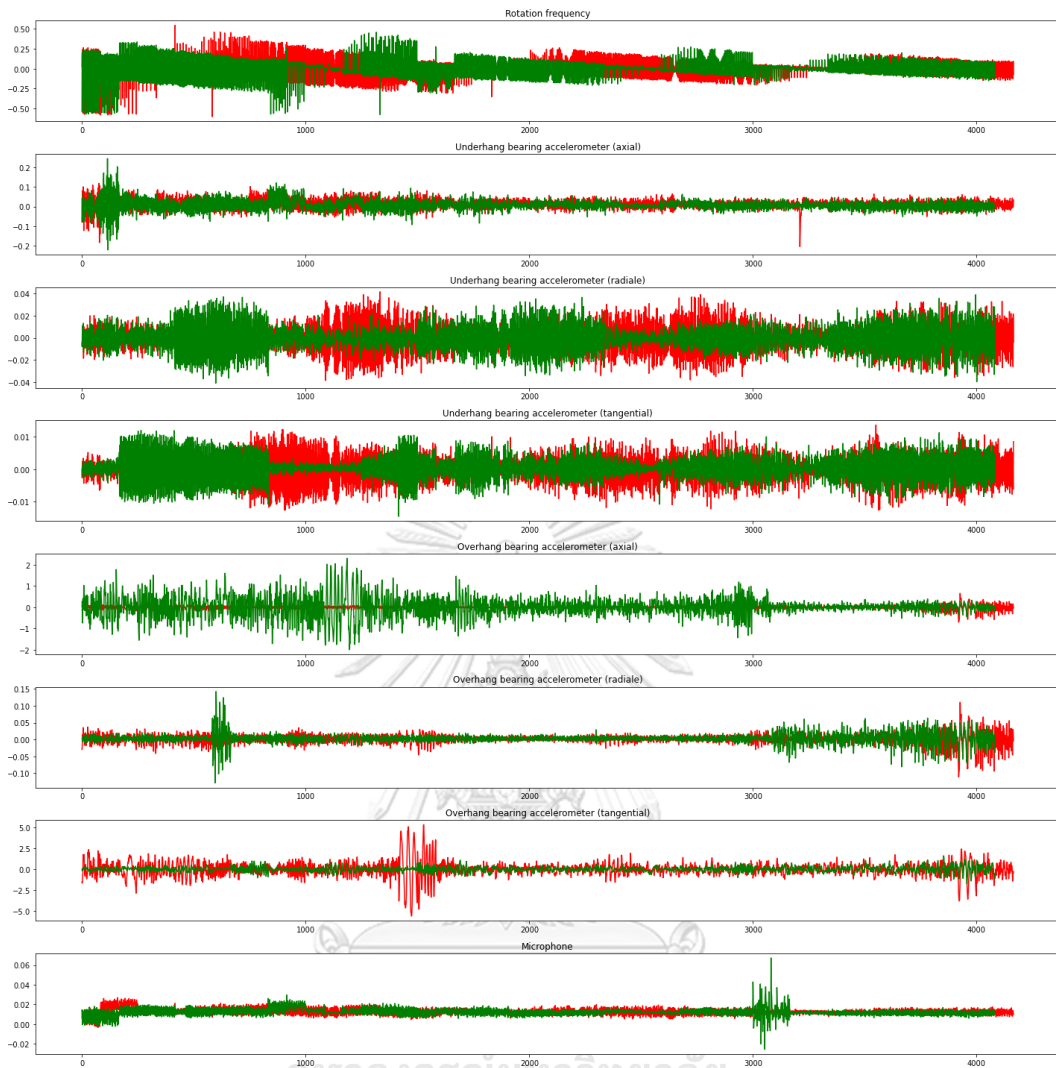
ตาราง 1 ข้อมูลที่วัดได้จากเซนเซอร์

ลำดับ	ค่าที่วัดได้
1	Rotation frequency
2	Underhang bearing accelerometer (axial)
3	Underhang bearing accelerometer (radiale)
4	Underhang bearing accelerometer (tangential)
5	Overhang bearing accelerometer (axial)
6	Overhang bearing accelerometer (radiale)
7	Overhang bearing accelerometer (tangential)
8	Microphone

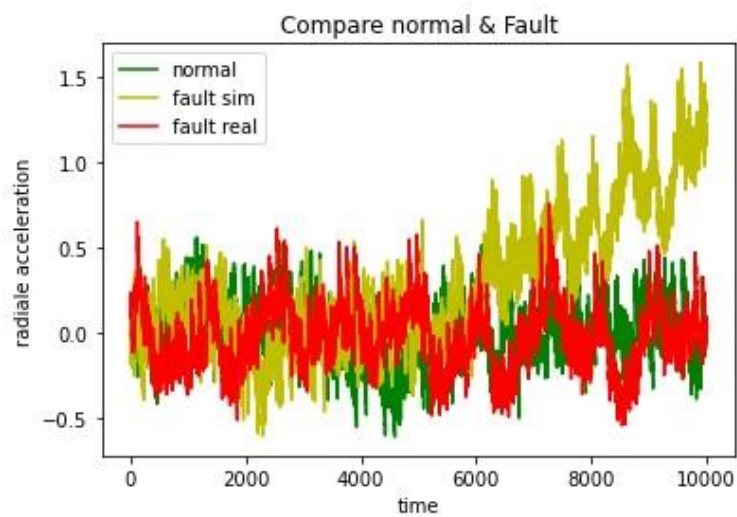
ผลที่ได้จากการใช้แบบจำลองพบว่า แบบจำลองสามารถแยกแยะข้อมูลปกติกับข้อมูลผิดปกติโดย  
เรียนรู้จากชุดข้อมูลที่จำลองความผิดปกติ แสดงดังตารางที่ 2 โดยจะมีความถูกต้องประมาณ 72 –  
82 % ซึ่งสามารถสรุปได้ว่าวิธีดังกล่าวสามารถนำมาใช้เพื่อจำลองความผิดปกติให้กับชุดข้อมูลที่ใช้ใน  
วิทยานิพนธ์นี้ได้

ตาราง 2 ความถูกต้องของแต่ละวิธีการ

Model	Accuracy	
	Train	Test
MLP	0.93	0.82
XGBoost	0.99	0.75
RF	0.89	0.72
SVM	0.91	0.74



รูป 26 เปรียบเทียบความสัมพันธ์ระหว่างข้อมูลปกติกับข้อมูลเชิงศูนย์ในแนวดิ่ง



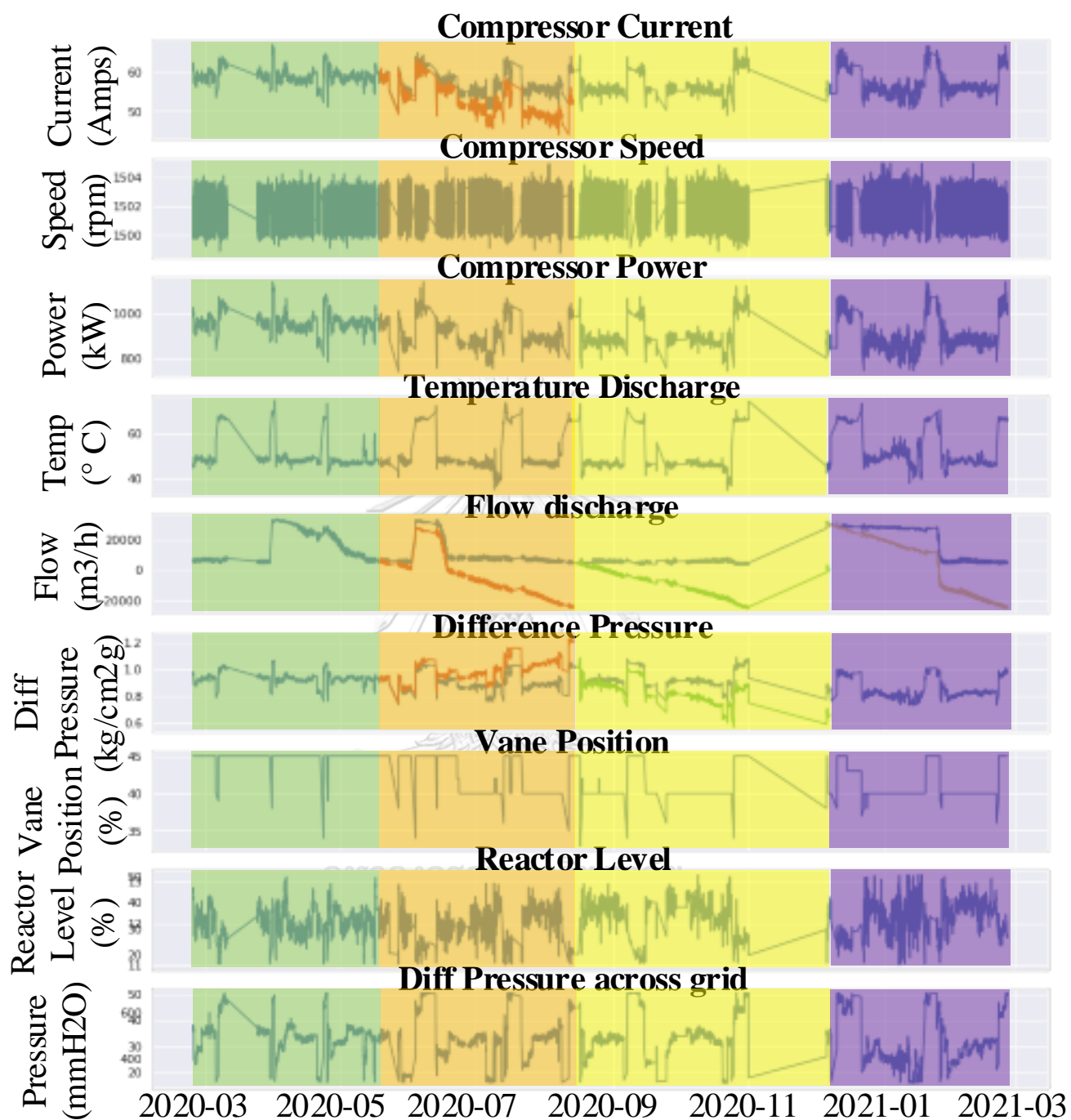
รูป 27 เปรียบเทียบลักษณะของข้อมูลระหว่างความเสียหายจริง ความเสียหายจำลอง และค่าปกติ



การจำลองความผิดปกติกับชุดข้อมูลเครื่องอัดอากาศแบบหมุนเหวี่ยง เมื่อได้รับชุดข้อมูลที่พร้อมใช้งานแล้ว ต่อมาจะทำการสร้างแบบจำลองในการเรียนรู้โดยใช้การเรียนรู้ของเครื่องเป็นเครื่องมือการเรียนรู้ประเภท การเรียนรู้แบบมีผู้สอน ซึ่งจำเป็นต้องใช้ข้อมูลที่มีสถานะข้อมูลกำกับไว้ แต่เนื่องจากข้อมูลที่ได้รับมาไม่มีสถานะกำกับ จึงมีการจำลองข้อมูลที่มีความผิดปกติใส่เข้าไปในข้อมูลที่มีแล้วกำหนดเป็นสถานะที่มีความผิดปกติ โดยจะแบ่งข้อมูลออกเป็น 4 ช่วง ดังนี้

1. ช่วงข้อมูลปกติ
2. ช่วงผิดปกติที่เกิดมาจากการอุดตันที่ทางเข้า
3. ช่วงผิดปกติที่ทำให้ประสิทธิภาพลดลง
4. ช่วงผิดปกติที่เกิดจากการอุดตันที่ทางออก

โดยช่วงข้อมูลปกติจะเป็นช่วงที่เครื่องจักรดำเนินการอย่างปกติ (ช่วงแถบสีเขียว) จะอยู่ในช่วงวันที่ 24 กุมภาพันธ์ 2020 เวลา 13.00 น. ถึง วันที่ 18 พฤษภาคม 2020 เวลา 21.00 น. ในการเพิ่มความเสียหายลงในข้อมูลจะทำโดยเลือกคุณลักษณะที่ส่งผลให้เกิดความผิดปกติในรูปนั้นๆได้ จากนั้นทำการเพิ่ม/ลดค่าดังกล่าว ทำโดยเพิ่มค่าขึ้นอย่างต่อเนื่องให้เหมือนกับการที่ความเสียหายค่อยๆรุนแรงขึ้นเรื่อย [37] โดยจะเริ่มจากศูนย์และเพิ่มขึ้นอย่างต่อเนื่องจนถึง 3 เท่าของความแปรปรวนของคุณลักษณะนั้น ในช่วงผิดปกติที่เกิดจากการอุดตันที่ทางเข้า (ช่วงแถบสีส้ม) [38] [39] เป็นความผิดปกติที่เกิดจากมีสิ่งผิดปกติขวางทางไหลที่ทางเข้า [40] [41] [42] [43] โดยอัตราการไหลของแก๊สที่ทางเข้าโดยจะพบว่าถ้ามีการอุดตันที่ทางเข้าจะส่งผลให้แก๊สไหลเข้าได้น้อยลง ทำให้แก๊สที่จะไหลเข้าไปทำให้ใบพัดหมุนน้อยลง จะมีการเพิ่ม/ลดค่าที่เป็นลักษณะเด่นของการเกิดการอุดตันที่ทางเข้าคือ ผลต่างความดันมีค่าเพิ่มขึ้น และ ลดค่าอัตราการไหลออกและกระแสไฟฟ้าจากมอเตอร์ลดลง ข้อมูลช่วงนี้อยู่ระหว่างวันที่ 18 พฤษภาคม 2020 เวลา 21.05 น. ถึงวันที่ 20 มิถุนายน 2020 เวลา 1.35 น. โดยค่าที่เพิ่ม/ลด เข้าไปในช่วงนี้จะแสดงเป็นสีแดง ช่วงผิดปกติที่ทำให้ประสิทธิภาพลดลง (ช่วงแถบสีเหลือง) เป็นความผิดปกติที่เกิดได้จากหลายสาเหตุ โดยคุณลักษณะที่มีจะเพิ่ม/ลดในช่วงนี้ได้แก่ อัตราการไหลออกและผลต่างความดันลดลง ข้อมูลช่วงนี้จะสิ้นสุดวันที่ 7 ธันวาคม 2020 เวลา 19.50 น. โดยค่าที่เพิ่ม/ลด เข้าไปในช่วงนี้จะแสดงเป็นสีเขียว ช่วงผิดปกติที่เกิดจากการอุดตันที่ทางออก (ช่วงแถบสีม่วง [44] [45]) เป็นความผิดปกติที่เกิดจากการสะสมของคราบสกปรก . โดยคุณลักษณะที่จะเพิ่ม/ลด คือ ผลต่างความดันระหว่างใบพัดเพิ่มขึ้น และ อัตราการไหลออกลดลง โดยค่าที่เพิ่ม/ลด เข้าไปในช่วงนี้จะแสดงเป็นสีแดง แสดงดังรูปที่ 28



รูป 28 การแบ่งช่วงของข้อมูลและค่าที่เพิ่ม/ลด

#### 4.5 แบบจำลอง

ในการสร้างแบบจำลองโดยใช้การเรียนรู้แบบมีผู้สอน เมื่อข้อมูลพร้อมแล้ว จะแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดข้อมูลที่ใช้เรียนรู้ (Training data) และ ชุดข้อมูลที่ใช้ทดสอบ (Testing data) โดยชุดข้อมูลที่ใช้เรียนรู้จะใช้เพื่อให้ข้อมูลเรียนรู้รูปแบบหรือจดจำรูปแบบ โดยจะมีขนาด 70 % ของชุดข้อมูลทั้งหมด และชุดข้อมูลทดสอบจะมีขนาด 30 % ของชุดข้อมูลทั้งหมด จากนั้นนำเข้าแบบจำลอง โดยแบบจำลองที่เลือกใช้จะมีทั้งหมด 4 ตัว ประกอบด้วย

- Random Forest (RF)
- XGBoost
- Support Vector Machine (SVM)
- Multilayer perceptron (MLP)

โดยเมื่อข้อมูลที่เตรียมมาผ่านขั้นตอนการเรียนรู้โดยแบบจำลองเรียนรู้ ก่อนที่จะนำเข้าสู่ขั้นตอนการทดสอบ จะมีการปรับค่าพารามิเตอร์เพื่อให้พารามิเตอร์ของแบบจำลองมีความถูกต้อง เหมาะสมกับชุดข้อมูลยิ่งขึ้น โดยการปรับค่าพารามิเตอร์จะใช้วิธีการค้นหาแบบกริด (Grid search) พารามิเตอร์ทุกชุดจะถูกนำมาพิจารณาและสังเกตความถูกต้อง เมื่อได้พิจารณาทั้งหมดแล้ว แบบจำลองที่มีชุดพารามิเตอร์ที่ให้ความถูกต้องมากที่สุดจะถือว่าเป็นชุดพารามิเตอร์ที่ดีที่สุด โดยแต่ละตัวจะมีการตั้งค่าพารามิเตอร์ที่ต่างกัน จึงสรุปมาเป็น ตารางที่ 3-6

ตาราง 3 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความเสียหายของเครื่องอัดอากาศ โดยใช้แบบจำลองป่าต้นไม้แบบสุ่ม

ลำดับ	ชื่อพารามิเตอร์	ค่า
1	Number of tree	500
2	Function criterion	gini
3	Bootstrap	False
4	Class_weight	balanced
5	Max_depth	10
6	Max_features	Sqrt
7	Max_sample	10

ตาราง 4 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความเสียหายของเครื่องอัดอากาศ โดยใช้แบบจำลองต้นไม้ตัดสินใจเอ็กซ์ทรีมเกรเดียนตึบู่ทติง

ลำดับ	ชื่อพารามิเตอร์	ค่า
1	Max_depth	6
2	Learning_rate	0.3
3	N_estimators	100
4	Booster	gbtree
5	Verbosity	1
6	objective	Multi:softprob
7	Gamma	0
8	Subsample	1

ตาราง 5 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความเสียหายของเครื่องอัดอากาศ โดยใช้แบบจำลองซัพพอร์ตเวคเตอร์แมชชีน

ลำดับ	พารามิเตอร์	ค่า
1	kernel	Sigmoid
2	Degree	3
3	Gamma	auto
4	Coef0	0
5	C	10

ตาราง 6 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความเสียหายของเครื่องอัดอากาศ โดยใช้แบบจำลอง MLP

ลำดับ	พารามิเตอร์	ค่า
1	Hidden layer sizes	(50,)
2	Activation function	Relu
3	solver	lbfgs
4	Batch size	500
5	Learning rate	invscaling

#### 4.6 ผลการทดลอง

การทดลองนี้มีวัตถุประสงค์เพื่อทำการเปรียบเทียบประสิทธิภาพในการตรวจจับและระบุรูปแบบความเสียหายที่ต่างกันของแบบจำลองด้วยค่า F1 ซึ่งผลการทดลองแสดงในตารางที่ 7 โดยการนำข้อมูลทดสอบมาใช้กับแบบจำลองที่นำเสนอมาพบว่า แบบจำลองการเรียนรู้แบบกลุ่ม (RF, XGBoost) ให้ประสิทธิภาพที่ดีกว่าและเวลาในการคำนวณที่สั้นกว่าการเรียนรู้ของเครื่อง และ มีค่าใกล้เคียงกับเครือข่ายประสาทเทียม โดยมีค่า F1 อยู่ที่ 93-96% ซึ่ง SVM และ MLP ได้ค่า F1 อยู่ที่ 73, 97% ตามลำดับ จากที่ SVM ได้คะแนนน้อยอาจจะมีผลมาจากการจำลองความผิดปกติในชุดข้อมูลทดสอบได้เพิ่มค่าความผิดปกติไป 1.5 เท่าของส่วนเบี่ยงเบนมาตรฐานแต่ชุดข้อมูลเรียนรู้ได้เพิ่มค่าความผิดปกติไป 3 เท่าของส่วนเบี่ยงเบนมาตรฐานส่งผลให้ support vector อาจจะไม่เรียนรู้ในช่วงที่ต่างกันทำให้ประสิทธิภาพในการเรียนรู้มีค่าน้อย ในด้านของเวลาในการประมวลผลพบว่า RF และ XGBoost ได้ 34 และ 7 วินาทีตามลำดับ SVM มีเวลาประมวลผลที่ 38 วินาทีซึ่งเป็นเวลาประมวลผลที่นานที่สุดในทางกลับกัน MLP มีเวลาประมวลผลที่ 34 วินาที โดยจะพบว่าวิธีการส่วนใหญ่มีเวลาที่ใช้ในการวิเคราะห์ที่ใกล้เคียงกัน เว้นแต่ XGBoost ที่มีความเร็วในการประมวลผลที่น้อยมากเนื่องการทำงานเป็นแบบขนานทำให้สามารถคำนวณได้เร็วกว่าวิธีการอื่น วิธีการที่กล่าวมาทั้งหมดเมื่อถ้าหากไม่ผ่านกระบวนการ Normalize ซึ่งเป็นการแปลงค่าให้มีขนาดใกล้เคียงกัน จะส่งผลกระทบต่อความสามารถในการแบ่งแยกข้อมูลให้มีประสิทธิภาพที่ลดลง ยกเว้นการเรียนรู้แบบหมู่ไม่ได้รับผลกระทบจากกระบวนการดังกล่าว แสดงดังตารางที่ 8

ตาราง 7 ประสิทธิภาพแต่ละแบบจำลองที่ผ่านกระบวนการ Normalization

Algorithm	Training Accuracy	Testing Accuracy	Precision	Recall	F1-score	Executed time
RF	0.99	0.92	0.93	0.9	0.93	34
XGBoost	1.00	0.96	0.96	0.96	0.96	7
SVM	0.81	0.72	0.80	0.73	0.73	<b>38</b>
MLP	0.98	0.97	0.98	0.97	0.97	34

ตาราง 8 ประสิทธิภาพแต่ละแบบจำลองที่ไม่ผ่านกระบวนการ normalize

Algorithm	Training Accuracy	Testing Accuracy	Precision	Recall	F1-score	Executed time
RF	0.99	0.95	0.94	0.94	0.95	15
XGBoost	1.00	0.95	0.94	0.94	0.96	3.2
SVM	0.92	0.88	0.89	0.85	0.88	<b>282</b>
MLP	0.48	0.44	0.11	0.25	0.44	2.3

อีกทั้งข้อมูลที่ใช้ในการทดลองมีขนาด 37,671 ข้อมูล โดยจะแบ่งเป็น

1. ช่วงข้อมูลปกติ 13,130 ข้อมูล
2. ช่วงผิดปกติการอุดตันที่ทางเข้า 9,489 ข้อมูล
3. ช่วงผิดปกติประสิทธิภาพลดลง 7,432 ข้อมูล
4. ช่วงผิดปกติการอุดตันที่ทางออก 7,620 ข้อมูล

ซึ่งจะเห็นว่ามีความต่างของขนาดข้อมูลแต่ละชนิด แสดงถึงความไม่สมดุลของข้อมูลแต่ยังมีความต่างที่ไม่มาที่จะส่งผลให้เห็นอย่างชัดเจน ซึ่งปัญหาดังกล่าวอาจส่งผลต่อการเรียนรู้ของแบบจำลองทำให้เรียนรู้รูปแบบที่มีขนาดข้อมูลน้อยได้ไม่ดี แต่การเรียนรู้แบบหมู่สามารถจัดการปัญหาดังกล่าวได้โดยจะพบว่า ตารางที่ 7 ค่าช่วงที่ 3-4 มีประสิทธิภาพที่ลดลงจากเดิม

ตาราง 9 ค่า F1-score แต่ละรูปแบบ

Algorithms	F1-score			
	Normal	Suction blockage	Efficiency problem	Discharge blockage
RF	0.96	0.96	0.96	0.88
XGBoost	0.97	0.98	0.95	0.95
SVM	0.70	0.50	0.83	0.73
MLP	0.95	0.96	1.00	0.97

## บทที่ 5

### สรุปและวิจารณ์ผลการทดสอบ

ในบทนี้จะกล่าวถึงผลการดำเนินงานวิทยานิพนธ์ ประกอบด้วยผลการทดลองขั้นตอนวิธีที่ใช้ปัญหาที่พบในการจัดทำวิทยานิพนธ์ และงานวิจัยในอนาคต

#### 5.1 สรุปและวิจารณ์ผลการทดสอบ

ในวิทยานิพนธ์นี้เป็นการประยุกต์วิธีการตรวจจับและระบุรูปแบบความผิดปกติของเครื่องอัดอากาศแบบหมุนเหวี่ยง โดยใช้การเรียนรู้แบบหมู่เป็นวิธีการหลักในการทดลองโดยใช้เทคนิคการปรับค่าพารามิเตอร์เพื่อให้ได้ค่าพารามิเตอร์ที่เหมาะสมในการทำนายส่งผลทำให้แบบจำลองที่ได้มีประสิทธิภาพมากขึ้น โดยจะพบว่าการเรียนรู้แบบหมู่มีความสามารถที่จะตรวจจับความผิดปกติแบบหลายความผิดปกติได้ดีเมื่อรู้ข้อมูลรูปแบบความผิดพลาดเมื่อเทียบกับการเรียนรู้ของเครื่อง และเครือข่ายประสาทเทียม อีกทั้งใช้เวลาในการประมวลผลที่สั้นเมื่อเทียบกับประสิทธิภาพที่ได้รับ โดยได้ทดลองปรับข้อมูลให้มีขนาดที่ต่างกันมาก ซึ่งจะก่อให้เกิดปัญหาข้อมูลไม่สมดุลพบว่าการเรียนรู้แบบหมู่สามารถจัดการปัญหาดังกล่าวได้ดีกว่า วิธีอื่นโดยสังเกตได้จากค่า F1-score ของแต่ละรูปแบบจะพบว่าการเรียนรู้แบบกลุ่มได้ค่าที่มากกว่า แต่ปัญหาที่พบคือในช่วงข้อมูลผิดที่เกิดจากการอดตันที่ทางเข้า รูปแบบข้อมูลผิดปกติกับข้อมูลปกติมีความใกล้เคียงกัน ส่งผลให้วิธีการเรียนรู้รูปแบบโดยใช้แบบจำลองอาจไม่สามารถแยกแยะได้ ทำให้ในรูปแบบผิดปกติดังกล่าวมีค่าความถูกต้องที่น้อย หากข้อมูลมีขนาดที่ต่างกันเยอะกระบวนการเตรียมข้อมูล เช่น Normalization จะมีความสำคัญต่อบางแบบจำลอง ผลจากการทดลองพบว่าการเรียนรู้แบบหมู่ไม่ได้รับผลจากขนาดข้อมูลที่ต่างกันมาก ในทางกลับกัน MLP มีประสิทธิภาพที่น้อยมากเมื่อเจอปัญหาดังกล่าว เนื่องจากการเรียนรู้แบบหมู่มีพื้นฐานจากต้นไม้ตัดสินใจซึ่งเป็นการวิเคราะห์คุณลักษณะที่มีผลต่อการแบ่งข้อมูลมากที่สุด ทำให้ได้รับผลกระทบจากขนาดข้อมูลที่ต่างกันมากได้น้อย ต่างกับวิธีของ MLP และ SVM ซึ่งมีพื้นฐานการคำนวณจากสมการโดยเมื่อข้อมูลมีขนาดต่างกันมากๆ ทำให้ส่งผลต่อการแบ่งแยกข้อมูลให้ประสิทธิภาพไม่ดีเท่าที่ควร

#### 5.2 ปัญหาที่พบในการจัดทำวิทยานิพนธ์

1. ชุดข้อมูลจริงของอุปกรณ์ดังกล่าวที่ปฏิบัติในอุตสาหกรรมจริง ไม่มีการเผยแพร่ทำให้ไม่สามารถหางานวิจัยที่ใช้ชุดข้อมูลเดียวกัน เพื่อนำมาใช้ในการเปรียบเทียบผลลัพธ์กับวิธีการอื่น

2. ข้อมูลของเครื่องอัดอากาศที่ได้รับไม่มีสถานะระบุที่ชัดเจน ทำให้ไม่สามารถยืนยันได้ว่า ข้อมูลที่ได้ไม่มีความเสียหายเกิดขึ้น

3. เครื่องจักรมีความซับซ้อนรวมถึงมีความสัมพันธ์กับอุปกรณ์อื่นในระบบซึ่งทำให้ยากที่จะ เข้าใจภาพรวมทางกายภาพของระบบ เพื่อที่จะคาดเดาค่าที่ควรจะเป็นของระบบ

### 5.3 งานวิจัยในอนาคต

เนื่องจากลักษณะข้อมูลเครื่องอัดอากาศที่ได้รับมีการเปลี่ยนแปลงค่าอย่างกระตั้นหัน เนื่องจากมีการปรับเปลี่ยนกำลังการผลิตและอาจจะมีปัจจัยภายนอกอื่นที่ควรนำมาพิจารณา เช่น สภาพแวดล้อมที่เปลี่ยนไปรวมถึงอาจจะมี ความเสียหายรูปแบบอื่นที่ไม่ได้ถูกนำมาใช้ในวิทยานิพนธ์ นี้ เพื่อช่วยให้วิศวกรซ่อมบำรุงสามารถตรวจสอบได้อย่างถูกต้องตรงจุดมากขึ้น

ในการจำลองข้อมูลผิดปกติขึ้นทำโดยการเพิ่มค่าบางตัวที่ส่งผลต่อความเสียหายในรูปแบบนั้น ๆ ซึ่งในความเป็นจริงอาจจะมีตัวแปรอื่นที่เปลี่ยนแปลงไปด้วย แต่ไม่ได้นำมาปรับค่าด้วย เพื่อให้มีความสมจริงยิ่งขึ้น



## บรรณานุกรม

1. Dafflon, B., *The challenges, approaches, and used techniques of CPS for manufacturing in Industry 4.0: a literature review*. The International Journal of Advanced Manufacturing Technology 2020.
2. Bloch, H.P. and A. Godse, *Compressors and modern process applications*. 2006: John Wiley & Sons.
3. Andrea, F., et al. *Compressor station facility failure modes: causes, taxonomy and effects*. 2020.
4. Amin, A.A., M.T. Maqsood, and K. Mahmood-ul-Hasan, *Surge protection of centrifugal compressors using advanced anti-surge control system*. Measurement and Control, 2021. 54(5-6): p. 967-982.
5. Batayev, N., B. Suleimenov, and S. Batayeva, *Centrifugal compressor anti-surge control system modelling*. International Journal of Electrical and Computer Engineering, 2022. 12(2): p. 1419.
6. Hafaifa, A., B. Rachid, and G. Mouloud, *Modelling of surge phenomena in a centrifugal compressor: experimental analysis for control*. Systems Science & Control Engineering: An Open Access Journal, 2014. 2(1): p. 632-641.
7. Moore, J.J., et al. *Transient Surge Measurements Of A Centrifugal Compressor Station During Emergency Shutdowns*. in *Proceedings of the 38th Turbomachinery Symposium*. 2009. Texas A&M University. Turbomachinery Laboratories.
8. Almasi, A., *Latest techniques and practical notes on anti-surge systems for centrifugal compressors*. Australian journal of mechanical engineering, 2012. 10(1): p. 81-90.
9. Aretakis, N. and K. Mathioudakis, *Classification of radial compressor faults using pattern-recognition techniques*. Control Engineering Practice, 1998. 6(10): p. 1217-1223.
10. Clements, J. and J. Bekkers, *Detection of spontaneous synaptic events with an optimally scaled template*. Biophysical journal, 1997. 73(1): p. 220-229.

11. Frank, P.M., *Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results*. *automatica*, 1990. 26(3): p. 459-474.
12. Tahan, M., M. Muhammad, and Z.A. Karim. *A framework for intelligent condition-based maintenance of rotating equipment using mechanical condition monitoring*. in *MATEC web of conferences*. 2014. EDP Sciences.
13. Laosiritaworn, W. and T. Bunjongjit, *Classification techniques for control chart pattern recognition: A case of metal frame for actuator production*. *Chiang Mai J. Sci*, 40 (4): 701, 2013. 712.
14. Zhao, Y., et al., *A robust pattern recognition-based fault detection and diagnosis (FDD) method for chillers*. *HVAC&R Research*, 2014. 20(7):p. 798-809.
15. Yinghua, Y., P. Yongkang, and Z. Liping. *Fault Monitoring Method Based on Mutual Information and Relative Principal Component Analysis*. in *2019 Chinese Control And Decision Conference (CCDC)*. 2019. IEEE.
16. Li, X., et al., *A novel diagnostic and prognostic framework for incipient fault detection and remaining service life prediction with application to industrial rotating machines*. *Applied Soft Computing*, 2019. 82: p. 105564.
17. El Koujok, M., et al., *A multiagent-based methodology for known and novel faults diagnosis in industrial processes*. *IEEE Transactions on Industrial Informatics*, 2020. 17(5): p. 3358-3366.
18. Aldrich, C. and L. Auret, *Fault detection and diagnosis with random forest feature extraction and variable importance methods*. *IFAC Proceedings Volumes*, 2010. 43(9): p. 79-86.
19. Puggini, L., J. Doyle, and S. McLoone, *Fault detection using random forest similarity distance*. *IFAC-PapersOnLine*, 2015. 48(21): p. 583-588.
20. TÜYSÜZÖĞLU, G. and D. Birant, *Enhanced bagging (eBagging): A novel approach for ensemble learning*. *International Arab Journal of Information Technology*, 2020. 17(4).
21. Zhang, D., et al., *A data-driven design for fault detection of wind turbines using random forests and XGboost*. *IEEE Access*, 2018. 6: p. 21020-21031.
22. Salunkhe, U.R. and S.N. Mali, *Classifier ensemble design for imbalanced data*

- classification: a hybrid approach*. *Procedia Computer Science*, 2016. 85: p. 725-732.
23. Deleplace, A., et al., *Ensemble Learning-based Fault Detection in Nuclear Power Plant Screen Cleaners*. *IFAC-PapersOnLine*, 2020. 53(2): p. 10354-10359.
  24. Patil, P.S., et al., *Ensembles of Ensemble Machine Learning approach for Fault detection of Bearing*. *Solid State Technology*, 2020. 63(6): p. 14442-14455.
  25. Korpela, S.A., *Principles of turbomachinery*. 2019: John Wiley & Sons.
  26. Hanlon, P.C., *Compressor handbook*. 2001: McGraw-Hill Professional.
  27. Isermann, R., *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. 2005: Springer Science & Business Media.
  28. Lughofer, E. and M. Sayed-Mouchaweh, *Predictive maintenance in dynamic systems: advanced methods, decision support tools and real-world applications*. 2019: Springer.
  29. Miljković, D. *Fault detection methods: A literature survey*. in 2011 *Proceedings of the 34th international convention MIPRO*. 2011. IEEE.
  30. Duda, R.O. and P.E. Hart, *Pattern classification*. 2006: John Wiley & Sons.
  31. Tufféry, S., *Data mining and statistics for decision making*. 2011: John Wiley & Sons.
  32. Yang, Y. and H. Lv, *Discussion of Ensemble Learning under the Era of Deep Learning*. arXiv preprint arXiv:2101.08387, 2021.
  33. Charbuty, B. and A. Abdulazeez, *Classification based on decision tree algorithm for machine learning*. *Journal of Applied Science and Technology Trends*, 2021. 2(01): p. 20-28.
  34. Tu, M.C., D. Shin, and D. Shin. *A comparative study of medical data classification methods based on decision tree and bagging algorithms*. in 2009 *Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing*. 2009. IEEE.
  35. Zanolli, S.M., G. Astolfi, and L. Barboni, *Principal component analysis based on anova test for multishaft centrifugal compressor fault detection and diagnosis*. *IFAC Proceedings Volumes*, 2010. 43(21): p. 90-96.
  36. Ribeiro, F.M., et al., *Rotating machinery fault diagnosis using similarity-based*

- models*. XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais-sbrt2017, 2017.
37. Li, X., et al., *Condition monitoring of rotating machines under time-varying conditions based on adaptive canonical variate analysis*. Mechanical Systems and Signal Processing, 2019. 131: p. 348-363.
  38. Bryce, C., et al., *Advanced two-stage compressor program design of inlet stage*. 1973.
  39. Furukawa, M., et al., *The role of tip leakage vortex breakdown in compressor rotor aerodynamics*. 1999.
  40. Al-Busaidi, W. and P. Pilidis, *Investigation of efficiency deterioration causes in process centrifugal compressor operation*. Journal of Failure Analysis and Prevention, 2016. 16(1): p. 19-36.
  41. Cavazzini, G., et al., *Using splitter blades to improve suction performance of centrifugal impeller pumps*. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, 2015. 229(3): p. 309-323.
  42. Mokhatab, S., W.A. Poe, and J. Mak, *Handbook of natural gas transmission and processing: principles and practices*. 2018: Gulf professional publishing.
  43. Rapur, J.S. and R. Tiwari. *A robust algorithm to detect multiple centrifugal pump faults with corrupted vibration and current signatures using continuous wavelet transform*. in *Proceedings of the 34th International Pump Users Symposium*. 2018. Turbomachinery Laboratory, Texas A&M Engineering Experiment Station.
  44. Boyce, M.P. *Principles Of Operation And Performance Estimation Of Centrifugal Compressors*. in *Proceedings of the 22nd Turbomachinery Symposium*. 1993. Texas A&M University. Turbomachinery Laboratories.
  45. Müller, M.W., H.-P. Schiffer, and C. Hah. *Effect of circumferential grooves on the aerodynamic performance of an axial single-stage transonic compressor*. in *Turbo Expo: Power for Land, Sea, and Air*. 2007.

## ประวัติผู้เขียน

ชื่อ-สกุล	ธิดิพัทธ์ สนธิโพธิ์
วัน เดือน ปี เกิด	16 มกราคม 2541
สถานที่เกิด	กรุงเทพมหานคร ประเทศไทย
วุฒิการศึกษา	ปริญญาตรีที่มหาวิทยาลัยมหิดล
ที่อยู่ปัจจุบัน	77/51 หมู่12 ต.ไร่ขิง อ.สามพราน จ.นครปฐม 73210
ผลงานตีพิมพ์	-
รางวัลที่ได้รับ	-



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY