



การบำรุงรักษาเชิงพยากรณ์ของกระบวนการผลิตด้วยการเรียนรู้ของเครื่อง
PREDICTIVE MAINTENANCE IN MANUFACTURING PROCESS USING MACHINE
LEARNING



วีรชิต ตั้งศิริวัฒนวงศ์

บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ

2564

การบำรุงรักษาเชิงพยากรณ์ของกระบวนการผลิตด้วยการเรียนรู้ของเครื่อง



สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตร์มหาบัณฑิต สาขาวิชาวิทยาการข้อมูล
คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2564

ลิขสิทธิ์ของมหาวิทยาลัยศรีนครินทรวิโรฒ

PREDICTIVE MAINTENANCE IN MANUFACTURING PROCESS USING MACHINE
LEARNING



WEERACHIT TUNGSIRIWATTANAWONG

A Master's Project Submitted in Partial Fulfillment of the Requirements
for the Degree of MASTER OF SCIENCE
(Data Science)

Faculty of Science, Srinakharinwirot University

2021

Copyright of Srinakharinwirot University

สารนิพนธ์

เรื่อง

การบำรุงรักษาเชิงพยากรณ์ของกระบวนการผลิตด้วยการเรียนรู้ของเครื่อง

ของ

วีรชิต ตั้งศิริวัฒนวงศ์

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล

ของมหาวิทยาลัยศรีนครินทรวิโรฒ

(รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)

คณบดีบัณฑิตวิทยาลัย

คณะกรรมการสอบปากเปล่าสารนิพนธ์

ที่ปรึกษาหลัก

ประธาน

(ผู้ช่วยศาสตราจารย์ ดร.จันตรี ผลประเสริฐ)

(อาจารย์ ดร.สุทธิพงศ์ รัชชพงษ์)

กรรมการ

(อาจารย์ ดร.ศิริสรรพ เหล่าหะเกียรติ)

ชื่อเรื่อง	การบำรุงรักษาเชิงพยากรณ์ของกระบวนการผลิตด้วยการเรียนรู้ของเครื่อง
ผู้วิจัย	วีรชิต ตั้งศิริวัฒนวงศ์
ปริญญา	วิทยาศาสตร์มหาบัณฑิต
ปีการศึกษา	2564
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. จันตรี ผลประเสริฐ

งานวิจัยนี้นำเสนอการใช้เทคโนโลยีการเรียนรู้ของเครื่องในการวิเคราะห์สัญญาณความสั่นสะเทือนของตลับลูกปืนเพื่อทำนายความผิดปกติที่จะเกิดขึ้นกับตลับลูกปืนจำนวน 3 ประเภทอันได้แก่ Inner, Outer และ Roller ระบบที่นำเสนอจะนำสัญญาณความสั่นสะเทือนที่วัดจาก accelerometer ที่ติดกับตลับลูกปืนมาคำนวณหาค่าทางสถิติพื้นฐาน เช่น ค่าเฉลี่ย ค่า Root Mean Square (RMS) ค่า Peak-to-Peak และอื่นๆ จำนวน 12 ค่าและนำมาใช้เป็นคุณลักษณะเฉพาะที่จะป้อนให้กับโมเดลการเรียนรู้ของเครื่องใช้ในการทำนายความผิดปกติของตลับลูกปืน เราทดสอบประสิทธิภาพของระบบที่นำเสนอกับชุดข้อมูล Intelligent Maintenance Systems (IMS) จากผลการทดลอง ระบบที่นำเสนอที่ใช้โมเดลการเรียนรู้ของเครื่องประกอบไปด้วย Support Vector Machine , KNN , Random forest และ Extreme Gradient Boosting มีค่าความแม่นยำในการทำนายความผิดปกติทั้ง 3 แบบที่ดีที่สุด โดยระบบที่นำเสนอสามารถตรวจจับความผิดปกติแบบ Inner, Outer และ Roller โดยมีค่าความแม่นยำ (Precision) เท่ากับ 79.66%, 90.99% และ 83.86% ตามลำดับ โดยคุณลักษณะเฉพาะที่สำคัญที่สุด 3 ลำดับแรกประกอบไปด้วยค่า RMS , Peak-to-Peak และ Clearance โดยมีความสำคัญเท่ากับ 0.46 , 0.14 และ 0.13 ตามลำดับ

คำสำคัญ : สั่นสะเทือน, ตลับลูกปืนชำรุด, การเรียนรู้ของเครื่อง

Title	PREDICTIVE MAINTENANCE IN MANUFACTURING PROCESS USING MACHINE LEARNING
Author	WEERACHIT TUNGSIRIWATTANAWONG
Degree	MASTER OF SCIENCE
Academic Year	2021
Thesis Advisor	Asst. Prof. Dr. Chantri Polprasert

This study investigates the performance of machine learning models in predictive maintenance problems. The vibration data (time-domain features) collected from accelerometers attached to the bearings used by the machine learning model to predict three types of bearing failures, including, inner, outer and roller failure. The proposed system extracted 12 simple statistical measures, such as mean, peak-to-peak and root-mean-square (RMS) values of the vibration signals and used as input features for machine learning models. The performance of the proposed system was tested on the Intelligent Maintenance Systems (IMS) dataset. The proposed system employed the Support Vector Machine, KNN, Random Forest and Extreme Gradient Boosting model yielded the best performance to detect inner, outer and roller failure and with precision equal to 79.66%, 90.99% and 83.86% respectively. The three most important features consisted of RMS, Peak-to-Peak and Clearance, with a feature importance equal to 0.46, 0.14 and 0.13, respectively.

Keyword : Vibration, Bearing failure, Machine learning

กิตติกรรมประกาศ

ขอกราบขอบพระคุณ บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ สำหรับทุนสนับสนุนการนำเสนอผลงานวิจัยของบัณฑิตศึกษา ทำให้ได้ประสบการณ์นำเสนอสารนิพนธ์และได้แลกเปลี่ยนความรู้ทางด้านเทคโนโลยีคณิตศาสตร์คอมพิวเตอร์และการใช้เทคนิคการเรียนรู้ของเครื่อง

วีรชิต ตั้งศิริวัฒนวงศ์

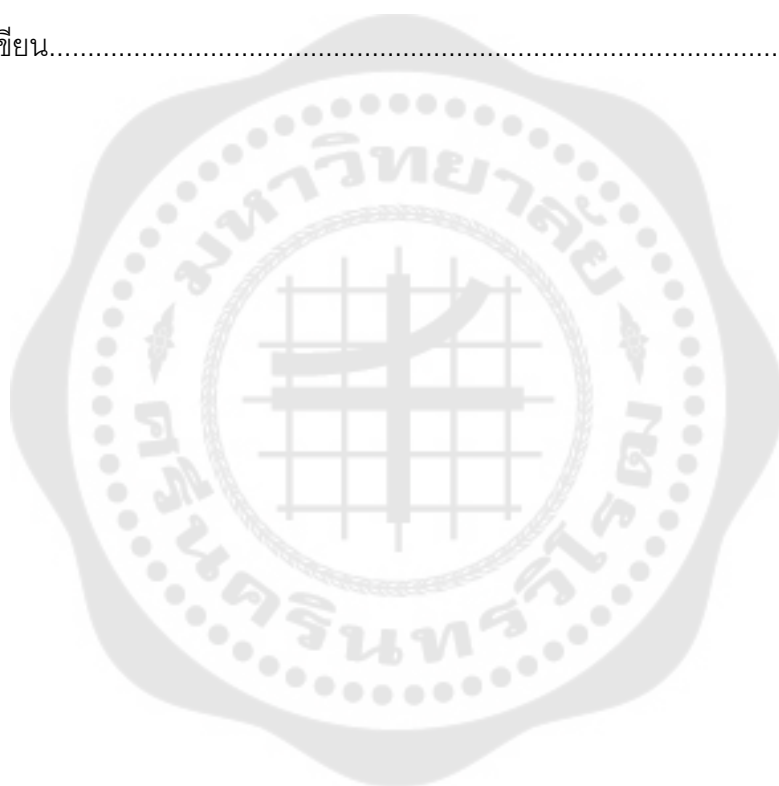


สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ	ช
สารบัญตาราง.....	ญ
สารบัญรูปภาพ	ฎ
บทที่ 1 บทนำ.....	1
1. ความสำคัญและความเป็นมาของงานวิจัย.....	1
1.1 การซ่อมบำรุงรักษาเชิงรับ (Reactive Maintenance)	1
1.2 การซ่อมบำรุงเชิงป้องกัน (Preventive maintenance)	1
1.3 การซ่อมบำรุงรักษาเชิงพยากรณ์ (Predictive maintenance).....	1
2. วัตถุประสงค์ของการวิจัย.....	2
3. ขอบเขตของการวิจัย และ การดำเนินงานวิจัย.....	3
4. ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย	3
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	4
1.1 ทฤษฎีเกี่ยวกับการเตรียมข้อมูล (Data Preprocessing)	4
1.1.1 การสันสะเทือน และวิธีการวัดการสันสะเทือน	4
1.1.2 รูปแบบการชำรุดของตลับลูกปืน.....	5
1.1.3 ค่าสันสะเทือนในคาบเวลา.....	6
1.2 ทฤษฎีเกี่ยวกับอัลกอริทึม	6
1.2.1 การเรียนรู้ของเครื่อง	6

1.2.2 K-Nearest Neighbor (KNN)	7
1.2.3 Support vector machine (SVM)	8
1.2.4 Random Forest (ป่าสุ่ม).....	9
1.2.5 Extreme Gradient Boosting (XG-Boost)	9
1.3 การวัดประสิทธิภาพของผลลัพธ์จากการวิเคราะห์ข้อมูล	10
1.4 งานวิจัยที่เกี่ยวข้อง	11
บทที่ 3 วิธีการดำเนินงานวิจัย.....	19
1. กระบวนการทำงานของแบบจำลอง.....	19
2. การเก็บรวบรวมข้อมูล	20
3. การเตรียมข้อมูล	22
3.1 การหาความสัมพันธ์ระหว่างคุณลักษณะเฉพาะ	25
3.2 การหาความสัมพันธ์ระหว่างชุดข้อมูลเดียวกันและต่างชุดข้อมูล	27
3.3 การหาคุณลักษณะเฉพาะที่สำคัญ (Feature important)	36
3.4 พิจารณาคลาสของข้อมูลแบบอนุกรมเวลา.....	37
3.5 พิจารณาคลาสของข้อมูลจากการแจกแจงแบบปกติ	40
3.6 การปรับปรุงพารามิเตอร์ของโมเดล และ การทำ Cross validation	42
บทที่ 4 ผลลัพธ์จากการวิจัย	44
1. ผลลัพธ์ลัพธ์ของทั้ง 4 โมเดล	44
1.1 โมเดล SVM	44
1.2 โมเดล KNN	46
1.3 โมเดลป่าสุ่ม	47
1.4 โมเดล XG-Boost.....	48
2. ผลลัพธ์ของการเปรียบเทียบโมเดลทั้ง 4 โมเดล	49

3. กราฟอธิบายผลการทดลองการแบ่งคลาสของโมเดล	49
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และ ข้อเสนอแนะ	53
1. สรุปผลการวิจัย.....	53
2. อภิปรายผลการวิจัย	54
3. ข้อเสนอแนะ	56
บรรณานุกรม	57
ประวัติผู้เขียน.....	60



สารบัญตาราง

	หน้า
ตาราง 1 รายละเอียดข้อมูล IMS	21
ตาราง 2 จำนวนข้อมูลสำหรับการเรียนรู้ของเครื่อง	22
ตาราง 3 ค่าทางสถิติของตลับลูกปืน 1 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 15 กุมภาพันธ์ 2004 0:52 ถึง 15 กุมภาพันธ์ 2004 4:04.....	28
ตาราง 4 ค่าทางสถิติของตลับลูกปืน 1 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 13 กุมภาพันธ์ 2004 10:02 ถึง 13 กุมภาพันธ์ 2004 02:13.....	28
ตาราง 5 ค่าทางสถิติของตลับลูกปืน 1 ชุดข้อมูล 3 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 05 มีนาคม 2004 22:42 ถึง 05 มีนาคม 2004 21:42.....	29
ตาราง 6 ค่าทางสถิติของตลับลูกปืน 2 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 15 กุมภาพันธ์ 2004 0:52 ถึง 15 กุมภาพันธ์ 2004 4:04.....	30
ตาราง 7 ค่าทางสถิติของตลับลูกปืน 2 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 13 กุมภาพันธ์ 2004 10:02 ถึง 13 กุมภาพันธ์ 2004 02:13.....	30
ตาราง 8 ค่าทางสถิติของตลับลูกปืน 2 ชุดข้อมูล 3 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 05 มีนาคม 2004 22:42 ถึง 05 มีนาคม 2004 21:42.....	31
ตาราง 9 ค่าทางสถิติของตลับลูกปืน 3 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 15 กุมภาพันธ์ 2004 0:52 ถึง 15 กุมภาพันธ์ 2004 4:04.....	32
ตาราง 10 ค่าทางสถิติของตลับลูกปืน 3 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 13 กุมภาพันธ์ 2004 10:02 ถึง 13 กุมภาพันธ์ 2004 02:13.....	32
ตาราง 11 ค่าทางสถิติของตลับลูกปืน 3 ชุดข้อมูล 3 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 05 มีนาคม 2004 22:42 ถึง 05 มีนาคม 2004 21:42.....	33
ตาราง 12 ค่าทางสถิติของตลับลูกปืน 4 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 15 กุมภาพันธ์ 2004 0:52 ถึง 15 กุมภาพันธ์ 2004 4:04.....	34

ตาราง 13 ค่าทางสถิติของตลับลูกปืน 4 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 13 กุมภาพันธ์ 2004 10:02 ถึง 13 กุมภาพันธ์ 2004 02:13.....	34
ตาราง 14 ค่าทางสถิติของตลับลูกปืน 4 ชุดข้อมูล 3 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 05 มีนาคม 2004 22:42 ถึง 05 มีนาคม 2004 21:42.....	35
ตาราง 15 ค่าทางสถิติของตลับลูกปืน 1 ชุดข้อมูล 2 อยู่ในคลาส 2 (Outer) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 19 กุมภาพันธ์ 2004 02:22 ถึง 19 กุมภาพันธ์ 2004 06:22.....	35
ตาราง 16 ค่าทางสถิติของตลับลูกปืน 3 ชุดข้อมูล 3 อยู่ในคลาส 2 (Outer) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 17 เมษายน 2004 22:42 ถึง 18 เมษายน 2004 2:42	36
ตาราง 17 แสดงถึงช่วงค่าของพารามิเตอร์ที่ใช้ในการปรับจูนของแต่ละโมเดล	42
ตาราง 18 การเปรียบเทียบ ค่าเฉลี่ยของค่าความถูกต้องของโมเดลจากการทำ Hyper parameter tuning และ 5-Fold Cross Validation.....	43
ตาราง 19 Classification report ของ SVM	45
ตาราง 20 Classification report KNN	46
ตาราง 21 Classification report ป่าสุ่ม	47
ตาราง 22 Classification report XG-Boost	48
ตาราง 23 เปรียบเทียบความถูกต้อง และ F1-score ของโมเดล.....	49

สารบัญรูปภาพ

	หน้า
ภาพประกอบ 1 ตำแหน่งจุดวัดสัญญาณการสั้นสะเทือน	5
ภาพประกอบ 2 โครงสร้างของตลับลูกปืน	6
ภาพประกอบ 3 การแบ่งคลาสของ KNN	8
ภาพประกอบ 4 การแบ่งกลุ่มแบบ SVM	8
ภาพประกอบ 5 โครงสร้างโมเดล Random forest.....	9
ภาพประกอบ 6 วิธีการทำงานของ XG-Boost	10
ภาพประกอบ 7 Confusion Matrix.....	11
ภาพประกอบ 8 กระบวนการทำงานแบบจำลอง.....	20
ภาพประกอบ 9 ตำแหน่งการติดตั้งตลับลูกปืน	20
ภาพประกอบ 10 ความสัมพันธ์ (Correlation) 12 คุณลักษณะเฉพาะ	26
ภาพประกอบ 11 ความสัมพันธ์ (Correlation) 8 คุณลักษณะเฉพาะ	27
ภาพประกอบ 12 เรียงลำดับคุณลักษณะเฉพาะที่มีความสำคัญของ Gradient boost Classification	36
ภาพประกอบ 13 เรียงลำดับคุณลักษณะเฉพาะที่มีความสำคัญของป่าสุ่ม	37
ภาพประกอบ 14 RMS แสดงอยู่ในรูปแบบกราฟเส้นแบ่งตามคลาส	38
ภาพประกอบ 15 Peak-to-Peak แสดงอยู่ในรูปแบบกราฟเส้นแบ่งตามคลาส	39
ภาพประกอบ 16 Clearance แสดงอยู่ในรูปแบบกราฟเส้นแบ่งตามคลาส	40
ภาพประกอบ 17 การแจกแจงแบบปกติของคุณลักษณะเฉพาะของ RMS.....	41
ภาพประกอบ 18 การแจกแจงแบบปกติของคุณลักษณะเฉพาะของ Peak to peak	41
ภาพประกอบ 19 การแจกแจงแบบปกติของคุณลักษณะเฉพาะของ Clearance.....	41
ภาพประกอบ 20 การทำ Fold Cross validation.....	43

ภาพประกอบ 21 Classification report SVM 45

ภาพประกอบ 22 Confusion matrix KNN 46

ภาพประกอบ 23 confusion matrix ป่าสุ่ม 47

ภาพประกอบ 24 confusion matrix XG-Boost..... 48

ภาพประกอบ 25 กราฟแสดงการกระจายตัว ระหว่าง Avg. RMS และ Avg. Peak-to-Peak..... 50

ภาพประกอบ 26 กราฟแสดงการกระจายตัวระหว่าง Avg. RMS และ Avg. Clearance 50

ภาพประกอบ 27 กราฟแสดงการกระจายตัวระหว่าง RMS และ Avg. Kurtosis..... 51

ภาพประกอบ 28 กราฟแสดงการกระจายตัวระหว่าง RMS และ Avg. Skew 52

ภาพประกอบ 29 กราฟแสดงการกระจายตัว ระหว่าง RMS และ Avg. Crest 52



บทที่ 1

บทนำ

1. ความสำคัญและความเป็นมาของงานวิจัย

งานซ่อมบำรุงเป็นงานประเภทการบำรุงรักษาเครื่องจักรให้เครื่องจักรเดินได้ต่อเนื่อง และป้องกันการหยุดเครื่องจักรโดยไม่วางแผนไว้หรือที่เรียกว่า เบรกดาวน์ (breakdown) ทำให้เกิดการสูญเสียเวลาในการผลิตสินค้าและผลิตภัณฑ์ โดยงานซ่อมบำรุงจะมี 3 แนวทางดังนี้

1.1 การซ่อมบำรุงรักษาเชิงรับ (Reactive Maintenance)

วิธีการซ่อมบำรุงเหมือนเครื่องจักรเกิดการชำรุด จึงดำเนินการแก้ไขจะใช้กับเครื่องจักรที่มีขนาดเล็ก ไม่สำคัญ หรือ การสูญเสียอยู่ในขอบเขตที่ยอมรับได้

1.2 การซ่อมบำรุงเชิงป้องกัน (Preventive maintenance)

วิธีการซ่อมบำรุงก่อนเครื่องจักรชำรุด แนวทางการจัดการจะมีทั้งแบบตามเวลาที่กำหนดเวลา หรือ ตามเงื่อนไข การซ่อมบำรุงตามเวลาที่กำหนดคือ การเปลี่ยนชิ้นส่วนของเครื่องจักรตามเวลาที่กำหนดไว้ ตามแผนงานของที่กำหนดเป็นมาตรฐาน ส่วนตามเงื่อนไขคือ เมื่อมีสิ่งผิดปกติเกิดขึ้นตามมาตรฐานคู่มือการซ่อมบำรุงหรือประสบการณ์ของผู้ปฏิบัติงาน ในการซ่อมบำรุงเชิงป้องกันยังมีข้อเสียที่สามารถปรับปรุงได้ เช่น เมื่อเครื่องจักรยังสามารถใช้งานได้อยู่ แต่ต้องทำการซ่อมบำรุงก่อนเพื่อลดความเสี่ยง การใช้ประสบการณ์ขึ้นอยู่กับตัวบุคคล การพบสิ่งผิดปกติแล้วจึงดำเนินการแก้ไขอาจล่าช้า เป็นต้น

1.3 การซ่อมบำรุงรักษาเชิงพยากรณ์ (Predictive maintenance)

วิธีการซ่อมบำรุงแบบแก้ไขล่วงหน้าก่อนจะเกิดความผิดปกติที่รุนแรง ทำให้ทราบล่วงหน้า มีการวางแผนงานซ่อมบำรุงมีประสิทธิภาพมากขึ้น การซ่อมบำรุงรักษาเชิงพยากรณ์เป็นการใช้เทคโนโลยีในการเข้าไปตรวจสอบและวิเคราะห์ข้อมูลสภาพปัจจุบันของเครื่องจักร รวมถึงการทำนายอนาคตจากค่าสถิติที่ช่วงบ่งบอกความผิดปกติที่เป็นปัจจุบัน โดยสร้างแบบจำลองการทำนายพร้อมกับการเรียนรู้ของเครื่อง (Machine learning) สำหรับการนำขั้นตอนวิธีที่ใช้ในการทำนาย ซึ่งจะช่วยลดต้นทุนการซ่อมบำรุงพร้อมกับเพิ่มอายุการใช้งานของเครื่องจักร หรืออุปกรณ์ช่วยในการวิเคราะห์ เพื่อลดการสูญเสียและนำข้อมูลของการเดินเครื่องมา ใช้ประโยชน์ปรับปรุง พัฒนาการควบคุมเครื่องจักร ลดต้นทุนในการผลิต และ พัฒนางานด้านการซ่อมบำรุง ในแต่ละอุตสาหกรรมจะมีกระบวนการผลิตที่แตกต่างกัน ส่วนประกอบหลักของ

เครื่องจักร จะประกอบด้วยส่วนที่เป็นเครื่องจักรกล และ แหล่งกำเนิดพลังงานในการขับเคลื่อนเครื่องจักร ส่วนที่เป็นเครื่องจักรกลมีได้หลายประเภท เช่น สายพานการผลิต , บีบ , โปพัดในถัง , เครื่องบดวัตถุดิบ , เครื่องกดวัตถุดิบ , เครื่องดูดอากาศ และ อื่นๆ ทั้งหมดนี้จะมีการเชื่อมต่อไปที่แหล่งกำเนิดพลังงานที่มีที่ผ่าน ระบบชุดขับเคลื่อน สายพาน หรือ ต่อโดยตรงกำลังแหล่งกำเนิดพลังงาน เครื่องจักรกลส่วนประกอบส่วนใหญ่เป็น วัสดุโลหะ มีการสึกหรนตามการใช้งานและเวลา ทำให้มีแรงสั่นสะเทือนเกิดขึ้น แรงสั่นสะเทือนอาจส่งผลทำให้เครื่องชำรุดและเสียหายได้ รวมไปถึงการติดตั้งเครื่องจักรอย่างไม่เหมาะสม หรือการจัดตำแหน่งเกินมาตรฐาน ส่งผลทำให้เกิดการสั่นสะเทือนทำให้เครื่องจักรชำรุด เครื่องจักรส่วนใหญ่ตัวขับเคลื่อนหรือเครื่องจักรที่มีการหมุนในแนวเส้นรอบวง จะมีตลับลูกปืนเป็นส่วนประกอบ ตลับลูกปืนเป็นอุปกรณ์สำคัญที่อยู่ในวัตถุที่มีการเคลื่อนที่ทางกลศาสตร์ ตลับลูกปืนมีชิ้นส่วนที่ใช้รองรับการเคลื่อนที่ของเพลาน้ำหนัก และถ่ายเทแรงที่เกิดขึ้นจากแกนหมุน หลักของตลับลูกปืนได้แก่ แหวนวงนอก (Outer ring) เม็ดลูกกลิ้ง (Ball or Roller) ตัวกั้นหรือรั้ง (Separator or Retainer) แหวนวงใน (Inner ring) โดยทั้งหมดนี้เป็นวัสดุที่มีส่วนประกอบของโลหะ ซึ่งเมื่อมีการใช้งานเป็นระยะเวลาอันยาวนาน วัสดุมีการสึกกร่อนไปตามการใช้งานและเวลา หากตลับลูกปืนมีการชำรุดจะส่งผลให้เครื่องจักร หรือสายการผลิตหยุดทันที คิดเป็นการสูญเสียโอกาสทางการผลิตประเภทหนึ่ง และมีโอกาสทำให้เครื่องจักรเสียหาย ในเครื่องจักรที่มีขนาดใหญ่หรือมีความสำคัญทางการผลิต จะต้องมีการตรวจจับหาสิ่งผิดปกติที่เกิดขึ้นกับตลับลูกปืน และเครื่องจักรสำคัญมีการติดอุปกรณ์เครื่องมือวัดสัญญาณการสั่นสะเทือนของเครื่องจักรมาวิเคราะห์ โดยให้ผู้ชำนาญเป็นคนตัดสินใจจากสัญญาณที่เกิดขึ้น แต่ปัจจุบันมีการนำเครื่องวัดแรงสั่นสะเทือนเพิ่มขึ้นการใช้ผู้เชี่ยวชาญวิเคราะห์สัญญาณปริมาณมากเป็นปัญหาที่เกิดขึ้น การเรียนรู้ของเครื่องมาทำนายให้รู้เรื่องล่วงหน้าจึงเป็นที่มาของงานวิจัยนี้

2. วัตถุประสงค์ของการวิจัย

1. ศึกษาการซ่อมบำรุงรักษาเชิงพยากรณ์ด้วยวิธีการเรียนรู้ของเครื่อง
2. ศึกษาการใช้การเรียนรู้ของเครื่องในการวิเคราะห์สัญญาณการสั่นสะเทือนเพื่อทำนายความผิดปกติของตลับลูกปืน
3. ศึกษาการใช้การเรียนรู้ของเครื่องในการทำนายรูปแบบการชำรุดของตลับลูกปืน

3. ขอบเขตของการวิจัย และ การดำเนินงานวิจัย

1. ในการศึกษาครั้งนี้ได้กำหนดขอบเขตศึกษาแนวทางการซ่อมบำรุงเชิงพยากรณ์กับตลับลูกปืน
 2. ใช้การเรียนรู้ของเครื่องมาทำนายหาสัญญาณผิดปกติที่ส่งผลทำให้ตลับลูกปืนชำรุด
 3. ชุดข้อมูลที่นำมาทดลองเป็นของ Intelligent Maintenance Systems เป็นสัญญาณสั่นสะเทือนในแวนอน และ แนวตั้ง
 4. คุณลักษณะที่ใช้ในงานวิจัยนี้คือ Mean, Standard derivation (Std), Skewness (Skew), Peak-to-Peak, Crest, Clearance, Shape, Impulse, Entropy, Kurtosis, Root mean square (RMS) และ Max
 5. การแบ่งคลาสของข้อมูลจะถูกกำหนด 4 คลาส คือ Normal, Inner, Outer และ Roller
 6. การเรียนรู้ของเครื่องที่ใช้ได้แก่ Support vector machine (SVM), K-Nearest Neighbor (KNN), Random Forest (RF) และ Extreme Gradient Boosting (XG-Boost)
 7. การวัดประสิทธิภาพของงานใช้ F1-score ในการตรวจวัดประสิทธิภาพของแต่ละโมเดลในรูปของค่าความแม่นยำ (Precision), ค่าความระลึก (recall), ค่าความถูกต้อง (Accuracy) และ macro average F1-score

4. ประโยชน์ที่คาดว่าจะได้รับการวิจัย

1. นำข้อมูลที่ได้ไปใช้ในการวิเคราะห์งานซ่อมบำรุง โดยการเรียนรู้ด้วยเครื่อง
2. ทำนายหาสัญญาณความผิดปกติของตลับลูกปืน เพื่อป้องกันการเกิดความเสียหายของเครื่องจักร
3. นำข้อมูลที่ได้จากการเรียนรู้ด้วยเครื่อง มาประกอบการพิจารณาในการวิเคราะห์หาสาเหตุของการชำรุดของตลับลูกปืนและเครื่องจักร

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

1.1 ทฤษฎีเกี่ยวกับการเตรียมข้อมูล (Data Preprocessing)

1.1.1 การสันสะเทือน และวิธีการวัดการสันสะเทือน

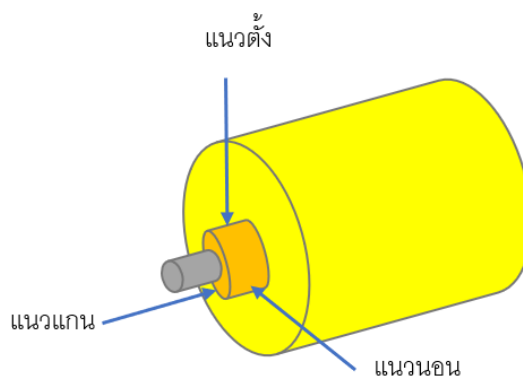
การสันสะเทือนคือ การเคลื่อนที่ของมวลทุกทิศทางในตำแหน่งอยู่กับที่ โดยเฉพาะอย่างยิ่งโครงสร้างเครื่องจักร ชิ้นส่วนอุปกรณ์ ที่ติดตั้งทำงานอยู่ในโรงงานทั่วไป เช่น การหมุนของเพลลา หรือตลับลูกปืนจะหมุนที่จุดศูนย์กลางของตลับลูกปืน ถ้าตลับลูกปืนสึกจะทำให้มีระยะห่างมาก การเคลื่อนที่ของเพลลาในทุกทิศทางก็มากเช่นกัน ทำให้เกิดการสันสะเทือน หรือ การเคลื่อนที่ที่กลับของตลับลูกปืนแทนรองรับ สาเหตุที่ทำให้เกิดการสันสะเทือน เช่น ชิ้นส่วนหมุนไม่สมดุล , การติดตั้งตลับลูกปืนรองรับเพลลาหมุน บิดงอ , เฟือง , ตลับลูกปืนสึกชำรุด , การเยื้องศูนย์กลาง , การยึดแน่นหลวมคลาย , การบิดตัวของชิ้นส่วน , ฐานรองรับบิดเอียงไม่เรียบ

การวัดความสันสะเทือนนั้นโดยปกติแล้วจะวัดใน 3 จุดคือ จุดวัดในแนวแกน (Axial, A) , จุดวัดในแนวนอน (Horizontal, H) , จุดวัดในแนวตั้ง (Vertical, V) ดังภาพประกอบ 1

1. แนวแกน (Axial, A) คือ จุดวัดที่อยู่ในแนวขนานกับเพลลาของเครื่องจักรที่ต้องการวัด ในการวัดนั้นเราต้องวัดในจุดที่อยู่ใกล้กับเพลลาให้มากที่สุด การสันสะเทือนที่เกิดขึ้นมากผิดปกติหรือเกิดขึ้นอย่างรุนแรงในแนวนี้มีสาเหตุมาจากการหมุนของเพลลาการคดงอของเพลลา และการจัดเรียงที่ส่วนต่าง ๆ ของเครื่องจักรหรือโครงสร้างไม่ตรงตำแหน่ง (Misalignment)

2. แนวนอน (Horizontal, H) คือ จุดวัดที่อยู่ในแนวนอนหรือแนวขนานกับพื้น หรือ ตั้งฉากกับจุดยึดของฐานมอเตอร์ เครื่องจักรซึ่งจะวัดทางด้านขวาหรือด้านซ้ายของชิ้นส่วน หรือเครื่องจักรที่ต้องการวัดก็ได้แต่มุมการวัดต้องได้ 90 องศา หรือใกล้เคียง กับจุดวัดที่เป็นแนวตั้ง (Vertical, V) การสันสะเทือนที่เกิดขึ้นมากผิดปกติหรือเกิดขึ้นอย่างรุนแรงในแนวนี้มีสาเหตุมาจากความสมดุลของเพลลา ไบพัด ชิ้นส่วนที่เคลื่อนที่เป็นวงกลม หรือการแกว่งของเพลลาของเครื่องจักรนั้น ๆ

3. แนวตั้งฉากหรือแนวตั้ง (Vertical, V) คือ จุดวัดที่อยู่ในแนวตั้งฉากกับพื้น หรือ ด้านที่อยู่ตรงกันข้ามกับจุดจับยึดมอเตอร์หรือเครื่องจักรโดยทั่วไป แล้วจุดที่ทำการวัดคือจุดวัดที่อยู่ด้านบนของเครื่องจักรที่ต้องการจะวัดเช่นด้านบนของมอเตอร์หรือปั๊ม โดยจุดที่วัดจะต้องมีมุมขนาน 90 องศา หรือใกล้เคียงกับจุดวัดในแนวนอน (LEGA Corporation, 2016)

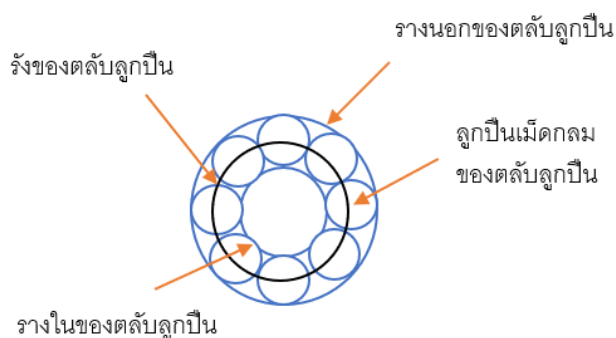


ภาพประกอบ 1 ตำแหน่งจุดวัดสัญญาณการสั่นสะเทือน

1.1.2 รูปแบบการชำรุดของตลับลูกปืน

ตลับลูกปืนประกอบไปด้วยรางด้านใน (Inner Race) , รางด้านนอก (Outer Race) และลูกปืน (Ball) อยู่ระหว่างสองแหวนในและนอก รางบนรัง (Cage) ของตลับลูกปืน ดังภาพประกอบ 2 การทำงานของตลับลูกปืนอาจจะเกิดความผิดปกติชำรุดเสียหาย อันเนื่องมาจากความล้าของตลับลูกปืน ผลที่ตามมาคือรางเกิดการเสียหาย ลูกปืนเกิดแตกสะเก็ดเล็กๆ ได้ และมันจะเป็นสาเหตุให้เกิดการสั่นและมีเสียงรบกวน หากตลับลูกปืนไม่ได้รับสารหล่อลื่นที่เพียงพอ หรือได้รับการติดตั้งอย่างไม่ถูกวิธี เกิด Misalignment ส่งผลให้เกิดการชำรุดของเครื่องจักรทั้งตัวโครงสร้างและตลับลูกปืน โดยความเสียหายที่เกิดขึ้นแบ่งออกเป็น 4 ประเภทดังนี้

1. รางนอกของตลับลูกปืนชำรุดเสียหาย (Outer Race Defects)
2. รางในของตลับลูกปืนชำรุดเสียหาย (Inner Race Defects)
3. ลูกปืนของตลับลูกปืนชำรุดเสียหาย (Ball Defects)
4. รังของตลับลูกปืนชำรุดเสียหาย (Cage Defects)



ภาพประกอบ 2 โครงสร้างของตลับลูกปืน

1.1.3 ค่าสัมประสิทธิ์ในคาบเวลา

การวัดสัญญาณตลับลูกปืนในคาบของเวลาจะมีการวัดเป็น 2 หน่วยวัด คือ Acceleration, Velocity ในการทดลองแบบจำลองเราใช้ Acceleration มาเป็นข้อมูลของการเรียนรู้ของเครื่อง Acceleration ใช้วัดความสั่นสะเทือนของเครื่องจักรที่ความถี่สูงกว่า 1,000 เฮิรตซ์ โดยส่วนมากมักจะใช้วัดและวิเคราะห์ความสั่นสะเทือนและการชำรุดเสียหายของลูกปืน

1.2 ทฤษฎีเกี่ยวกับอัลกอริทึม

1.2.1 การเรียนรู้ของเครื่อง

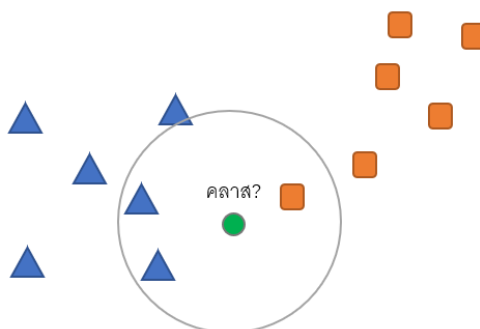
เป็นการศึกษาขั้นตอนวิธีการของที่มีการพัฒนา การเรียนรู้ของเครื่องเป็นส่วนหนึ่งของปัญญาประดิษฐ์ โดยขั้นตอนวิธีการสร้างแบบจำลองทางคณิตศาสตร์จากข้อมูลเพื่อที่จะคาดการณ์หรือตัดสินใจ การเรียนรู้ของเครื่องพัฒนาจากการศึกษาการรู้จำแบบ เกี่ยวข้องกับการศึกษาและการสร้างขั้นตอนวิธีการที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้ ขั้นตอนวิธีการนั้นจะทำงานโดยอาศัยโมเดลที่สร้างมาจากชุดข้อมูลเพื่อการทำนายหรือตัดสินใจ แทนที่จะทำงานตามลำดับของคำสั่งโปรแกรมคอมพิวเตอร์ การเรียนรู้ของเครื่องมีความเกี่ยวข้องอย่างมากกับสถิติศาสตร์ สาขาศึกษาการวิเคราะห์ข้อมูลเพื่อการทำนาย (Wikipedia., 2021b)

การเรียนรู้ของเครื่องถูกแบ่งเป็น 3 กลุ่มคือ

1. Supervised Learning คือ การเรียนรู้ของเครื่องที่ต้องมีมนุษย์มาฝึก คอมพิวเตอร์เกิดการเรียนรู้ก่อน คอมพิวเตอร์จะสามารถหาคำตอบหรือผลลัพธ์ได้ด้วยตัวเอง โดยหลักๆ การเรียนรู้ ของ Supervised Learning มี 2 ประเภทได้แก่ คือ แบบที่ 1 การสอนให้ คอมพิวเตอร์ รู้เป้าหมายของ Classification Output จะเป็น Categories และ แบบที่ 2 Regression คือการเรียนรู้ แบบตัวเลขซึ่งมีข้อมูลจำนวนมาก หาความสัมพันธ์ระหว่าง 2 ตัวแปรได้ คือ ตัวแปรหลัก และตัวแปรรองคือสิ่งที่เราพยายามจะทำนาย (Prediction)
2. Unsupervised Learning คือ เป็นการเรียนรู้แบบของขั้นตอนวิธีการแบบไม่มี ผู้สอน โดยจะไม่มีใครระบุให้คำตอบ ให้คอมพิวเตอร์หาความสัมพันธ์จากข้อมูลด้วยตนเองหรือทำ หน้าที่สร้าง Label ให้กับข้อมูล การเรียนรู้แบบไม่มีผู้สอน(Clustering) คือ การจับกลุ่มของข้อมูล เป็นกระบวนการแบ่งกลุ่มข้อมูลที่อยู่ด้วยกันเหมือนกันคล้ายกันภายในกลุ่ม
3. Reinforcement learning คือ การให้การเรียนรู้ของเครื่องเรียนรู้ โดยตัว โครงสร้าง ประกอบด้วย Agent คือ program หรือ AI ทำหน้าที่ตอบสนองกับ Environment ใน Environment มีการกำหนด state ให้ Agent ตัดสินใจ และ Action การกระทำที่สั่งให้ Agent ทำ แบบไหนควรเลือกทำอะไร Agent จะลองผิดลองถูกจนเกิดการเรียนรู้ว่าแบบไหนให้ Maximize expected discount reward ทำแบบไหนได้reward จาก Environment มากที่สุด

1.2.2 K-Nearest Neighbor (KNN)

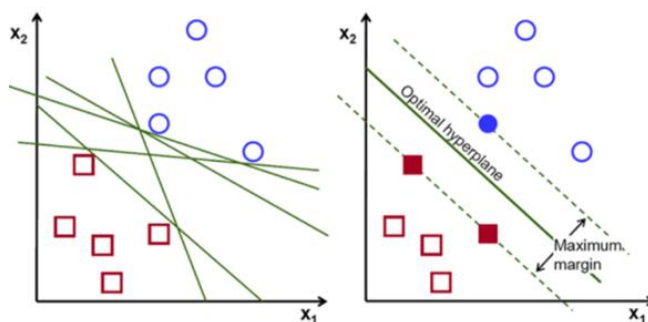
ขั้นตอนวิธีการเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor Algorithm) เป็นวิธีที่ ใช้ในการจัดแบ่งคลาส โดยเทคนิคนี้จะตัดสินใจว่า คลาสใดที่จะแทนเงื่อนไขหรือกรณีใหม่ๆ ได้ บ้าง โดยการตรวจสอบจำนวนบางจำนวน ในขั้นตอนวิธีการเพื่อนบ้านใกล้ที่สุด ของกรณีหรือ เงื่อนไขที่เหมือนกันหรือใกล้เคียงกันมากที่สุด โดยจะหาผลรวม (Count Up) ของจำนวนเงื่อนไข หรือกรณีต่างๆ สำหรับแต่ละคลาส และกำหนดเงื่อนไขใหม่ๆ ให้คลาสที่เหมือนกันกับคลาสที่ ใกล้เคียงกันมากที่สุด (นที ไทยยาธรรม, 2019) ภาพประกอบ 3 แสดงถึงวงกลมสีเขียวพิจารณาว่า ตัวเองจะอยู่ในคลาส สามเหลี่ยมสีน้ำเงินหากกำหนดระยะทางเป็นวงกลมที่บจะพบว่ามี สามเหลี่ยมสีน้ำเงินจำนวนที่มากกว่าสี่เหลี่ยมสีส้ม ในขั้นตอนนี้ใช้เพื่อพิจารณาการแบ่งกลุ่มข้อมูล ในการแบ่งกลุ่มข้อมูลในการแบ่งแยกคลาสของดัลบลูกปืนเพื่อพิจารณากลุ่มของข้อมูล



ภาพประกอบ 3 การแบ่งคลาสของ KNN

1.2.3 Support vector machine (SVM)

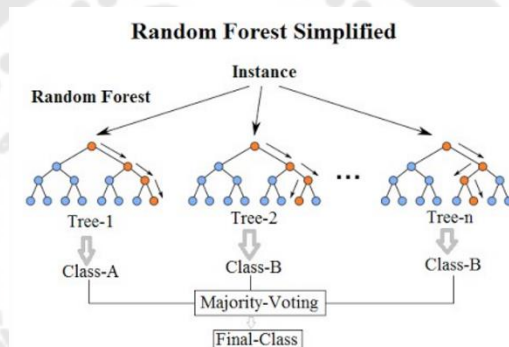
เป็นการเรียนรู้ของเครื่องกระบวนการแบ่งกลุ่มในเชิงเส้นตรง โดยมี decision boundary เป็นตัวแบ่งกลุ่ม โมเดลจะทำการเพิ่มประสิทธิภาพในการแบ่งกลุ่มให้มี margin สูงที่สุด (Maximum margin) และ Error น้อยที่สุด ภาพประกอบ 4 แสดงถึง SVM พยายามจะแบ่งกลุ่มข้อมูลโดยกำหนดเส้น Decision boundary จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า SVM เป็นวิธีการที่มีความนิยมในการนำมาใช้แบ่งแยกคลาส จากภาพประกอบ 4 จะพบว่า โมเดลจะทำการแบ่งกลุ่มข้อมูลโดยมีเส้น Optimal Hyperplane ในการแบ่งกลุ่ม และ Maximum margin คือ ระยะห่างของแต่ละกลุ่ม จะสังเกตได้ว่าถ้า Maximum margin มีค่ามากจะแบ่งจะแบ่งกลุ่มสีแดงและน้ำเงินได้ชัดเจน SVM สามารถที่จะใช้แบ่งกลุ่มข้อมูลที่ไม่เป็นเส้นตรงได้จึงมีความเป็นไปได้ที่จะมาจัดการเรื่องการทำนายคลาสของตลับลูกปืน



ภาพประกอบ 4 การแบ่งกลุ่มแบบ SVM

1.2.4 Random Forest (ป่าสุ่ม)

Random forest เป็นแบบจำลอง supervised learning เป็นหนึ่งในกลุ่มของโมเดลที่เรียกว่า Ensemble learning พัฒนามาจาก Decision tree แต่ได้เพิ่มจำนวนต้นไม้ (Tree) ในละต้นเป็นอิสระต่อกัน โดยใช้ชุดข้อมูลเรียนรู้ที่ได้จากการสุ่มตัวอย่างขึ้นมาใช้ในการทดสอบ อาจมีการสุ่มเลือกซ้ำ เมื่อได้ผลของ Class ในแต่ละต้นไม้ (Tree) ในรูปประกอบ 5 Tree-1 , Tree-2 และ Tree-3 จะได้ผลลัพธ์ออกมาว่าเป็นคลาสใด แล้วเอาการตัดสินใจของโมเดลเหล่านั้นมาโหวตกันให้ได้ผลลัพธ์ของโมเดล ปัญหาที่เกี่ยวกับ Classification วิธีการ Random Forest เป็นวิธีการหนึ่งที่ทำงานได้ดี และจากประสบการณ์ Random Forest สามารถนำไปใช้งานได้จริงกับงานประเภทอุตสาหกรรมการผลิต จึงมีความเป็นไปได้ในการนำมาประยุกต์ใช้กับการทดลอง

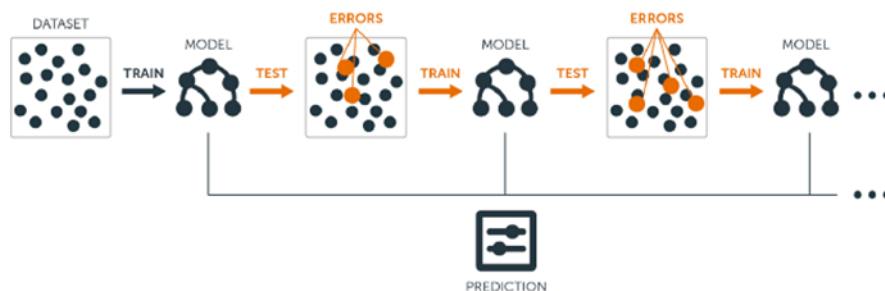


ภาพประกอบ 5 โครงสร้างโมเดล Random forest

ที่มา: (Wikipedia., 2021a)

1.2.5 Extreme Gradient Boosting (XG-Boost)

XGBOOST เป็น Ensemble learning method ที่ถูกปรับขนาดและพัฒนาจาก gradient boosting algorithm คำนวณได้เร็วขึ้น และโมเดลมีประสิทธิภาพเพิ่มขึ้น Boosting คือ การเรียนแบบเป็นลำดับ การสร้าง Independent predictor จาก multiple dataset คนละชุด ใช้ weight average ทำงานแบบ Multiple predictor Sequential fashion จากภาพประกอบ 6 วิธีการทำงานของ Boosting นำเอาผลการเรียนรู้ก่อนหน้า แล้วนำเอาผล ปรับปรุงโดยการเพิ่ม Weight ที่ทำนายผิด และลด Weight ที่ทำนายถูกส่งไปการเรียนรู้ครั้งต่อไป เพื่อลดข้อผิดพลาดจากการเรียนรู้ก่อนหน้า ส่งผลให้มีค่าความถูกต้องดีขึ้น วิธี XG-Boost ถือเป็นวิธีที่ถูกกล่าวถึงในงานวิจัย แนะนำให้ใช้กับปัญหาการแบ่งคลาส



ภาพประกอบ 6 วิธีการทำงานของ XG-Boost

ที่มา : (Wordpress., 2021b)

1.3 การวัดประสิทธิภาพของผลลัพธ์จากการวิเคราะห์ข้อมูล

Confusion Matrix ในภาพประกอบ 7 คำนวณ Accuracy, Precision, Recall, F-Score วัดประสิทธิภาพของแบบจำลองการทำนายเหตุการณ์เกิดความผิดปกติของดัลตับถูกเป็น

True Positive (TP) คือ ทำนายตรงกับสิ่งที่เกิดขึ้นจริง

True Negative (TN) คือ ทำนายตรงกับสิ่งที่เกิดขึ้นจริงคือทำนายว่าไม่จริง
สิ่งที่เกิดขึ้นไม่จริง

False Positive (FP) คือ ทำนายไม่ตรงกับสิ่งที่เกิดขึ้นจริงคือทำนายว่าจริง
แต่สิ่งที่เกิดขึ้นไม่จริง

False Negative (FN) คือ ทำนายไม่ตรงกับสิ่งที่เกิดขึ้นจริงคือทำนายว่าไม่จริง
แต่สิ่งที่เกิดขึ้นจริง

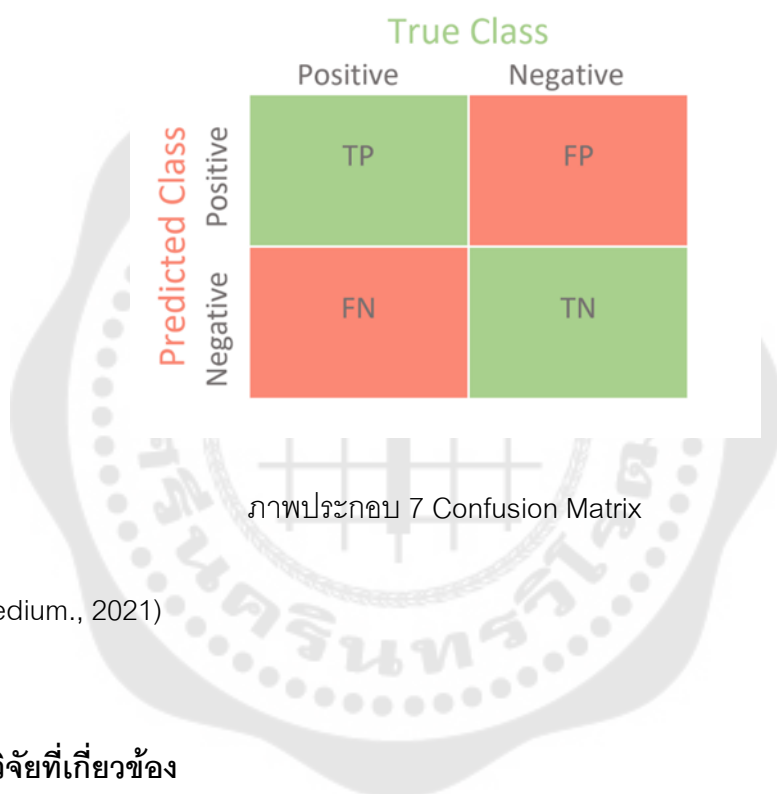
False Negative (FN) คือ ทำนายไม่ตรงกับที่ที่เกิดขึ้นจริงคือทำนายว่าไม่จริง แต่สิ่งที่
เกิดขึ้นจริง

Accuracy คือ ความถูกต้องที่เราทายได้ตรงกับสิ่งที่เกิดขึ้นจริง ดังสมการ (1)

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

F1 score คือ ค่าเฉลี่ยแบบ harmonic mean ระหว่าง precision และ recall จุดประสงค์ของการสร้าง F1 ขึ้นเพื่อเป็น Single metric ที่วัดความสามารถของโมเดล ดังสมการ (3)

$$F1 - score = \frac{2X(Precision \times Recall)}{Precision + Recall} \tag{2}$$



ที่มา: (Medium., 2021)

1.4 งานวิจัยที่เกี่ยวข้อง

(1) Fault Diagnosis from Raw Sensor Data Using Deep Neural Networks Considering Temporal Coherence (Zhang, Peng, Wu, Yao, & Guan, 2017)

บทความนี้นำเสนอการวินิจฉัยความผิดพลาดโดยใช้ Deep Neural Networks (DNN) ซึ่งเป็นโมเดลสามารถวินิจฉัยความผิดพลาดโดยใช้ข้อมูลจากเซ็นเซอร์อย่างอนุกรมเวลา โดยไม่ต้องเลือกคุณสมบัติและประมวลผลสัญญาณและมีการ temporal coherence ของข้อมูลมาใช้ประโยชน์ ซึ่งต่างการวินิจฉัยความผิดพลาดแบบเดิมที่เน้นไปที่การสกัดคุณลักษณะพร้อมกับหลักการประมวลผลสัญญาณ อย่าง time domain and frequency domain feature representation, Empirical Mode Decomposition (EMD) , Intrinsic Mode Function (IMF), Discrete Wavelet Transform (DWT) , Hilbert Huang Transform (HHT) ซึ่งในการวินิจฉัย

ข้อผิดพลาดแบบเดิมมักพบปัญหาเรื่องต้องใช้ความเชี่ยวชาญในการเลือกคุณลักษณะ และไม่มี การพิจารณา temporal coherence ของข้อมูลอนุกรมเวลา

หลักการการทำงานของโมเดลวินิจฉัยความผิดพลาดโดยอิงจาก DNN คือ ขั้นแรกนำข้อมูล อนุกรมเวลาที่เซ็นเซอร์เก็บได้ มาแยกโดยการ Data segmentation เพื่อมาใช้ Train โมเดล จนกระทั่งได้ค่า Cost Function ที่เป็นค่าต่ำสุด ขั้นตอนต่อมา นำข้อมูลทดสอบมาใช้ทดสอบความ ถูกต้องของการจำแนกประเภทของ DNN บนข้อมูลอนุกรมเวลา ขั้นตอนที่สุดท้าย วินิจฉัย ข้อผิดพลาดโดยพิจารณาถึงความสอดคล้องชั่วคราวกับข้อมูลอนุกรมเวลาเดิมร่วมด้วย

จากการทดลองพบว่าการวินิจฉัยความผิดพลาดรูปแบบใหม่โดยอิงจาก DNN สามารถจำแนก ความผิดปกติของตลับลูกปืนได้ 100% ซึ่งถือว่าวินิจฉัยความผิดพลาดโดยอิงจาก DNN มี ประสิทธิภาพในการจับความผิดปกติของตลับลูกปืน แต่จะพบว่า DNN มีความซับซ้อนในการ พิจารณาผลลัพธ์ ทำให้ยากที่จะอธิบายความสัมพันธ์ระหว่างตัวแปรและผลการทำนาย ความ ถูกต้องของโมเดล 100% มีความเป็นไปได้เมื่อนำไปใช้ทดสอบกับข้อมูลชุดอื่นจะเกิด Overfit

(2) Power Wind Mill Fault Detection via one-class -SVM Vibration Signal Analysis (Martinez-Rego, Fontenla-Romero, & Alonso-Betanzos, 2011)

บทความนี้ได้ทำการศึกษาการคาดการณ์ความผิดปกติที่เกิดกับเครื่องจักรหมุนความเร็ว สูง โดยก่อนหน้านี้ มีการศึกษาเกี่ยวกับตรวจหาข้อบกพร่องของเครื่องจักรที่หมุนด้วยความเร็วสูง อยู่หลากหลาย แต่ผลงานส่วนใหญ่ยังไม่สามารถนำมาปรับใช้ในสถานการณ์จริงได้ เนื่องจากใน การศึกษาก่อนหน้านั้นยังพบข้อจำกัดในการแยกข้อมูลระหว่างความผิดปกติและข้อมูลปกติ ดังนั้น ความสามารถและวิวัฒนาการของการตรวจจับความผิดปกติ จะช่วยให้ตรวจหาข้อบกพร่องของ เครื่องจักรมีความแม่นยำมากขึ้น เพื่อแก้ปัญหาข้อจำกัดที่กล่าวไปข้างต้น โมเดล one-class-SVM จึงถูกนำมาใช้ในการศึกษาครั้งนี้ โดยการทดสอบได้แบ่งออกเป็นสามสถานการณ์ได้แก่ สถานการณ์จำลอง สถานการณ์ที่มีการควบคุมโดยมีข้อมูลการสั่นสะเทือนจริง และสถานการณ์ จริงโดยมีข้อมูลการสั่นสะเทือนที่บันทึกได้จากเครื่องกังหันลมที่ติดตั้งในกังหันลมทางตะวันตกเฉียงเหนือของสเปน

จากการทดลองพบว่าในทั้งสามสถานการณ์ one-class-V-SVM สามารถตรวจจับความ ผิดปกติในการทำงานของกังหันลม ยิ่งไปกว่านั้นยังสามารถแสดงตัวบ่งชี้เชิงคุณภาพเกี่ยวกับ วิวัฒนาการของความผิดปกติได้อีกด้วย เมื่อเทียบกับ Novelty detection method แล้ว one-class SVM ได้รับการพิสูจน์แล้วว่าสามารถให้ผลรับที่แม่นยำมากกว่า

(3) Fault diagnosis of ball bearings using machine learning methods (Kankar, Sharma, & Harsha, 2011)

บทความนี้ได้ทำการศึกษาระบบการตรวจจับเสียหายของตลับลูกปืน และแยกประเภทความเสียหายด้วยการเรียนรู้ของเครื่อง 2 วิธีได้แก่ Artificial Neural Network (ANN) และ SVM เนื่องจากสองวิธีนี้สามารถทำงานได้กับ Non-linear classifications โดยข้อมูลการสั่นสะเทือนจะถูกเก็บรวบรวมโดยเครื่องวัดความเร่งแบบตาม time domain signals ของตลับลูกปืนที่ใช้งานได้ดี และตลับลูกปืนที่เสียหายแบบต่างๆ โดยความเสียหายได้แก่ รอยร้าวบนตลับลูกปืนรางนอก, ผิวของคัลล์ลูกปืนรางในเกิดผิวขรุขระ และการกัดกร่อนจนเกิดรูพรุนบนลูกปืน ใช้วิธีแยกการคุณลักษณะเฉพาะจาก time domain โดยวิธีการทางสถิติ คุณลักษณะเฉพาะจะถูกคัดแยกด้วย supervised attribute filter เพื่อนำมาใช้ในการเลือกคุณลักษณะ และคุณลักษณะที่ถูกเลือกด้วย Supervise learning ที่ได้นำมาใช้ในการ train โมเดลด้วย ANN และ SVM ผลการทดลองพบว่า ANN และ SVM สามารถใช้ตรวจจับความเสียหายของตลับลูกปืนอัตโนมัติได้ โดย SVM ให้ผลลัพธ์ที่แม่นยำมากกว่า นอกจากนี้เมื่อตลับลูกปืนมีความเสียหายมากกว่าหนึ่งอย่าง การเรียนรู้ของเครื่อง จะให้ผลลัพธ์ที่แม่นยำน้อยลงอาจจะเป็นเพราะข้อมูลใช้ในการ Train โมเดล เนื่องจากในการศึกษาคั้งนี้คาดว่าในสถานการณ์จริงจะมีข้อมูลที่บันทึกเกี่ยวกับความเสียหายกับตลับลูกปืนน้อย

(4) MACHINE PROGNOSTICS BASED ON HEALTH STATE ESTIMATION USING SVM (Kim, Tan, Mathew, Kim, & Choi, 2012)

บทความนี้ได้อธิบายเทคนิคการคาดการณ์อายุที่เหลืออยู่ของเครื่องจักร (remnant life of machines) ด้วย SVM และใช้ระบบ Expert knowledge ใน closed loop prognostics ในการการจำแนกข้อบกพร่อง (Fault Classification) และประเมินการเสื่อมสภาพของลูกปืนจำแนกเป็น 6 ระดับ มีการวิเคราะห์ข้อมูลย้อนหลังของการชำรุดและเหตุการณ์ของตลับลูกปืน เพื่อแยกเพื่อแยกรูปแบบความชำรุดโดยใช้ Expert knowledge system ในการแยกคุณลักษณะเฉพาะที่มีประสิทธิภาพและระบุกระบวนการของความผิดปกติ เพื่อประเมินความชำรุดที่จะเกิดขึ้นต่อไป และเพื่อที่จะทดสอบว่าโมเดลที่นำเสนอสามารถใช้งานได้จริง ได้มีการนำข้อมูลคลาสความชำรุดของตลับลูกปืนในปั๊ม High Pressure Liquefied Natural Gas (LNG) มาวิเคราะห์จนได้ระดับ Degradation stages หก ระดับ หลังจากนั้น pre-determined failure stages มาใช้ในการ

ประเมินอายุการใช้งานของเครื่องจักรที่เหลืออยู่ด้วยโมเดล SVM จากการทดลองพบว่าผลลัพธ์ที่ได้ โมเดลสามารถใช้ในคาดการณ์อายุการใช้งานของเครื่องจักรที่เหลืออยู่ได้

(5) Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati (Gousseau, Girardin, & Griffaton, 2016)

บทความนี้ได้อธิบายเทคนิค การวินิจฉัยตามการสั่นสะเทือนและการพยากรณ์ของตลับลูกปืนเม็ดกลม เพื่อสรุปให้ได้ความชัดเจนจากการวิเคราะห์สัญญาณ ใช้เทคนิคการวินิจฉัยหลายอย่าง ได้แก่ การวิเคราะห์อนุกรมเวลา , การวิเคราะห์สเปกตรัม , Blind deconvolution , สเปกตรัม coherence และ สเปกตรัม envelope มีการใช้ Pre-whitening และ Blind deconvolution ในการเตรียมข้อมูล เป็นการ Normalize โดยการใส่ค่าเฉลี่ย และกำจัดช่วงค่าที่เกิดจากการกำพอน (Resonance) ที่เกิดขึ้นจากระบบสายพาน และ การทำสเปกตรัม coherence เพื่อทำให้สัญญาณ Impulse เห็นชัดเจนขึ้น เป็นการทำให้ค่า Kurtosis มีค่าสูงสุด

จากการทดลองครั้งนี้ ทดสอบการเสื่อมสภาพตามธรรมชาติของความเสียหาย โดยไม่ใช้การเรียนรู้ของเครื่อง สร้างวิธีการที่ครอบคลุมซึ่งสามารถแยกข้อมูลการวินิจฉัยออกมาได้ สิ่งนี้น่าจะเป็นประโยชน์ที่สุดสำหรับการใช้งานให้นักวิจัยท่านอื่น สัญญาณที่บันทึกพร้อมกันโดยตัวตรวจวัดความเร่ง ค้นพบว่าไม่แนะนำให้ตรวจหาความถี่พื้นฐานของ Ball Pass Frequency Outer (BPFO) เพื่อวัตถุประสงค์ในการวินิจฉัยและการพยากรณ์ในชุดข้อมูล 1,2 พบความผิดปกติ Outer race ความถี่นี้มีตั้งแต่เริ่มการทดสอบและสอดคล้องกับ ความถี่ธรรมชาติ (Nature Frequency) ของตลับลูกปืน จากการวิเคราะห์รูปแบบสเปกตรัมของความผิดปกติภายในชุดข้อมูล 1 ส่วนใหญ่จะประกอบด้วย ฮาร์โมนิกที่ 2 และ 3 ของ BPFI (Ball Pass Frequency Inner) สามารถตรวจจับความผิดปกติที่ Inner race ในตลับลูกปืน 3 และ ความผิดปกติที่ Roller ตลับลูกปืน 4 ในส่วนของเครื่องตรวจวัดที่ 1 และ 2 มีความเป็นไปได้ที่จะตรวจพบความผิดปกติ Outer race ในตลับลูกปืน 1 จากตัวตรวจวัดความเร่ง 2 , 3 และ 4

(6) IMS Bearing Dataset_ Extracting Failure modes from vibration signals

(Miltadis, 2018)

ข้อมูลการทดลองที่เกี่ยวข้องกับตลับลูกปืน 4 ตลับลูกปืนบนเพลารับน้ำหนัก 6000 ปอนด์ หมุนด้วยความเร็วคงที่ 2,000 รอบต่อนาที การทำงานกับสัญญาณการสั่นสะเทือนแบบ ข้อมูลดิบ ไม่ใช่วิธีที่ดีที่สุดที่เราสามารถทำได้ กระบวนการชำรุดที่เกิดขึ้นและสะสมอยู่ภายใน สัญญาณที่มีไดนามิกสูง ภูมิปัญญาดั้งเดิมกำหนดให้ใช้เทคนิคการประมวลผลสัญญาณในรูปแบบคลื่นเพื่อบีบอัดวิเคราะห์และตีความข้อมูลซึ่งใช้ประสบการณ์ของคนในการวิเคราะห์ วัตถุประสงค์ในการวินิจฉัยนี้ใช้แยกข้อมูลที่เป็นประโยชน์สำหรับการวิเคราะห์และการพยากรณ์ ปกติแล้วรูปคลื่นจะถูกวิเคราะห์โดยแยกลักษณะเฉพาะ โดเมนเวลาและโดเมนความถี่ โดยข้อมูล ถูกแบ่งเป็น 4 คลาสคือ Early , Normal , Suspect และ Failure เราได้สร้างตัวแยกประเภทที่สามารถกำหนดสถานะสุขภาพของตลับลูกปืนเม็ดกลมได้เช่นเดียวกับการจดจำประเภทของความผิดปกติที่กำลังเกิดเพิ่มมากขึ้น เราได้ทดลองด้วยการแยกคุณลักษณะ ซึ่งทำให้เราต้องเลือก คุณลักษณะเฉพาะ 8 คุณลักษณะจากสัญญาณการสั่น 2 แบบ ในแนวแกน x และ y จากการ Test โมเดลได้ค่าความแม่นยำ 90% แต่คาดว่าจะเกิดข้อผิดพลาดในการทำงานมีความซับซ้อน เล็กน้อยระหว่างคลาส Early และ Normal มีความซับซ้อนเล็กน้อยระหว่าง Suspect และ Failure ความซับซ้อนส่วนใหญ่ดูเหมือนจะอยู่ในคลาส Suspect แต่นั่นเป็นสิ่งที่เข้าใจได้ เมื่อพิจารณาว่า คลาส Suspect คือการเปลี่ยนจากรูปแบบปกติไปเป็นรูปแบบความล้มเหลว

(7) Nonparametric time series modelling for industrial prognostics and health management (Mosallam, Medjaher, & Zerhouni, 2013)

วิธีการพยากรณ์และการจัดการสุขภาพ มุ่งเป้าไปที่การตรวจจับการเสื่อมสภาพ วินิจฉัย ข้อบกพร่อง และคาดการณ์เวลาที่ระบบหรือส่วนประกอบจะไม่ทำงาน โดยใช้แบบจำลองทาง กายภาพหรือการแยกข้อมูล เป็นเทรนด์เป็นหนึ่งในวิธีการที่ใช้ในการดึงข้อมูลสำคัญที่มีอยู่ใน สัญญาณ การดึงข้อมูลดังกล่าวจากข้อมูลที่รวบรวมมาในทางปฏิบัติ เป็นความท้าทายอย่างยิ่ง เนื่องจากสัญญาณมักจะมีหลายมิติและถูกบดบังด้วยสัญญาณรบกวน

แนวคิด คือการดึงพารามิเตอร์จากคุณสมบัติหลายมิติที่เป็นข้อมูลต่อเนื่องได้มาจาก สัญญาณของเครื่องจักร โดยไม่ต้องทำการสันนิษฐานใดๆ เกี่ยวกับแหล่งที่มาของสัญญาณ และ จำนวนของคุณสมบัติที่แยกออกมา สามารถดึงออกมาจากข้อมูลดิบ คุณสมบัติเหล่านี้สามารถ ได้มาจากโดเมนเวลา และโดเมนความถี่หรือโดเมนความถี่ร่วม วิธีการที่ใช้คือ Unsupervised

feature selection , Features compression , Signal trend extraction และ Health Status Modelling จากการทดลองแสดงให้เห็นสถานะความสมบูรณ์ของเครื่องจักรได้ และถูกนำมาใช้สำหรับการสร้างแบบจำลองการเสื่อมสภาพทางสุขภาพ

(8) Prediction of Bearing Remaining Useful Life With Deep Convolution Neural Network Internet of Things (IoT) (Ren, Sun, Wang, & Zhang, 2018)

ทางอุตสาหกรรม เป็นแนวทางที่ช่วยขับเคลื่อนข้อมูล แต่ความแม่นยำในการทำนายยังไม่เป็นที่น่าพอใจ บทความนี้เสนอวิธีการใหม่ในการคาดการณ์ตลับลูกปืน ของชีวิตที่เหลืออยู่ หรือ RUL (Remaining Useful Life) การนำเสนอวิธีการแยกคุณลักษณะใหม่เพื่อให้ได้เวกเตอร์ลักษณะเฉพาะ Eigenvector ซึ่งมีชื่อว่า Spectrum-principal-energy-vector โดย Eigenvector เหมาะกับ Convolutional Neural Network (CNN) ขั้นตอนการคาดการณ์เสนอวิธีการจัดการกับปัญหาความไม่ต่อเนื่อง พีเจอร์สัญญาณแบบดั้งเดิมมีการสูญเสียข้อมูลจำนวนมากเช่น คุณสมบัติโดเมนเวลา คุณสมบัติโดเมนความถี่เวลา และคุณสมบัติโดเมนความถี่ มันเป็นสิ่งสำคัญที่จะจัดการข้อมูลจาก ข้อมูลเดิมให้มากที่สุด อย่างไรก็ตาม หากใช้ตัวสัญญาณการสั่นสะเทือน 2560 มิติข้อมูล

การทำนายจากคุณลักษณะเฉพาะทั้งหมด สำหรับ Remaining Useful Life ใช้เวลาในการทำนาย เวลาที่นำมาใช้ไม่ควรยาวมาก ควรจะมีค่าน้อยกว่า 100 ในการทดลองนำเอา 2560 มิติข้อมูล คำนวณหา spectrum-principal-energy-vector โดยใช้ Fast Fourier Transform (FFT) ได้ 64 มิติข้อมูล และ เวกเตอร์ในช่วงเวลารวมเป็น 64 มิติข้อมูล แล้วจึงนำเข้าโมเดล CNN

การวิเคราะห์และการทดลองของเราแสดงความแม่นยำในการทำนายที่ดีขึ้นอย่างมาก ด้วยวิธีการแยกคุณลักษณะ Spectrum-Principal-Energy Vector สามารถแสดงข้อมูลจากข้อมูลดิบได้ดีขึ้น เวกเตอร์ลักษณะเฉพาะนี้สามารถใช้ร่วมกับการทำนายที่แตกต่างกันได้ แบบจำลองในสถานการณ์ต่างๆ สำหรับข้อมูลประเภทต่างๆ ใน ทั้งโดเมนเวลาและความถี่

(9) Supervised Health Stage Prediction Using Convolutional Neural Networks for Bearing Wear (S. Suh, Jang, Won, Jha, & Lee, 2020)

การตรวจหาข้อผิดพลาดในระบบเครื่องจักรที่หมุนได้ตั้งแต่แรกเริ่ม เป็นสิ่งสำคัญในการป้องกันความล้มเหลวของระบบเพิ่มความปลอดภัยและลดต้นทุนการบำรุงรักษา วิธีการปัจจุบันของการตรวจจับข้อบกพร่อง และขาดวิธีการแยกคุณลักษณะที่มีประสิทธิภาพ ความจำเป็นในการ

กำหนดขีดจำกัดซึ่งทำให้เกิดความผิดพลาดน้อยที่สุด อัตราการเตือนภัยและความต้องการความรู้ โดเมนของผู้เชี่ยวชาญซึ่งมีค่าใช้จ่ายสูง

สถานะสุขภาพดี (Health Stage) หรือ ช่วงปกติ มีจำนวนของข้อมูลมาก แตกต่างจากการตรวจจับข้อบกพร่องหรือการวินิจฉัยข้อบกพร่องในขณะที่ส่วนหลังมีจุดมุ่งหมายเพื่อตรวจจับลักษณะที่ปรากฏ และความรุนแรงของความผิดปกติ โดยนำข้อมูลดิบเป็น FFT แล้วแปลงเป็นข้อมูลทางรูปภาพ โดยใช้เทคนิค Correlation นำข้อมูลที่ได้เข้า CNN ผลการทดลอง การผสมผสานของ NSP ซึ่งช่วยให้เรารวมการวิเคราะห์ที่ช่วยแยกคุณลักษณะที่ผิดพลาด และ CNN-based binary regression พบว่าโมเดล CNN-HS ของเราแยกแยะอย่างมีประสิทธิภาพระหว่าง สถานะสุขภาพดีและสถานะสุขภาพไม่ดี ระบุได้ชัดเจน โดยเปลี่ยนปัญหาการตรวจจับความผิดปกติเป็น Binary regression ผลการทดลองแสดงให้เห็นว่าการรวมกันของสองวิธีนี้ตรวจพบข้อบกพร่องเร็วกว่าวิธีก่อนหน้า

(10) Generative Oversampling Method for Imbalanced Data on Bearing Fault Detection and Diagnosis (Sungho Suh, Lee, Jo, Lukowicz, & Lee, 2019)

พัฒนาการตรวจจับและวินิจฉัยข้อผิดพลาดที่ซับซ้อนด้วยข้อมูล หรือ fault detection and diagnosis(FDD) วิธีการสำหรับความผิดพลาดของตลับลูกปืนในมอเตอร์เหนี่ยวนำที่ข้อมูลสภาพความผิดปกติไม่สมดุล ตรวจจับข้อผิดพลาดของตลับลูกปืนตามโครงข่ายประสาทเทียมแบบ CNN ซึ่งสัญญาณการสั่นสะเทือนจากแท่นทดสอบจะถูกใช้เป็นอินพุตในขั้นตอนการแปลงภาพ ผลการทดลองแสดงให้เห็นว่าตัวแยกประเภทที่เสนอสำหรับ FDD ทำงานได้ดีมีความแม่นยำถึง 88% และ 99% แม้ว่าปริมาณของข้อมูลสภาวะปกติและความผิดปกติจะไม่สมดุล อัตราส่วนความไม่สมดุลแตกต่างกันตั้งแต่ 20:1 ถึง 200:1 นอกจากนี้ โมเดลยังช่วยลดระดับความไม่สมดุลของข้อมูลจากการสุ่มตัวอย่างอย่าง Over sampling ผลลัพธ์ที่ได้ปรับปรุงความแม่นยำของ FDD มากถึง 99% เมื่อเกิดความไม่สมดุลอย่างรุนแรง การทดลองใช้มอเตอร์เหนี่ยวนำสามเฟส ขนาด 3 กิโลวัตต์ ที่มีแรงดันไฟฟ้า 400 โวลต์ กระแส 6.4 แอมป์ การกำหนดค่าตามการตั้งค่าโดยเชื่อมต่อกับมอเตอร์ไปยังตัวควบคุมผ่านอินเวอร์เตอร์เพื่อควบคุมความเร็วและแรงบิด สภาพแวดล้อมการทำงานมี 25 Hz ในการทำงานและแรงบิด 10 นิวตันเมตร มอเตอร์และอินเวอร์เตอร์ถูกติดตั้งบนรางเหล็ก หลังจากได้ข้อมูลนำข้อมูลมาแปลงเป็นภาพ การแปลงภาพของข้อมูลอนุกรมเวลาที่มีความสัมพันธ์สำหรับการวิเคราะห์สเปกตรัมหลายตัวแปรและการเรียนรู้ของเครื่อง จากการสร้างรูปภาพขนาดคงที่ที่สร้างขึ้นผ่านสามขั้นตอน คือ การบีบอัดในกลุ่มที่ซ้อนกัน , การสร้างภาพ และ

สะสม การใช้ BF-NSP การแก้ปัญหา FDD โดยการวิเคราะห์ข้อมูลอนุกรมเวลากลายเป็นปัญหาในการจดจำภาพ เราใช้ตัวแยกประเภท CNN เพราะมันให้ประสิทธิภาพที่โดดเด่นในการจดจำและจำแนกรูปภาพ ประสิทธิภาพของ FDD ขึ้นอยู่กับคลื่นความถี่ที่กำหนดใน BF-NSP และสถาปัตยกรรมของแบบจำลอง CNN-based bearing fault detector (CBFD) ในที่นี้ เน้นวิธีการ Train ที่ใช้สำหรับเงื่อนไขความไม่สมดุลของข้อมูล การใช้งาน CNN ทำงานได้ดีเมื่อการกระจายคลาสมีความสมดุลโดยประมาณ ข้อมูลสภาพที่ผิดพลาดโดยทั่วไปจะมีปริมาณต่ำกว่าข้อมูลสภาวะปกติ ความไม่สมดุลดังกล่าวทำให้ความแม่นยำในการจดจำของกลุ่มที่มีจำนวนน้อยลดลง ในกรณีนี้คือเงื่อนไขที่ผิดพลาด ข้อมูล ปรากฏการณ์นี้มีความสำคัญ เนื่องจากอัตราการรับรู้ของสภาวะที่ผิดพลาดเป็นการวัดผลทางปฏิบัติที่สำคัญที่สุดของ FDD

การสุ่มตัวอย่างแบบ Over sampling ด้วย Generative adversarial networks (GANs) ช่วยพัฒนาความถูกต้อง GANs เป็นตัวแทนของคลาสของแบบจำลองกำเนิดตามสถานการณ์ ทฤษฎี Generator และ Discriminator โดย Generator สร้างภาพปลอมที่เหมือนจริง และใช้สำหรับสุ่มตัวอย่างแบบ Over sampling แต่กระบวนการนี้แสดงให้เห็นว่าไม่เสถียรอย่างยิ่งและฝึกปฏิบัติได้ยาก

ในการศึกษานี้ ได้พัฒนา วิธีการสุ่มตัวอย่างเพื่อระบุข้อมูล ปัญหาความไม่สมดุลของตลับลูกปืน ข้อบกพร่องของตลับลูกปืน ในมอเตอร์เหนี่ยวนำถูกจำแนก โดยใช้ CNN ที่ออกแบบไว้ ความแม่นยำในการจำแนกประเภท นั้นสูงกว่างานก่อนหน้า 2.4 ถึง 25% นอกจากนี้ วิธีการนี้ ยังให้ความแม่นยำประมาณ 90% แม้ว่าอัตราส่วนความไม่สมดุลจะเป็น 50:1 ความแม่นยำลดลงเหลือ 80% เมื่ออัตราส่วนความไม่สมดุลเพิ่มขึ้นเป็น 1,000:1

บทที่ 3

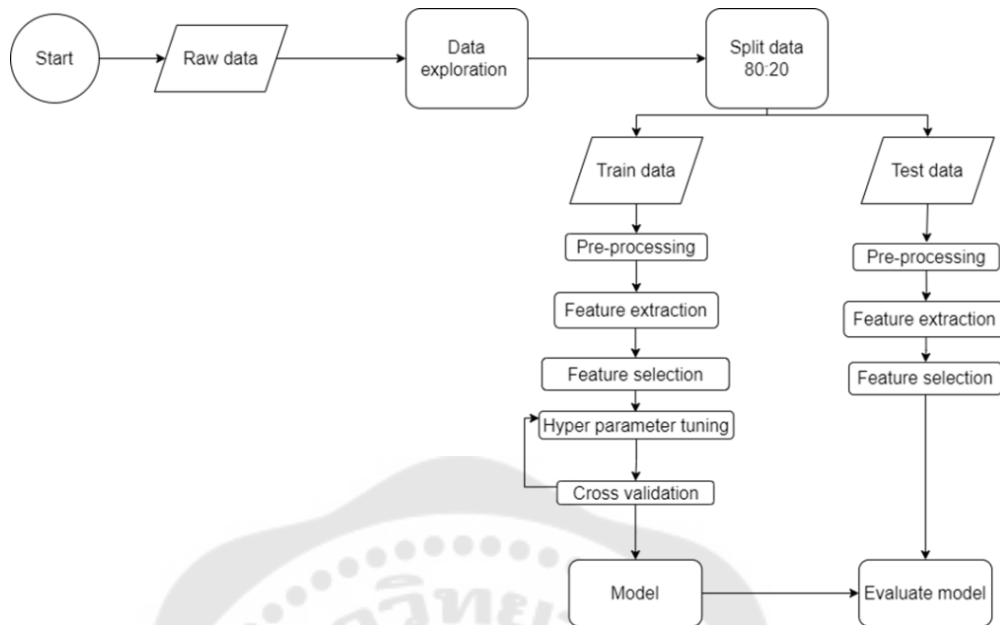
วิธีการดำเนินงานวิจัย

ในงานวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการตามขั้นตอนมีขั้นตอนดังนี้

1. กระบวนการทำงานของแบบจำลอง
2. การเก็บรวบรวมข้อมูล
3. การเตรียมข้อมูล
4. การสำรวจข้อมูล และการวิเคราะห์ข้อมูล
5. การปรับปรุงพารามิเตอร์ของโมเดล และการทำ Cross validation

1. กระบวนการทำงานของแบบจำลอง

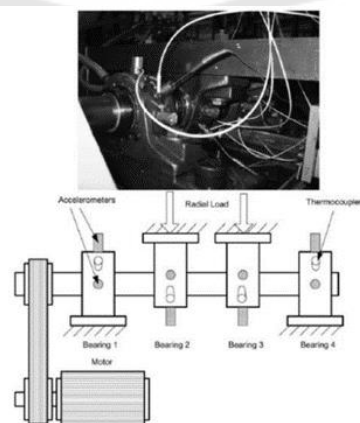
ภาพประกอบ 8 อธิบายถึงกระบวนการสร้างแบบจำลอง พิจารณาข้อมูลดิบจากการพิจารณากราฟ พร้อมกับพิจารณาเลือกตัวอย่างข้อมูล ให้ในแต่ละคลาสให้มีปริมาณข้อมูลใกล้เคียงกัน นำกลุ่มข้อมูลมาแบ่งเป็น Train 80% และ Test 20% ข้อมูล Train ถูกนำไปแยกคุณลักษณะเฉพาะ (Feature Exaction) เป็น 12 คุณลักษณะ ตามด้วยกระบวนการเตรียมข้อมูล (Preprocessing) , การเลือกคุณลักษณะเฉพาะ (Feature selection) เพื่อจัดการและหาความสัมพันธ์ระหว่างคุณลักษณะ หลังจากนั้นนำข้อมูล Train เข้าโมเดล หาค่า Hyper parameter และ ทำ Cross validation เพื่อปรับปรุงให้มีความถูกต้องเพิ่มขึ้น หลังจากนั้นนำข้อมูลที่ใช้ Test แยกคุณลักษณะเฉพาะ , กระบวนการเตรียมข้อมูล และ การเลือกคุณลักษณะเฉพาะ เพื่อให้มีจำนวนคุณลักษณะตามข้อมูล Train หลังจากนั้นนำข้อมูลที่ได้ไปเข้าโมเดลเพื่อทำการ Evaluation model



ภาพประกอบ 8 กระบวนการทำงานแบบจำลอง

2. การเก็บรวบรวมข้อมูล

ในงานวิจัยครั้งนี้ใช้ข้อมูลตลับลูกปืนของ Intelligent Maintenance Systems (IMS) เป็นข้อมูลของการทดสอบตลับลูกปืนแบบจากการทำงานปกติจนตลับลูกปืนเกิดการชำรุดเสียหาย (test-to-failure) ภาพประกอบ 9 ตลับลูกปืนถูกติดตั้งหมุนด้วยความเร็ว 2000 รอบต่อนาที มีมอเตอร์ไฟฟ้าขับผ่านสายพาน โหลดเครื่องจักรอยู่ที่ 600 ปอนด์ ได้ทำการติดตั้งเครื่องมือตรวจวัดแรงสั่นสะเทือน โดยมีทั้งหมด 4 ตลับลูกปืน



ภาพประกอบ 9 ตำแหน่งการติดตั้งตลับลูกปืน

จากตาราง 1 ข้อมูลจาก IMS มีด้วยกันทั้งหมด 3 ชุดข้อมูล โดยข้อมูลชุดที่ 1 มีจำนวน 2,516 แฟ้มข้อมูล ข้อมูลด้วยความถี่ 20 kHz ในแต่ละแฟ้มข้อมูลมีทั้งหมด 20,480 ค่า ใน 1 แฟ้มข้อมูลคือ 1 วินาที ของการวัด ในแต่ละแฟ้มมีความห่างของช่วงเวลาอยู่ที่ทุกๆ 5 นาที หรือ 10 นาที ค่าที่เก็บไว้สัญญาณตลับลูกปืนเป็นค่าการสั่นสะเทือน จากเครื่องวัดความเร่ง มีหน่วยวัดเป็น millimeters/square second (mm/s^2) ติดตั้งในแนวแกนนอนและตั้ง ทั้งหมด 8 จุดติดตั้ง จึงได้ 8 คุณลักษณะตามชุดข้อมูล 1 แต่ชุดข้อมูล 2,3 ติดตั้งเฉพาะแกนนอนจึงมีแค่ 4 คุณลักษณะ ใน 1 แฟ้มข้อมูลคือ 1 วินาที ของการวัดในแต่ละแฟ้มมีความห่างของช่วงเวลาอยู่ที่ทุกๆ 10 นาที ตลับลูกปืนถูกทดสอบจนกระทั่งตลับลูกปืนชำรุด เมื่อตลับลูกปืนชำรุดจะไม่สามารถเดินเครื่องจักรได้ โดยส่วนใหญ่ตลับลูกปืนจะเริ่มชำรุดแต่เครื่องจักรยังสามารถเดินได้ แต่ความเสียหายจะเพิ่มขึ้นเรื่อยๆ สามารถวิเคราะห์ข้อมูลได้จากเครื่องวัดความเร่ง ข้อมูลจาก IMS ข้อมูลชุด 1 ระบุการชำรุดเป็น Inner race คือเกิดการชำรุดที่ตลับลูกปืนบริเวณรางใน และการชำรุดเป็น Rolling element คือเกิดการชำรุดที่ตลับลูกปืนที่ลูกปืนเม็ดกลม ข้อมูลชุด 2 ระบุการชำรุดเป็น Outer race คือ เกิดการชำรุดที่ตลับลูกปืนบริเวณรางนอก และข้อมูลชุด 3 ระบุการชำรุดเป็น Outer race คือ เกิดการชำรุดที่ตลับลูกปืนบริเวณรางนอก สาเหตุของการชำรุดที่ระบุในเอกสาร IMS นำไปใช้ระบุคลาสของกระบวนการเรียนรู้ของเครื่องได้ในการทดลอง การวัดสัญญาณสั่นสะเทือนจะติดตามตลับลูกปืนหากเป็นข้อมูลชุดเดียวกันจะวัดในช่วงเวลาเดียวกัน เวลาที่บันทึกคือเวลาตั้งแต่เริ่มวัดจนไปถึงช่วงเวลาที่ชำรุด

ตาราง 1 รายละเอียดข้อมูล IMS

ชุดข้อมูล	จำนวนแฟ้มข้อมูล	จำนวนคุณลักษณะ	เวลายับบันทึกทั้งหมด	สาเหตุของการชำรุดเสียหาย
ชุดข้อมูล 1	2,156	8	49,680 นาที	ตลับลูกปืน 3: Inner race ตลับลูกปืน 4: Rolling element
ชุดข้อมูล 2	984	4	9,840 นาที	ตลับลูกปืน 1: Outer race
ชุดข้อมูล 3	4,448	4	44,480 นาที	ตลับลูกปืน 3: Outer race

ที่มา: (J. Lee, Yu, Lin, IMS, & Cincinatti., 2007)

3. การเตรียมข้อมูล

จากชุดข้อมูลดิบ นำข้อมูลของค่าการสั่นสะเทือนมาจัดกลุ่มให้แต่ละคลาสมีปริมาณข้อมูลใกล้เคียงกัน และจำนวนเวลาที่นำมาใช้เท่ากัน โดยข้อมูลทั้งหมดมี 13,600 แกว ทำการแบ่งข้อมูลเป็น Train 80% มี 10,880 แกว นำไปเข้าโมเดลการเรียนรู้ของเครื่องเพื่อสอน และ Test 20% มี 2,720 แกว เพื่อทดสอบการทำงานให้แน่ใจว่า โมเดลจะมีค่าความถูกต้อง และน่าเชื่อถือในการนำไปใช้งาน คลาสที่กำหนดจะมีคลาสที่อยู่ในช่วงตลับลูกปืนทำงานปกติคลาส 1 (Normal) มี 2,726 และ ค่าสำหรับ Train 674 ค่า ในส่วนของคลาสที่เหลือ เป็นคลาสที่ชำรุด คือ คลาส 0 (Inner) , คลาส 2 (Outer) และ คลาส 3 (Roller) ในสำหรับ Test โดยมีการแบ่งข้อมูลตามตาราง 2 ได้ดังนี้

ตาราง 2 จำนวนข้อมูลสำหรับการเรียนรู้ของเครื่อง

คลาส	Train	Test
คลาส 0 (Inner)	2713	687
คลาส 1 (Normal)	2726	674
คลาส 2 (Outer)	2716	684
คลาส 3 (Roller)	2725	685

การนำสัญญาณสั่นสะเทือน ในรูปแบบของคาบเวลา เป็นคุณสมบัติทางสถิติได้ 12 คุณลักษณะเฉพาะคือ Mean , Standard derivation (Std) , Skewness , Peak-to-Peak , Crest , Clearance, Shape , Impulse, Entropy , Kurtosis , Root mean square (RMS) และ Max เพื่อให้สัญญาณ แบ่งออกเป็นหลายส่วนได้ การแบ่งสัญญาณออกเป็นหลายคุณลักษณะ ทำให้การทำนายมีความถูกต้องเพิ่มขึ้นและช่วยแยกความแตกต่างระหว่างตลับลูกปืนที่แข็งแรง และชำรุดได้ชัดเจนขึ้น การเปลี่ยนแปลงในคุณสมบัติเหล่านี้สามารถบ่งบอกถึงการเปลี่ยนแปลงในสถานะความสมบูรณ์ของระบบ เพื่อเป็นคุณลักษณะสำหรับการวินิจฉัยโดยค่าที่นำมาใช้ได้แก่ เป็น 12 คุณลักษณะเฉพาะ ดังนี้

Mean (\bar{x}) คือ ค่าเฉลี่ยสัมบูรณ์ของทั้งหมด โดยที่ x_i คือ ค่าของสมาชิกของชุดข้อมูล และ N คือ จำนวนสมาชิกของชุดข้อมูล ดังสมการ (3)

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N |x_i| \quad (3)$$

Std (σ) คือ ส่วนเบี่ยงเบนมาตรฐานของข้อมูล โดยที่ x_i คือ ค่าของสมาชิกของชุดข้อมูล และ N คือ จำนวนสมาชิกของชุดข้อมูล ดังสมการ (4)

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (4)$$

Skewness (Sk) คือ ความไม่สมมาตรของการกระจายสัญญาณ ความผิดปกติอาจส่งผลกระทบต่อการกระจายตัวของข้อมูล ทำให้ระดับของ Skewness เพิ่มขึ้น โดยที่ x_i คือ ค่าของสมาชิกของชุดข้อมูล และ N คือ จำนวนสมาชิกของชุดข้อมูล ดังสมการ (5)

$$Sk = \frac{1}{N} \sum_{i=1}^N \frac{(x_i - \bar{x})^3}{\sigma^3} \quad (5)$$

Kurtosis (K) คือ ความยาวของปลายการกระจายตัวของข้อมูล บ่งบอกถึงปริมาณของ Outlier หากมีจำนวนค่าความผิดปกติเพิ่มมากขึ้นจะทำให้ค่า Kurtosis สูงขึ้น โดยที่ x_i คือ ค่าของสมาชิกของชุดข้อมูล และ N คือ จำนวนสมาชิกของชุดข้อมูล ดังสมการ (6)

$$K = \frac{1}{N} \sum_{i=1}^N \frac{(x_i - \bar{x})^4}{\sigma^4} \quad (6)$$

Entropy (H) คือ เป็นตัวแปรสุ่มระดับค่าเฉลี่ยของข้อมูล เป็นค่าความไม่แน่นอนที่อยู่ใ ผลลัพธ์ที่เป็นไปของตัวแปร โดยกำหนด $P(x_i)$ เป็นฟังก์ชันความน่าจะเป็น และ N คือ จำนวนสมาชิกของชุดข้อมูล ดังสมการ (7)

$$H = -\sum_{i=1}^N P(x_i) \log P(x_i) \quad (7)$$

RMS (x_{rms}) คือ รากที่สองของค่าเฉลี่ยกำลังสองใช้ในการเปรียบเทียบระดับการสั่นสะเทือนของสัญญาณ RMS มีแนวโน้มที่จะสัมพันธ์กับพลังงานการสั่นสะเทือน โดยที่ x_i คือ ค่าของสมาชิกของชุดข้อมูล และ N คือ จำนวนสมาชิกของชุดข้อมูล ดังสมการ (7)

$$x_{rms} = \sqrt{\left(\frac{1}{N}\right) \sum_{i=1}^N (x_i)^2} \quad (8)$$

ที่มา: (Hanly, 2021)

Max คือ ค่าสูงสุดของข้อมูล Max value

Peak-to-Peak (x_p) คือ ช่วงของข้อมูลจากการแกว่งตัวของข้อมูลโดยนำค่า สูงสุดของข้อมูลจนถึงค่าต่ำสุดบ่งบอกถึงแรงสั่นสะเทือน ดังสมการ (9)

$$x_p = \text{Max value} - \text{Min value} \quad (9)$$

Crest Factor (x_{crest}) คือ ค่าสูงสุดหารด้วย RMS ค่าความผิดปกติจะทำให้เกิดการเปลี่ยนแปลงจะพบได้จากค่าสูงสุดของสัญญาณ ค่า Crest Factor สามารถให้การเตือนล่วงหน้าสำหรับข้อมูลผิดพลาด ดังสมการ (10)

$$x_{crest} = \frac{\text{Max value}}{x_{rms}} \quad (10)$$

Clearance Factor (x_{crest}) คือ ค่าพิศหารด้วยค่าเฉลี่ยกำลังสองของสแควร์รูทของแอมพลิจูดสัมบูรณ์ เป็นค่าที่วัดเครื่องจักรที่หมุนเป็นเส้นรอบวงตลับลูกปืนใหม่จะมีค่า Clearance Factor ที่สูงและจะลดลงเมื่อตลับลูกปืนเริ่มชำรุดค่า Clearance Factor จะลดลงที่ Roller, Outer race และ Inner race ของตลับลูกปืนตามลำดับ Clearance Factor จะเป็นค่าที่ใช้แยกความผิดปกติที่ Inner race ได้ดีโดยที่ x_i คือ ค่าของสมาชิกของชุดข้อมูล และ N คือ จำนวนสมาชิกของชุดข้อมูล ดังสมการ (11)

$$x_{clear} = \frac{x_p}{\left(\frac{1}{N} \sum_{i=1}^N 2\sqrt{|x_i|}\right)^2} \quad (11)$$

Shape Factor (Xsf) คือ RMSหารด้วยค่าเฉลี่ยของค่าสัมบูรณ์ ขึ้นอยู่กับรูปร่างของสัญญาณ ไม่สนใจขนาดของสัญญาณ ดังสมการ (12)

$$Xsf = \frac{x_{rms}}{\bar{x}} \quad (12)$$

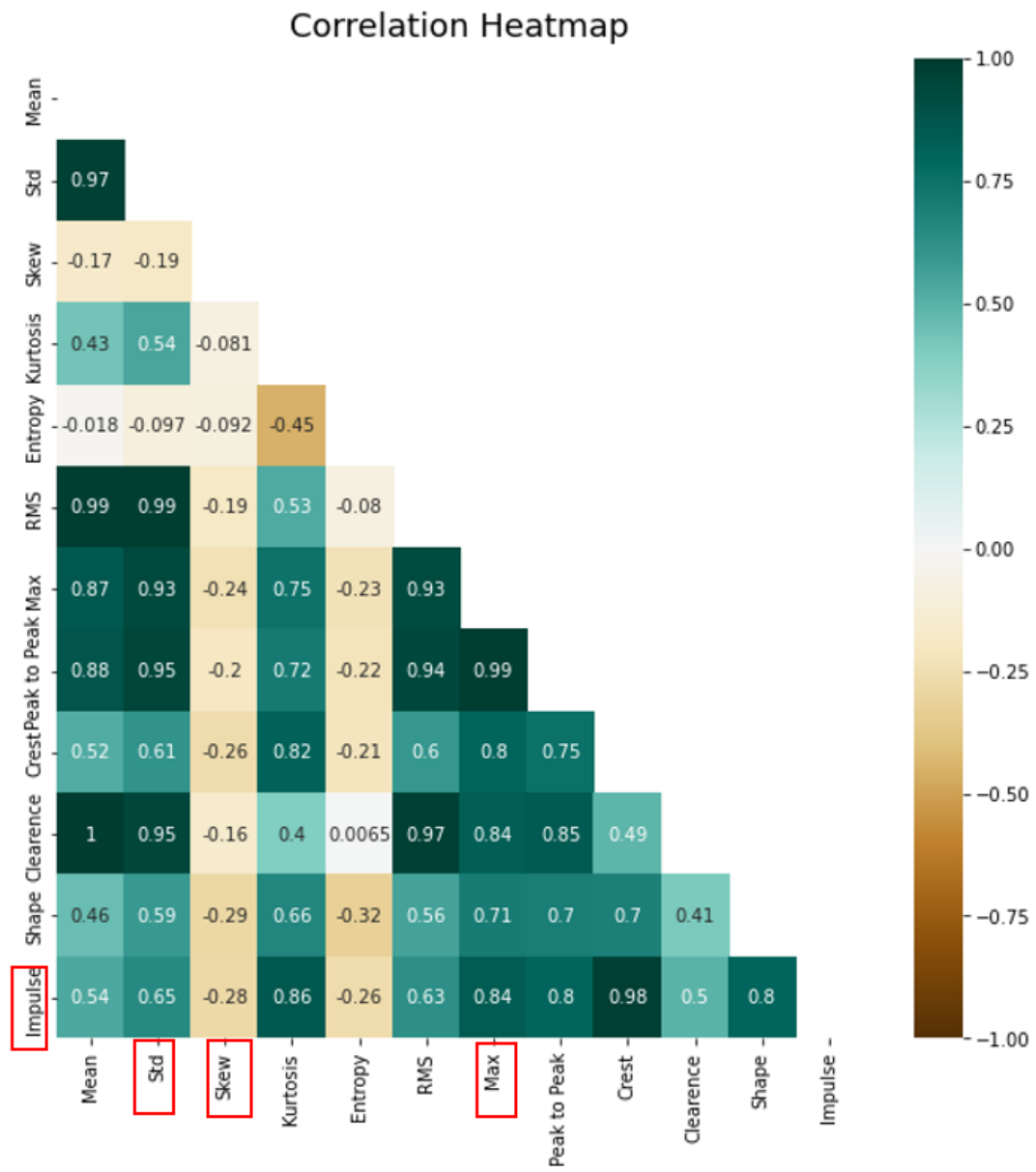
Impulse ($Ximpulse$) คือ การเปรียบเทียบความสูงของจุดสูงสุดกับระดับกลางของสัญญาณ ดังสมการ (13)

$$Ximpulse = \frac{Max\ value}{\bar{x}} \quad (13)$$

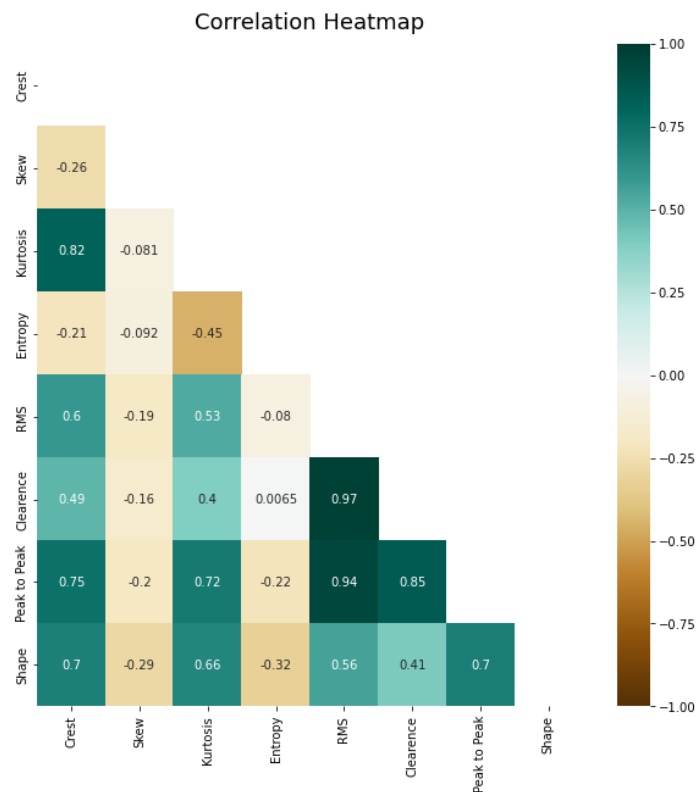
ที่มา:(MathWorks, 2021)

3.1 การหาความสัมพันธ์ระหว่างคุณลักษณะเฉพาะ

ภาพประกอบ 10 เป็นการหาความสัมพันธ์ระหว่างข้อมูล (Correlation) เพื่อลดความจำวนคุณลักษณะเฉพาะที่มีความคล้ายคลึงกัน ลดการทำงานของโมเดล จากรูปภาพจะพบว่า RMS สัมพันธ์กับ Mean และ Std. ถึง 99% ในส่วนของ Crest สัมพันธ์กับ Impulse ถึง 98% ทำการตัด Impulse เนื่องจากสูตรการคำนวณของ Crest มีค่า RMS เป็นส่วนประกอบ ทางทฤษฎีการวิเคราะห์ที่ลดรูปขึ้นให้ความสำคัญกับ RMS มากกว่า Mean และอีกหนึ่งความสัมพันธ์ Peak-to-Peak สัมพันธ์กับ Max ถึง 99 % ทำการตัด Max เนื่องจาก Peak-to-Peak ให้ผลลัพธ์ทางการทำนายได้ดีกว่า ดังนั้นจึงได้ตัดค่า Mean , Std , Impulse และ Max ออก ทำให้ตัวแปรจาก 12 คุณลักษณะเฉพาะ เหลือ 8 คุณลักษณะเฉพาะ ดังภาพประกอบ 11 โดยมีคุณลักษณะเฉพาะ คือ Skew , Crest , Kurtosis , Entropy , Shape , RMS , Clearance และ Peak-to-Peak การลดจำวนคุณลักษณะช่วยให้ลดความซับซ้อน ของการทำงานของโมเดล แต่พิจารณาไม่ให้คุณลักษณะที่ลดลงส่งผลกระทบต่อค่าความถูกต้องของโมเดล



ภาพประกอบ 10 ความสัมพันธ์ (Correlation) 12 คุณลักษณะเฉพาะ



ภาพประกอบ 11 ความสัมพันธ์ (Correlation) 8 คุณลักษณะเฉพาะ

3.2 การหาความสัมพันธ์ระหว่างชุดข้อมูลเดียวกันและต่างชุดข้อมูล

จากการสำรวจข้อมูลโดยการเปรียบเทียบคลาสของข้อมูลระหว่าง ข้อมูลที่นำไปเข้าการเรียนรู้ของเครื่อง และ ชุดข้อมูลอื่นที่ไม่ใช่ช่วงข้อมูลที่นำไปเข้าการเรียนรู้ของเครื่อง โดยทำการพิจารณาคลาส 1 (Normal) และ คลาส 2 (Outer) เป็นตัวแทนของคลาสชำรุด ในการเปรียบเทียบ ได้ทำการเปรียบเทียบในข้อมูลชุดเดียวกันในแต่ละตลับลูกปืนเวลาที่แตกต่างกัน และ เปรียบเทียบตลับลูกปืนตำแหน่งเดียวกันที่ชุดข้อมูลต่างกัน เพื่อดูค่าทางสถิติมีความแตกต่างหรือคล้ายกันเพื่อเป็นแนวทางในการพัฒนาโมเดลและการนำไปใช้กับชุดข้อมูลอื่นได้

เปรียบเทียบในข้อมูลชุดเดียวกัน จากตาราง 3 แสดงถึงค่าทางสถิติคุณลักษณะเฉพาะทั้ง 8 คุณลักษณะ ในช่วงที่เป็นคลาส 1 (Normal) ของชุดข้อมูล 2 ตลับลูกปืน 1 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และค่าเฉลี่ยเท่ากับ 0.077 เปรียบเทียบกับตาราง 4 ชุดข้อมูล 2 ตลับลูกปืน 1 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 เหมือนกัน และค่าเฉลี่ยเท่ากับ 0.076 ค่าใกล้เคียงกันและคำนวณหา T-test ของ 8 คุณลักษณะพบว่ามีค่าน้อยกว่า 0.05 สามารถยอมรับสมมุติฐานว่าทั้ง 2 ข้อมูลมีความใกล้เคียงกัน

เปรียบเทียบตลับลูกปืนตำแหน่งเดียวกันที่ชุดข้อมูลต่างกัน ในช่วงที่เป็นคลาส 1 (Normal) จากตาราง 3 คุณลักษณะของตลับลูกปืน 1 ชุดข้อมูล 2 ตลับลูกปืน 1 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และ ค่าเฉลี่ยเท่ากับ 0.077 เปรียบเทียบกับตาราง 5 ชุดข้อมูล 3 ตลับลูกปืน 1 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 เหมือนกันและค่าเฉลี่ยเท่ากับ 0.077 เท่ากัน และคำนวณหา T-test ของ 8 คุณลักษณะพบว่ามีย่าน้อยกว่า 0.05 สามารถยอมรับสมมุติฐานว่าทั้ง 2 ชุดข้อมูลมีความใกล้เคียงกัน

ตาราง 3 ค่าทางสถิติของตลับลูกปืน 1 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 15 กุมภาพันธ์ 2004 0:52 ถึง 15 กุมภาพันธ์ 2004 4:04

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.006	0.209	3.432	0.077	0.416	3.006	0.051	1.260
ค่ากลาง	0.006	0.079	3.444	0.076	0.406	2.934	0.051	1.257
ค่าแปรปรวน	0.066	0.395	0.012	0.000	0.005	0.195	0.000	0.001
ค่าต่ำสุด	-1.01	-0.860	2.816	0.051	0.256	2.006	0.036	1.157
ค่าสูงสุด	1.215	6.923	3.700	0.123	0.862	5.407	0.078	1.439

ตาราง 4 ค่าทางสถิติของตลับลูกปืน 1 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 13 กุมภาพันธ์ 2004 10:02 ถึง 13 กุมภาพันธ์ 2004 02:13

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	-0.004	0.252	3.423	0.076	0.418	3.040	0.051	1.261
ค่ากลาง	-0.005	0.112	3.432	0.076	0.409	2.973	0.051	1.258
ค่าแปรปรวน	0.071	0.471	0.012	0.000	0.006	0.218	0.000	0.001
ค่าต่ำสุด	-1.245	-0.977	2.856	0.053	0.240	2.072	0.033	1.162
ค่าสูงสุด	1.118	7.526	3.686	0.125	0.855	5.721	0.082	1.430

ตาราง 5 ค่าทางสถิติของตลับลูกปืน 1 ชุดข้อมูล 3 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 05 มีนาคม 2004 22:42 ถึง 05 มีนาคม 2004 21:42

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.008	0.298	3.423	0.077	0.419	3.044	0.051	1.266
ค่ากลาง	0.006	0.139	3.434	0.075	0.403	2.961	0.050	1.262
ค่าแปรปรวน	0.083	0.575	0.013	0.000	0.008	0.221	0.000	0.001
ค่าต่ำสุด	-1.151	-0.922	2.546	0.047	0.217	2.123	0.031	1.164
ค่าสูงสุด	1.730	14.276	3.746	0.130	1.235	5.591	0.082	1.612

เปรียบเทียบในข้อมูลชุดเดียวกันในช่วงที่เป็นคลาส 1 (Normal) จากตาราง 6 แสดงถึงค่าทางสถิติของคุณลักษณะเฉพาะทั้ง 8 คุณลักษณะของชุดข้อมูล 2 ตลับลูกปืน 2 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และค่าเฉลี่ยเท่ากับ 0.093 เปรียบเทียบกับตาราง 7 ชุดข้อมูล 2 ตลับลูกปืน 2 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 เหมือนกัน และค่าเฉลี่ยเท่ากับ 0.095 ค่าใกล้เคียงกัน ทำการคำนวณหา T-test ของ 8 คุณลักษณะพบว่ามิต้าน้อยกว่า 0.05 สามารถยอมรับสมมติฐานว่าทั้ง 2 ข้อมูลมีความใกล้เคียงกัน

เปรียบเทียบตลับลูกปืนตำแหน่งเดียวกันที่ชุดข้อมูลต่างกันในช่วงที่เป็นคลาส 1 (Normal) จากตาราง 6 ชุดข้อมูล 2 ตลับลูกปืน 2 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และค่าเฉลี่ยเท่ากับ 0.093 เปรียบเทียบกับตาราง 8 ชุดข้อมูล 3 ตลับลูกปืน 2 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 เหมือนกัน และค่าเฉลี่ยเท่ากับ 0.086 ค่าแตกต่างกันเล็กน้อย หากพิจารณาค่า Peak-to-Peak , skew และ Kurtosis มีค่าแตกต่างกัน และคำนวณหา T-test ของ 8 คุณลักษณะพบว่ามิต้าน้อยกว่า 0.05 สามารถยอมรับสมมติฐานว่าทั้ง 2 ข้อมูลมีความใกล้เคียงกัน

ตาราง 6 ค่าทางสถิติของตลับลูกปืน 2 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 15 กุมภาพันธ์ 2004 0:52 ถึง 15 กุมภาพันธ์ 2004 4:04

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.104	0.061	3.459	0.093	0.494	2.927	0.063	1.253
ค่ากลาง	0.105	-0.020	3.466	0.092	0.486	2.863	0.062	1.251
ค่าแปรปรวน	0.051	0.227	0.009	0.000	0.005	0.154	0.000	0.001
ค่าต่ำสุด	-0.759	-0.885	3.061	0.068	0.298	2.109	0.043	1.169
ค่าสูงสุด	1.065	3.812	3.728	0.129	0.962	4.891	0.088	1.372

ตาราง 7 ค่าทางสถิติของตลับลูกปืน 2 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 13 กุมภาพันธ์ 2004 10:02 ถึง 13 กุมภาพันธ์ 2004 02:13

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.101	0.070	3.455	0.095	0.507	2.946	0.064	1.252
ค่ากลาง	0.096	-0.013	3.463	0.094	0.498	2.876	0.064	1.250
ค่าแปรปรวน	0.050	0.257	0.009	0.000	0.006	0.165	0.000	0.001
ค่าต่ำสุด	-0.762	-0.977	2.956	0.068	0.332	2.027	0.045	1.153
ค่าสูงสุด	1.235	5.793	3.708	0.129	1.011	5.679	0.095	1.365

ตาราง 8 ค่าทางสถิติของตลับลูกปืน 2 ชุดข้อมูล 3 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 05 มีนาคม 2004 22:42 ถึง 05 มีนาคม 2004 21:42

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.134	0.215	3.434	0.086	0.466	3.036	0.058	1.258
ค่ากลาง	0.129	0.073	3.446	0.085	0.454	2.949	0.057	1.255
ค่าแปรปรวน	0.068	0.534	0.012	0.000	0.007	0.240	0.000	0.001
ค่าต่ำสุด	-1.543	-0.884	2.920	0.055	0.254	2.030	0.038	1.172
ค่าสูงสุด	2.149	13.574	3.690	0.127	0.979	6.866	0.088	1.413

เปรียบเทียบในข้อมูลชุดเดียวกันในช่วงที่เป็นคลาส 1 (Normal) จากตาราง 9 แสดงถึงค่าทางสถิติของคุณลักษณะเฉพาะทั้ง 8 คุณลักษณะของชุดข้อมูล 2 ตลับลูกปืน 3 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และค่าเฉลี่ยเท่ากับ 0.099 เปรียบเทียบกับตาราง 10 ชุดข้อมูล 2 ตลับลูกปืน 3 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 เหมือนกัน และค่าเฉลี่ยเท่ากับ 0.102 ค่าใกล้เคียงกัน ทำการคำนวณหา T-test ของ 8 คุณลักษณะพบว่ามีความน้อยกว่า 0.05 สามารถยอมรับสมมติฐานว่าทั้ง 2 ข้อมูลมีความใกล้เคียงกัน

เปรียบเทียบตลับลูกปืนตำแหน่งเดียวกันที่ชุดข้อมูลต่างกันในช่วงที่เป็นคลาส 1 (Normal) จากตาราง 9 คุณลักษณะของชุดข้อมูล 2 ตลับลูกปืน 3 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และค่าเฉลี่ยเท่ากับ 0.099 เปรียบเทียบกับตาราง 11 ชุดข้อมูล 3 ตลับลูกปืน 3 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 เหมือนกัน และค่าเฉลี่ยเท่ากับ 0.066 ค่าแตกต่างกันมาก หากพิจารณาค่า Peak-to-Peak ,skew และ Kurtosis มีค่าแตกต่างกัน ทำการคำนวณหา T-test ของ 8 คุณลักษณะพบว่ามีความน้อยกว่า 0.05 สามารถยอมรับสมมติฐานว่าทั้ง 2 ข้อมูลมีความใกล้เคียงกัน

ตาราง 9 ค่าทางสถิติของตลับลูกปืน 3 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 15 กุมภาพันธ์ 2004 0:52 ถึง 15 กุมภาพันธ์ 2004 4:04

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.037	0.920	3.352	0.099	0.596	3.357	0.064	1.283
ค่ากลาง	0.027	0.385	3.385	0.097	0.557	3.153	0.064	1.275
ค่าแปรปรวน	0.144	3.130	0.027	0.000	0.025	0.555	0.000	0.002
00	-2.521	-0.863	2.554	0.069	0.342	2.122	0.044	1.164
ค่าสูงสุด	2.795	22.358	3.723	0.156	1.587	7.780	0.094	1.560

ตาราง 10 ค่าทางสถิติของตลับลูกปืน 3 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 13 กุมภาพันธ์ 2004 10:02 ถึง 13 กุมภาพันธ์ 2004 02:13

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.011	0.964	3.347	0.102	0.618	3.388	0.066	1.285
ค่ากลาง	0.013	0.450	3.375	0.101	0.583	3.213	0.066	1.278
ค่าแปรปรวน	0.158	3.415	0.027	0.000	0.025	0.558	0.000	0.002
ค่าต่ำสุด	-2.679	-0.881	2.436	0.071	0.334	2.166	0.046	1.176
Maximum	4.197	35.634	3.720	0.156	1.491	8.381	0.095	1.701

ตาราง 11 ค่าทางสถิติของตลับลูกปืน 3 ชุดข้อมูล 3 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 05 มีนาคม 2004 22:42 ถึง 05 มีนาคม 2004 21:42

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.013	0.101	3.445	0.066	0.353	2.948	0.044	1.255
ค่ากลาง	0.011	0.015	3.452	0.066	0.345	2.885	0.044	1.254
ค่าแปรปรวน	0.056	0.291	0.010	0.000	0.003	0.170	0.000	0.001
ค่าต่ำสุด	-0.928	-0.981	2.930	0.045	0.208	2.031	0.030	1.162
ค่าสูงสุด	1.411	7.780	3.759	0.100	0.764	5.982	0.071	1.418

เปรียบเทียบในข้อมูลชุดเดียวกันในช่วงที่เป็น คลาส 1 (Normal) จากตาราง 12 แสดงถึงค่าทางสถิติของคุณลักษณะเฉพาะทั้ง 8 คุณลักษณะของชุดข้อมูล 2 ตลับลูกปืน 4 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และค่าเฉลี่ยเท่ากับ 0.055 เปรียบเทียบกับ ตาราง 13 ชุดข้อมูล 2 ตลับลูกปืน 4 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และค่าเฉลี่ยเท่ากับ 0.055 เหมือนกัน คำนวณหา T-test ของ 8 คุณลักษณะพบว่ามีความน้อยกว่า 0.05 สามารถยอมรับสมมติฐานว่าทั้ง 2 ข้อมูลมีความใกล้เคียงกัน

เปรียบเทียบตลับลูกปืนตำแหน่งเดียวกันที่ชุดข้อมูลต่างกันในช่วงที่เป็น คลาส 1 (Normal) จากตาราง 12 คุณลักษณะของชุดข้อมูล 2 ตลับลูกปืน 4 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 และ ค่าเฉลี่ย 0.055 เปรียบเทียบกับตาราง 14 ชุดข้อมูล 3 ตลับลูกปืน 4 คุณลักษณะเฉพาะ RMS มีค่าแปรปรวนเท่ากับ 0 เหมือนกัน และค่าเฉลี่ยเท่ากับ 0.052 ค่าใกล้เคียงกัน และคำนวณหา T-test ของ 8 คุณลักษณะพบว่ามีความน้อยกว่า 0.05 สามารถยอมรับสมมติฐานว่าทั้ง 2 ข้อมูลมีความใกล้เคียง

ตาราง 12 ค่าทางสถิติของตลับลูกปืน 4 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 15 กุมภาพันธ์ 2004 0:52 ถึง 15 กุมภาพันธ์ 2004 4:04

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	-0.007	0.000	3.463	0.055	0.288	2.873	0.037	1.252
ค่ากลาง	-0.010	-0.071	3.467	0.054	0.283	2.823	0.037	1.250
ค่าแปรปรวน	0.053	0.192	0.008	0.000	0.002	0.131	0.000	0.001
ค่าต่ำสุด	-0.899	-0.952	3.038	0.036	0.170	2.019	0.024	1.162
ค่าสูงสุด	0.845	3.717	3.714	0.081	0.545	4.872	0.054	1.382

ตาราง 13 ค่าทางสถิติของตลับลูกปืน 4 ชุดข้อมูล 2 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 13 กุมภาพันธ์ 2004 10:02 ถึง 13 กุมภาพันธ์ 2004 02:13

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	-0.007	-0.011	3.462	0.055	0.287	2.866	0.037	1.251
ค่ากลาง	-0.004	-0.089	3.469	0.054	0.283	2.816	0.037	1.249
ค่าแปรปรวน	0.051	0.203	0.009	0.000	0.002	0.132	0.000	0.001
ค่าต่ำสุด	-0.917	-0.922	3.035	0.039	0.171	2.070	0.025	1.162
ค่าสูงสุด	0.912	3.420	3.702	0.089	0.539	5.079	0.059	1.378

ตาราง 14 ค่าทางสถิติของตลับลูกปืน 4 ชุดข้อมูล 3 อยู่ในคลาส 1 (Normal) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 05 มีนาคม 2004 22:42 ถึง 05 มีนาคม 2004 21:42

statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	0.008	0.075	3.446	0.052	0.276	2.920	0.035	1.255
ค่ากลาง	0.007	-0.012	3.455	0.051	0.271	2.864	0.034	1.254
ค่าแปรปรวน	0.058	0.269	0.009	0.000	0.002	0.150	0.000	0.001
ค่าต่ำสุด	-0.990	-0.984	2.885	0.034	0.158	1.981	0.022	1.148
ค่าสูงสุด	1.126	6.858	3.702	0.088	0.611	5.422	0.060	1.454

เปรียบเทียบตลับลูกปืนตำแหน่งเดียวกันที่ชุดข้อมูลต่างกัน จากตาราง 15 คุณลักษณะของตลับลูกปืน ในช่วงที่เป็น คลาส 2 (Outer) ชุดข้อมูล 2 ตลับลูกปืน 1 คุณลักษณะเฉพาะ RMS เปรียบเทียบกับตาราง 16 ชุดข้อมูล 3 ตลับลูกปืน 3 จะพบว่า Skew , Kurtosis และ Clearance มีค่าแตกต่างกัน และคำนวณหา T-test ของ 8 คุณลักษณะพบว่า มีค่าน้อยกว่า 0.05 สามารถยอมรับสมมติฐานว่าทั้ง 2 ข้อมูลมีความใกล้เคียง

ตาราง 15 ค่าทางสถิติของตลับลูกปืน 1 ชุดข้อมูล 2 อยู่ในคลาส 2 (Outer) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 19 กุมภาพันธ์ 2004 02:22 ถึง 19 กุมภาพันธ์ 2004 06:22

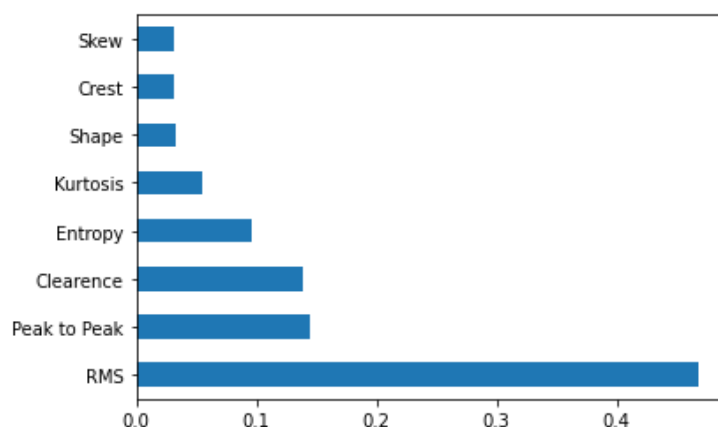
Statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	-0.092	2.647	3.029	0.346	2.417	3.598	0.208	1.338
ค่ากลาง	-0.059	0.774	3.308	0.329	1.875	3.238	0.215	1.298
ค่าแปรปรวน	0.653	18.669	0.613	0.028	2.946	1.483	0.007	0.023
ค่าต่ำสุด	-3.662	-2.027	0.122	0.001	0.002	1.014	0.000	1.014
ค่าสูงสุด	5.957	33.930	3.698	0.908	9.607	8.038	0.428	3.705

ตาราง 16 ค่าทางสถิติของตลับลูกปืน 3 ชุดข้อมูล 3 อยู่ในคลาส 2 (Outer) โดยมีช่วงเวลาของข้อมูลอยู่ระหว่าง 17 เมษายน 2004 22:42 ถึง 18 เมษายน 2004 2:42

Statistics	skew	Kurtosis	Entropy	RMS	Peak-to- -Peak	Crest	Clearance	Shape
ค่าเฉลี่ย	-0.274	3.340	3.094	0.401	2.877	4.056	0.248	1.321
ค่ากลาง	-0.085	1.318	3.223	0.393	2.555	3.705	0.247	1.305
ค่าแปรปรวน	0.736	27.785	0.311	0.019	2.160	1.917	0.006	0.010
ค่าต่ำสุด	-4.441	-2.027	0.415	0.003	0.005	1.067	0.002	1.027
ค่าสูงสุด	3.058	37.653	3.699	1.219	9.463	8.781	0.929	1.803

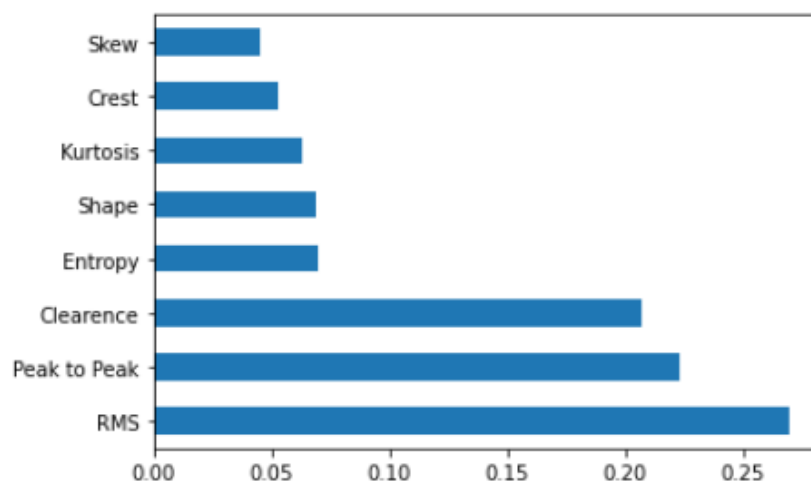
3.3 การหาคุณลักษณะเฉพาะที่สำคัญ (Feature important)

การพิจารณาคุณลักษณะเฉพาะที่สำคัญจะช่วยให้สามารถทราบได้ว่าคุณลักษณะเฉพาะใดมีความสำคัญกับโมเดลที่เลือกใช้งาน ทั้งหมด 8 คุณลักษณะเฉพาะทำการเปรียบเทียบคุณลักษณะเฉพาะ โดยใช้โมเดล Gradient boost Classification ดังภาพประกอบ 12 เราจะพบว่า คุณลักษณะเฉพาะ Clearance เท่ากับ 0.13 , Peak-to-Peak เท่ากับ 0.14 และ RMS เท่ากับ 0.46 มีค่าสูงเป็น 3 อันดับแรก เราจึงวิเคราะห์ข้อมูลของทั้ง 3 คุณลักษณะเฉพาะโดยพิจารณาจากกราฟ



ภาพประกอบ 12 เรียงลำดับคุณลักษณะเฉพาะที่มีความสำคัญของ Gradient boost Classification

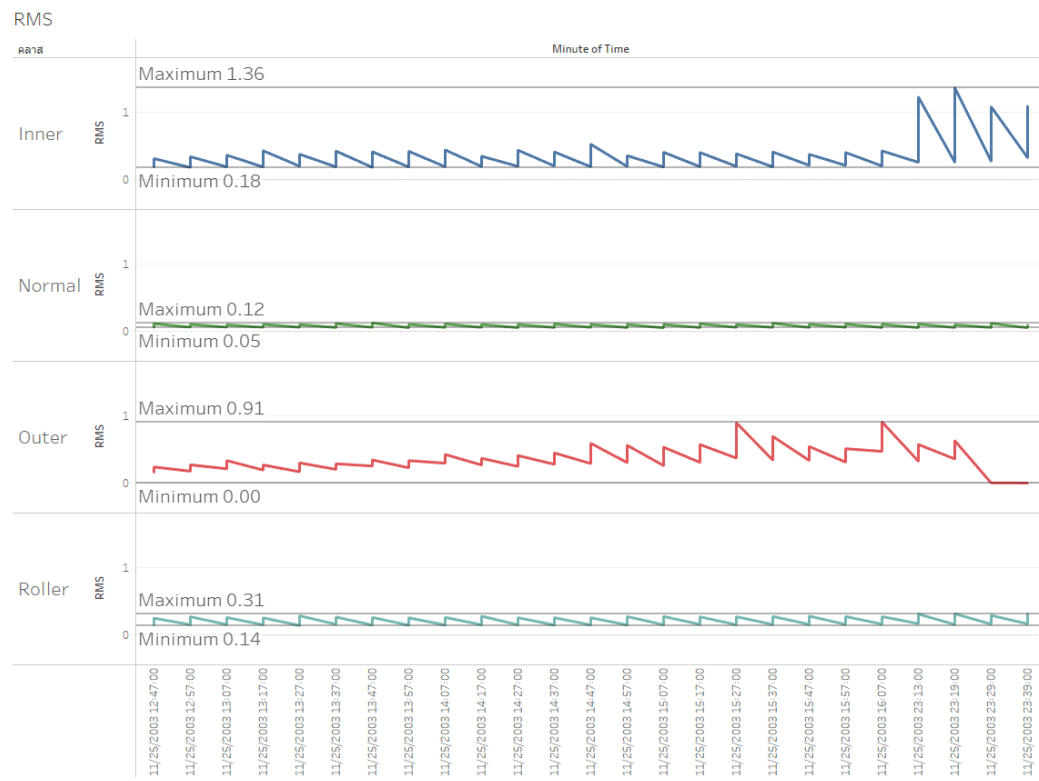
ภาพประกอบ 13 คุณลักษณะเฉพาะที่มีความสำคัญของป่าส้ม เราจะพบว่า คุณลักษณะเฉพาะ Clearance เท่ากับ 0.242 , Peak-to-Peak เท่ากับ 0.244 และ RMS เท่ากับ 0.251 มีค่าสูงเป็น 3 อันดับแรก เหมือนของการหาคุณลักษณะเฉพาะคล้าย Gradient boost Classification



ภาพประกอบ 13 เรียงลำดับคุณลักษณะเฉพาะที่มีความสำคัญของป่าส้ม

3.4 พิจารณาคลาสของข้อมูลแบบอนุกรมเวลา

คุณลักษณะเฉพาะของข้อมูล RMS จากการเปรียบเทียบค่าของข้อมูล จาก ภาพประกอบ 14 จะพบว่าคลาส 1 (Normal) แสดงอยู่ในในช่วงค่า 0.05 ถึง 0.12 และ คลาส 3 (Roller) แสดงอยู่ในช่วงค่า 0.14 ถึง 0.31 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวนต่ำกว่า 0.001 จะพบว่าค่าเปลี่ยนแปลงเพียงเล็กน้อยในคลาส 0 (Inner) แสดงอยู่ในช่วงค่า 0.18 ถึง 1.36 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวน 0.01 และมีค่าสูงขึ้นในช่วงท้ายของข้อมูล และ คลาส 2 (Outer) แสดงอยู่ในช่วงค่า 0.00 ถึง 0.91 ตลอดช่วงค่าความแปรปรวน 0.03 ค่ามีแนวโน้มสูงขึ้น ตั้งแต่ช่วงต้นของข้อมูลและเปลี่ยนแปลงอย่างรวดเร็วในช่วงกลางและท้ายตามอนุกรมเวลา



ภาพประกอบ 14 RMS แสดงอยู่ในรูปแบบกราฟเส้นแบ่งตามคลาส

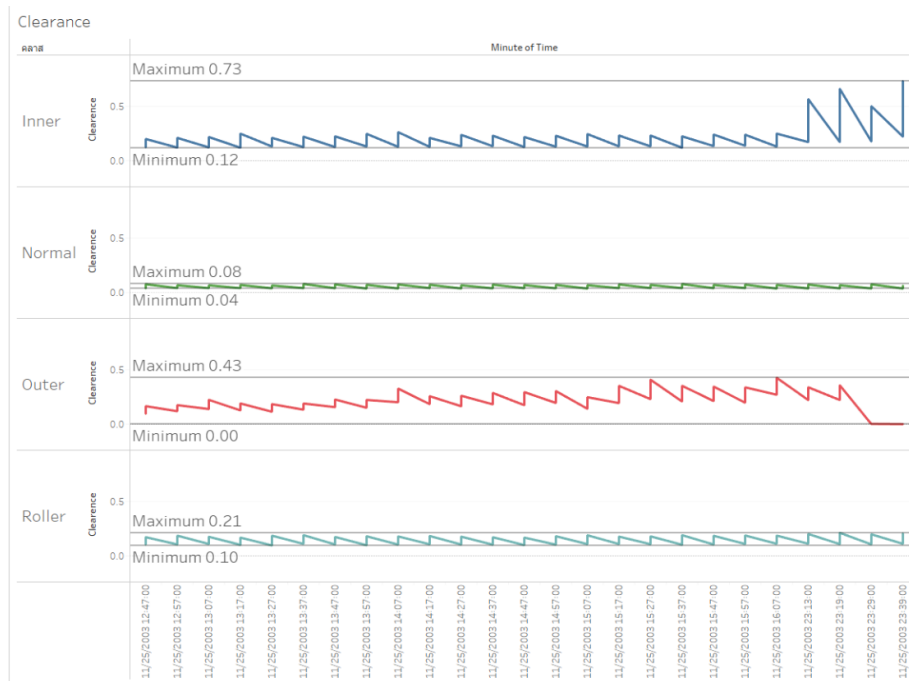
คุณลักษณะเฉพาะของข้อมูล Peak-to-Peak จากการเปรียบเทียบค่าของข้อมูล จากภาพประกอบ 15 จะพบว่าคลาส 1 (Normal) แสดงอยู่ในช่วงค่า 0.26 ถึง 0.86 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวน 0.005 ในคลาส 3 (Roller) แสดงอยู่ในช่วงค่า 0.48 ถึง 1.97 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวน 0.04 ส่วนคลาส 0 (Inner) แสดงอยู่ในช่วงค่า 0.64 ถึง 9.47 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวน 1.12 ค่ามีการเปลี่ยนแปลงสูงในช่วงแรก ลดลงในช่วงกลางของอนุกรมเวลาและมีค่าสูงขึ้นในช่วงท้ายของข้อมูล และ คลาส 2 (Outer) แสดงอยู่ในช่วงค่า 0 ถึง 9.61 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวน 2.95 ค่ามีแนวโน้มสูงขึ้นตั้งแต่ช่วงกลางของข้อมูลและเปลี่ยนแปลงรวดเร็ว พบการขึ้นลงของข้อมูลได้ชัดเจนกว่าคลาสอื่น



ภาพประกอบ 15 Peak-to-Peak แสดงอยู่ในรูปแบบกราฟเส้นแบ่งตามคลาส

คุณลักษณะเฉพาะของข้อมูล Clearance จากการเปรียบเทียบค่าของข้อมูล

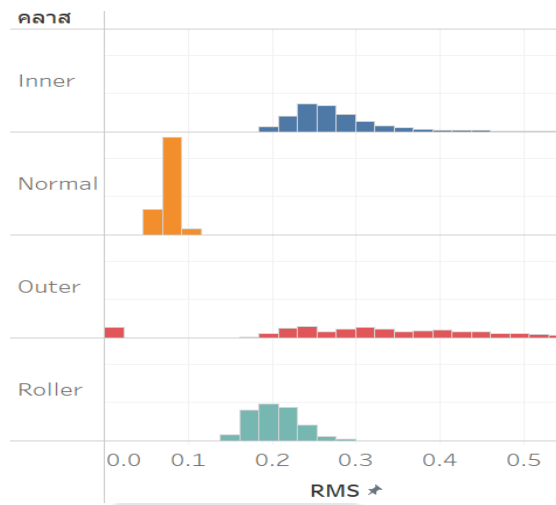
จากภาพประกอบ 16 จะพบว่าคลาส 1 (Normal) แสดงอยู่ในช่วงค่า 0.12 ถึง 0.73 และคลาส 3 (Roller) แสดงอยู่ในช่วงค่า 0.10 ถึง 0.21 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวนน้อยกว่า 0.001 มีการเปลี่ยนแปลงน้อย ส่วนคลาส 0 (Inner) แสดงอยู่ในช่วงค่า 0.12 ถึง 0.73 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวน 0.004 ค่ามีการเปลี่ยนแปลงสูงในช่วงท้ายของข้อมูล และคลาส 2 (Outer) แสดงอยู่ในช่วงค่า 0 ถึง 0.43 ตลอดทั้งช่วงของข้อมูลมีค่าความแปรปรวน 0.007 ค่ามีแนวโน้มสูงขึ้นตามการเปลี่ยนแปลงของเวลา



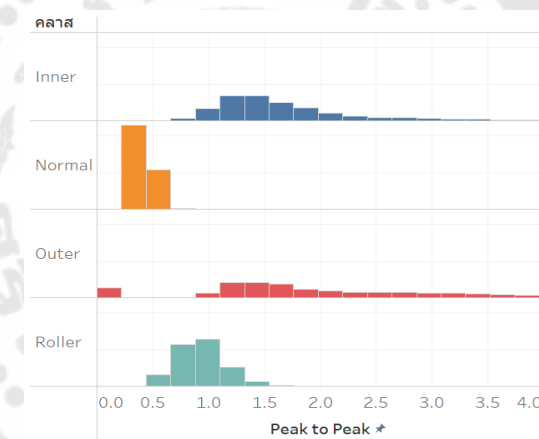
ภาพประกอบ 16 Clearance แสดงอยู่ในรูปแบบกราฟเส้นแบ่งตามคลาส

3.5 พิจารณาคลาสของข้อมูลจากการแจกแจงแบบปกติ

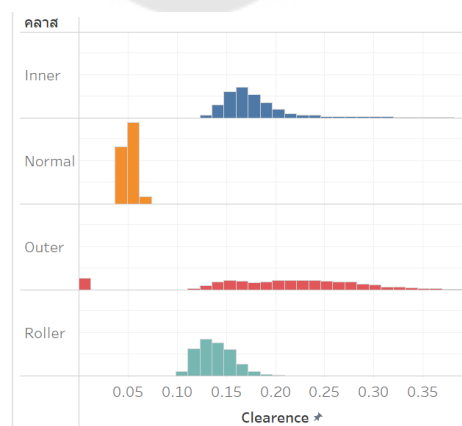
จากภาพประกอบ 17 ,18 และ 19 จากการแจกแจงแบบปกติของ RMS , Peak-to-Peak และ Clearance พบว่าคลาส 1 (Normal) มีช่วงของข้อมูลแยกออกจากคลาสชำรุดทั้ง 3 คลาสได้อย่างชัดเจนในคลาส 2 จะมีการกระจายตัวของข้อมูลในช่วงค่าที่กว้างกว่าคลาสนั้น มีความแปรปรวนสูงและในทั้ง 3 คลาสที่ชำรุดจะพบว่า มีบางช่วงของข้อมูลทับซ้อนกัน ทั้งใน RMS ,Peak-to-Peak และ Clearance



ภาพประกอบ 17 การแจกแจงแบบปกติของคุณลักษณะเฉพาะของ RMS



ภาพประกอบ 18 การแจกแจงแบบปกติของคุณลักษณะเฉพาะของ Peak to peak



ภาพประกอบ 19 การแจกแจงแบบปกติของคุณลักษณะเฉพาะของ Clearance

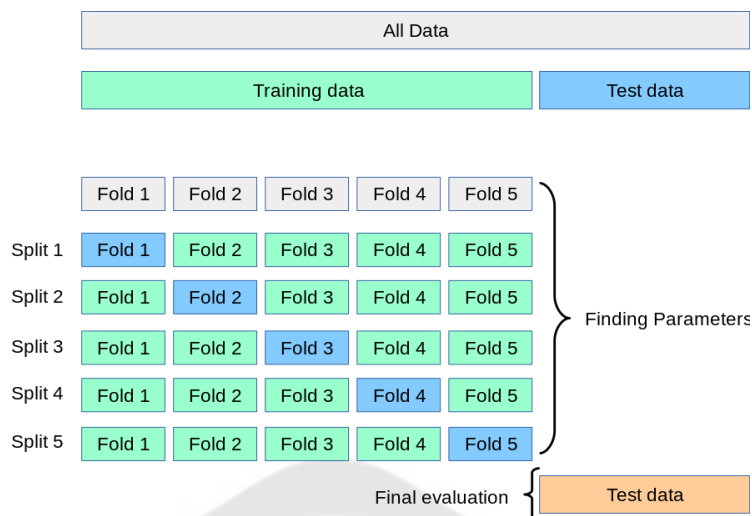
3.6 การปรับจูนพารามิเตอร์ของโมเดล และ การทำ Cross validation

การทำปรับจูนพารามิเตอร์ (Hyper parameter tuning) คือ การกำหนดช่วงค่าของพารามิเตอร์ในช่วงค่าของแต่ละพารามิเตอร์ที่สนใจ เพื่อหาพารามิเตอร์ที่สามารถทำงานร่วมกับข้อมูลและโมเดลนั้นได้ดีที่สุด ค่าที่ได้จากการปรับจูนเรียกว่า พารามิเตอร์ที่ดีที่สุด (Best parameter) ซึ่งพารามิเตอร์ทำให้โมเดลมีค่าความถูกต้องมากที่สุด โดยทั้ง 4 โมเดลกำหนดค่าไว้ดังตาราง 17

ตาราง 17 แสดงถึงช่วงค่าของพารามิเตอร์ที่ใช้ในการปรับจูนของแต่ละโมเดล

การเรียนรู้ของเครื่อง	Hyper parameter tuning
SVM	C :[0.5,0.75,1, 1.5] , kernel :[linear, rbf]
KNN	neighbors : [1 :15]
RF	max_depth: [2,4,6,8,10] , n_estimators:[100,200,300,400,500] , criterion:['gini','entropy']
XG-Boost	max_depth :[2,4,6,8,10] , n_estimators:[100,200,300,400,500] learning_rate': [0.01,0.05]objective : softmax, softprob , booster:btree , tree_method: gpu_hist

การทำ Cross validation คือการแบ่งจำนวนข้อมูลออกเป็นตามกลุ่มข้อมูลที่ตั้งค่าไว้ ในการทดลองนี้กำหนดค่า เป็น 5-Fold Cross validation เป็นการแบ่งข้อมูลเป็น 5 กลุ่ม แล้วนำข้อมูล 4 ส่วนไปทำการสอนในโมเดล และ 1 ส่วนไว้สำหรับ Test ข้อมูล จะดำเนินการสลับกลุ่มข้อมูลไปเรื่อยๆจนครบ 5 ครั้งตามการทดลองนี้ ทำให้ในทุกกลุ่มข้อมูลจะได้เป็นทั้งข้อมูลในการ Train และ Validation ประโยชน์ที่ได้จากการทำ Cross validation คือ ทำให้การ Train ข้อมูลลดการ Bias ของโมเดลเพราะ ทุกกลุ่มข้อมูลได้เป็นทั้ง train และ Validation ร่วมไปถึงโมเดลที่มีการทำซ้ำตามจำนวนครั้งที่กำหนดค่าไว้ ในการทดลองนี้กำหนดไว้ 5 ครั้ง ความถูกต้องของโมเดลจะนำมาเฉลี่ยกัน เพื่อใช้ในการเปรียบเทียบโมเดลประสิทธิภาพของโมเดล โดยปกติแล้วจะนิยมทำ Cross validation ควบคู่ไปกับการปรับจูนพารามิเตอร์ ดังภาพประกอบ 20



ภาพประกอบ 20 การทำ Fold Cross validation

ที่มา: (scikit-learn., 2022)

ผลลัพธ์ของการเปรียบเทียบโมเดลทั้ง 4 โมเดล จากการทำปรับพารามิเตอร์ (Hyper parameter tuning) และ 5-Fold Cross Validation ได้ค่าเฉลี่ยของค่าความถูกต้องของแต่ละโมเดล จะพบว่า XG-Boost มีค่าเฉลี่ยความถูกต้องสูงสุดอยู่ที่ 87.42% ดังตาราง 18

ตาราง 18 การเปรียบเทียบ ค่าเฉลี่ยของค่าความถูกต้องของโมเดลจากการทำ Hyper parameter tuning และ 5-Fold Cross Validation

การเรียนรู้ของเครื่อง	Hyper parameter tuning	ค่าเฉลี่ยความถูกต้อง
SVM	C :11 , kernel :rbf	81.85%
KNN	neighbors : 11	82.10%
RF	max_depth:10 , Estimators : 500 , Criterion : entropy	86.59%
XG-Boost	max_depth:6,Estimators:500,Objective: softmax , learning_rate:0.05, booster:btree , tree_method: gpu_hist	87.42%

บทที่ 4

ผลลัพธ์จากการวิจัย

ในการวิจัยเรื่องการทำนาย คลาสของตลับลูกปืนโดยเทคนิคการเรียนรู้ของเครื่อง ผู้วิจัยได้เปรียบเทียบประสิทธิภาพการเรียนรู้ของเครื่อง บรรลุวัตถุประสงค์ของงานวิจัยได้กำหนดดังนี้

1. ผลลัพธ์ลัพธ์ของทั้ง 4 โมเดล
2. ผลลัพธ์ของการเปรียบเทียบโมเดลทั้ง 4 โมเดล
3. กราฟอธิบายผลการทดลองการแบ่งคลาสของโมเดล

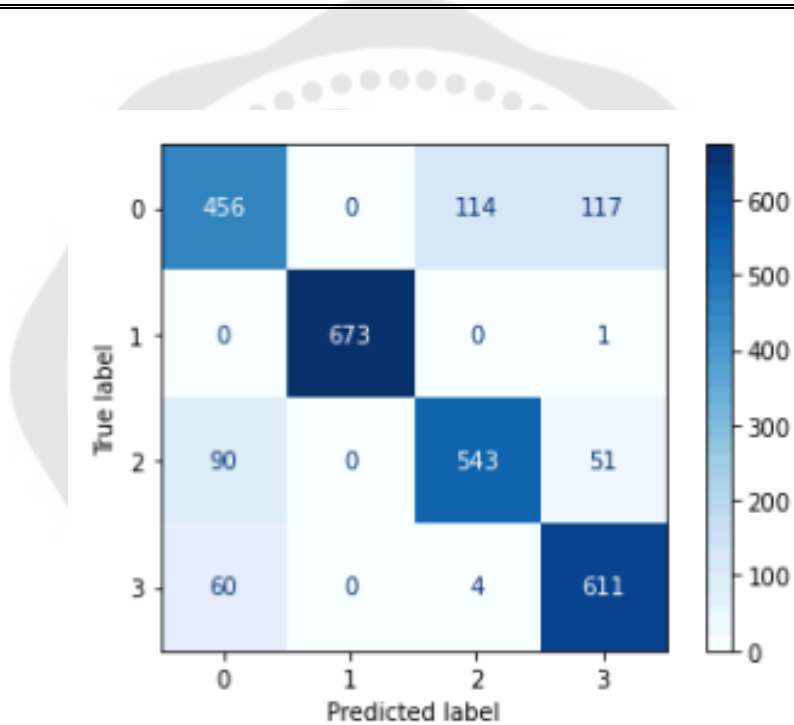
1. ผลลัพธ์ลัพธ์ของทั้ง 4 โมเดล

1.1 โมเดล SVM

ในการสอนโมเดล กำหนดพารามิเตอร์ที่ต้องการแล้วพบว่า พารามิเตอร์ที่ดีที่สุดคือ C เท่ากับ 1.5 , kernel เป็น RBF และทำการ Cross validation 5 ครั้ง ผลปรากฏว่า ค่าความถูกต้องเฉลี่ยเป็น 81.75% จากตาราง 19 ได้ทำการวัด Classification report ของ SVM พบว่าค่าความแม่นยำสูงสุดเป็นของคลาส 1 (Normal) ได้ 100% และค่าความแม่นยำต่ำสุดเป็นของคลาส 0 (Inner) ได้ 75.25% จากภาพประกอบ 21 คลาส 0 (Inner) ทำนายผิดเป็น คลาส 2 อยู่ 114 ค่า ทำนายผิดเป็น คลาส 3 อยู่ 117 ค่า

ตาราง 19 Classification report ของ SVM

คลาส	0 (Inner)	1 (Normal)	2 (Outer)	3 (Roller)
Precision	75.25%	100.00%	82.15%	78.33%
Recall	66.38%	99.85%	79.39%	90.37%
F1-score	70.53%	99.93%	80.74%	87.08%



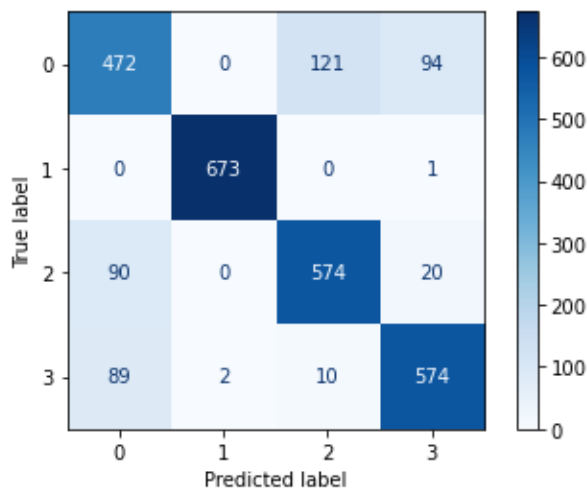
ภาพประกอบ 21 Classification report SVM

1.2 โมเดล KNN

ในการสอนโมเดล กำหนดพารามิเตอร์ที่ต้องการจูน แล้วพบว่าพารามิเตอร์ที่ดีที่สุดคือ $n_neighbors = 11$ และทำการ Cross validation 5 ครั้ง ผลปรากฏว่า ค่าความถูกต้องเฉลี่ยเป็น 82.10% จากตาราง 20 ได้ทำการวัด Classification report ของ KNN พบว่าค่าความแม่นยำสูงสุดเป็นของคลาส 1 (Normal) ได้ 99.70% และค่าความแม่นยำต่ำสุดเป็นของคลาส 0 (Inner) ได้ 72.50% จากภาพประกอบ 22 คลาส 0 (Inner) ทำนายผิดเป็นคลาส 3 อยู่ 1 ค่า และคลาส 0 ทำนายผิดเป็น คลาส 2 อยู่ 121 ค่า ทำนายผิดเป็น คลาส 3 อยู่ 94 ค่า

ตาราง 20 Classification report KNN

คลาส	0 (Inner)	1 (Normal)	2 (Outer)	3 (Roller)
Precision	75.50%	99.70%	81.42%	83.31%
Recall	68.70%	99.85%	83.92%	85.04%
F1-score	70.55%	99.78%	20.65%	84.16%



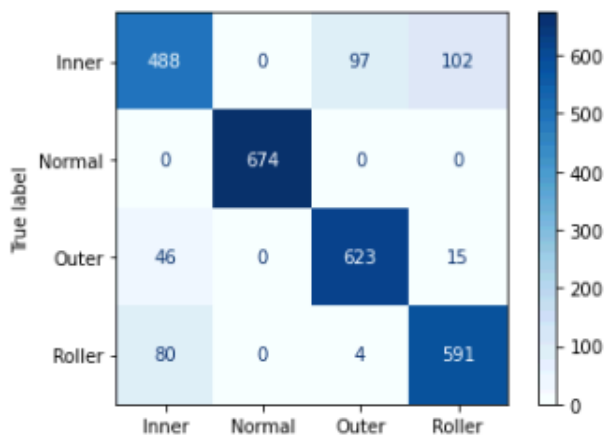
ภาพประกอบ 22 Confusion matrix KNN

1.3 โมเดลป่าสุ่ม

ในการสอนโมเดล กำหนดพารามิเตอร์ที่ต้องการจนแล้วพบว่าพารามิเตอร์ที่ดีที่สุดคือ max_depth:10 , Estimators:500 , Criterion: entropy และทำการ Cross validation 5 ครั้ง ผลปรากฏว่า ค่าความถูกต้องเฉลี่ยเป็น 86.69% จากตาราง 21 ได้ทำการวัด Classification report ของ ป่าสุ่ม พบว่าค่าความแม่นยำสูงสุดเป็นของคลาส 1 (Normal) ได้ 100% และค่าความแม่นยำต่ำสุดเป็นของคลาส 0 (Inner) ได้ 79.59% จากภาพประกอบ 23 คลาส 0 ทำนายผิดเป็น คลาส 2 อยู่ 97 ค่า และ ทำนายผิดเป็น คลาส 3 อยู่ 102 ค่า

ตาราง 21 Classification report ป่าสุ่ม

คลาส	0 (Inner)	1 (Normal)	2 (Outer)	3 (Roller)
Precision	79.59%	100.00%	89.38%	83.64%
Recall	73.80%	100.00%	92.25%	87.11%
F1-score	76.59%	100.00%	90.79%	85.34%



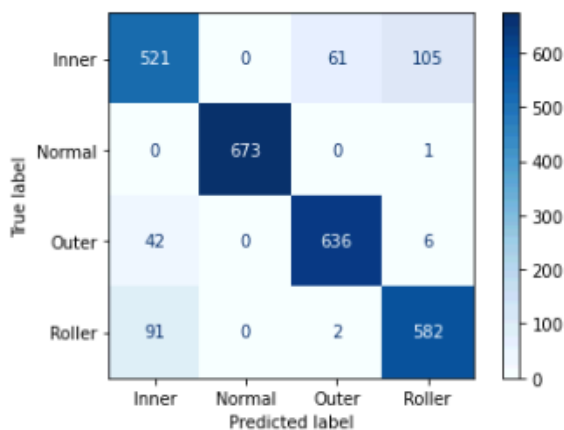
ภาพประกอบ 23 confusion matrix ป่าสุ่ม

1.4 โมเดล XG-Boost

ในการสอนโมเดลกำหนดพารามิเตอร์ที่ต้องการจน แล้วพบว่า พารามิเตอร์ที่ดีที่สุดคือ max_depth:6, Estimators:500, Objective:softmax, learning_rate:0.05, booster: btree ,tree_method:gpu_hist และทำการ Cross validation 5 ครั้ง ผลปรากฏว่าค่าความถูกต้อง 87.52% จากตาราง 22 ได้ทำการวัด Classification report ของ XG-Boost พบว่าค่าความแม่นยำสูงสุดเป็นของคลาส 1 (Normal) ได้ 100% และค่าความแม่นยำต่ำสุดเป็นของคลาส 0 (Inner) ได้ 78.60% จากภาพประกอบ 24 คลาส 1 (Normal) ทำนายไม่ผิด และคลาส 0 ทำนายผิดเป็น คลาส 2 อยู่ 58 ค่า ทำนายผิดเป็นคลาส 3 อยู่ 111 ค่า

ตาราง 22 Classification report XG-Boost

คลาส	0 (Inner)	1 (Normal)	2 (Outer)	3 (Roller)
Precision	79.66%	100.00%	90.99%	83.86%
Recall	75.84%	99.85%	92.28%	86.22%
F1-score	77.70%	99.93%	91.97%	85.03%



ภาพประกอบ 24 confusion matrix XG-Boost

2. ผลลัพธ์ของการเปรียบเทียบโมเดลทั้ง 4 โมเดล

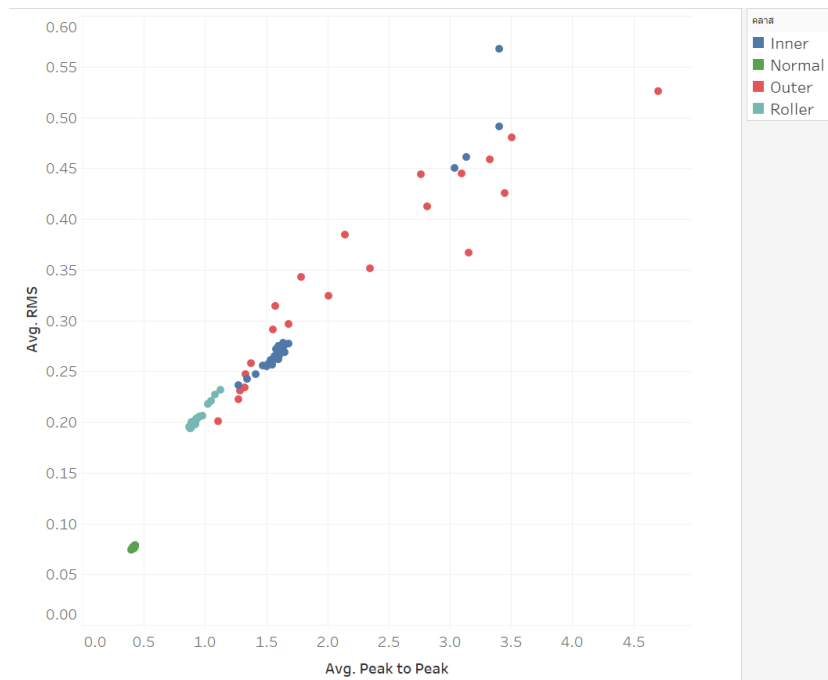
ตาราง 23 นำโมเดลทั้ง 4 มาทำการทดสอบ และการเปรียบเทียบสรุปผล Macro avg. F1-score และ ค่าความถูกต้องของทุกคลาส การทดลอง พบว่าโมเดล XG-Boost มีค่าความถูกต้อง 88.68% เปรียบเทียบ SVM เป็นโมเดลพื้นฐานในการวิจัยนี้ เพราะในหลายงานวิจัยที่เกี่ยวข้องกับตลับลูกปืนใช้โมเดล SVM ให้ค่าความถูกต้องและความแม่นยำสูง ประกอบกับโมเดล SVM สามารถแยกคลาสในรูปแบบลักษณะของข้อมูลไม่เป็นเชิงเส้นได้ จากการทดลองนี้ค่าความถูกต้องของ SVM อยู่ที่ 83.93%

ตาราง 23 เปรียบเทียบความถูกต้อง และ F1-score ของโมเดล

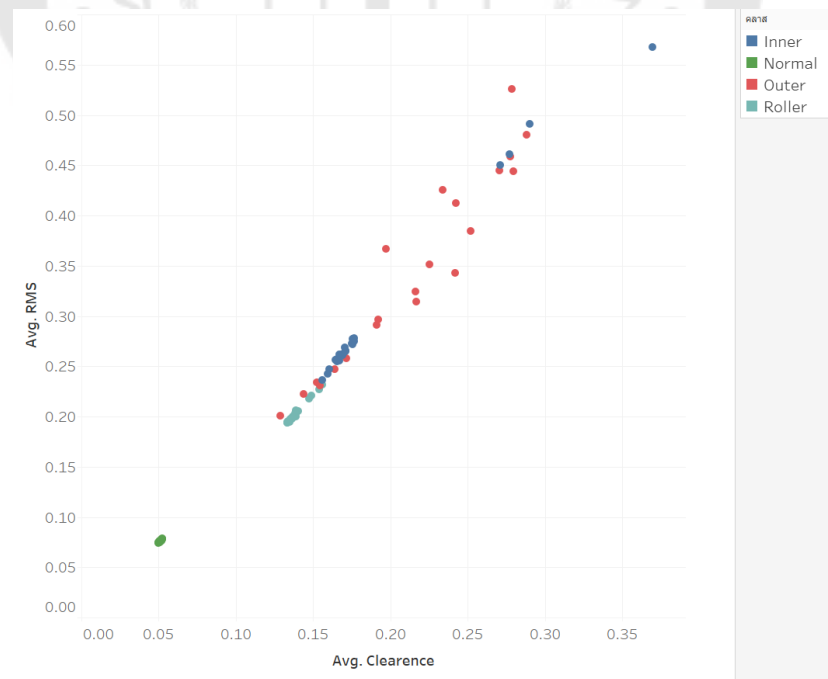
คลาส	SVM(Baseline)	KNN	ป่าสุ่ม	XG-boost
Macro avg.F1	83.80%	84.30%	87.98%	88.66%
Accuracy	83.93%	84.29%	87.93%	88.68%

3. กราฟอธิบายผลการทดลองการแบ่งคลาสของโมเดล

กราฟแสดงการกระจายตัวของข้อมูลระหว่าง คุณลักษณะเฉพาะ โดยแต่ละจุดเป็น ค่าเฉลี่ยค่าเฉลี่ยต่อ 1 นาที และใช้สีในการระบุคลาสของข้อมูล โดยคลาส 0 (Inner) สีน้ำเงิน , คลาส 1 (Normal) สีเขียว , คลาส 2 (Outer) สีแดง และ คลาส 3 (Roller) สีฟ้า จากภาพประกอบ 25 และ 26 กราฟทั้งสองมีลักษณะใกล้เคียงกัน กราฟมีความเป็นเส้นตรง เมื่อพิจารณาข้อมูลจาก คลาส 1 (Normal) สีเขียว ข้อมูลแยกตัวจากกลุ่มอื่นอย่างชัดเจน และกลุ่มข้อมูลเกาะกลุ่มกันมากกว่าคลาสนอื่น ในส่วนคลาสตลับลูกปืนชำรุด คลาส 0 (Inner) สีน้ำเงิน , คลาส 2 (Outer) สีแดง และคลาสน 3 (Roller) สีฟ้า มีบางส่วนของข้อมูลมีส่วนที่ทับซ้อนกัน มีโอกาสที่การทำนายของ คลาสทั้ง 3 ทำนายผิดระหว่างคลาสน

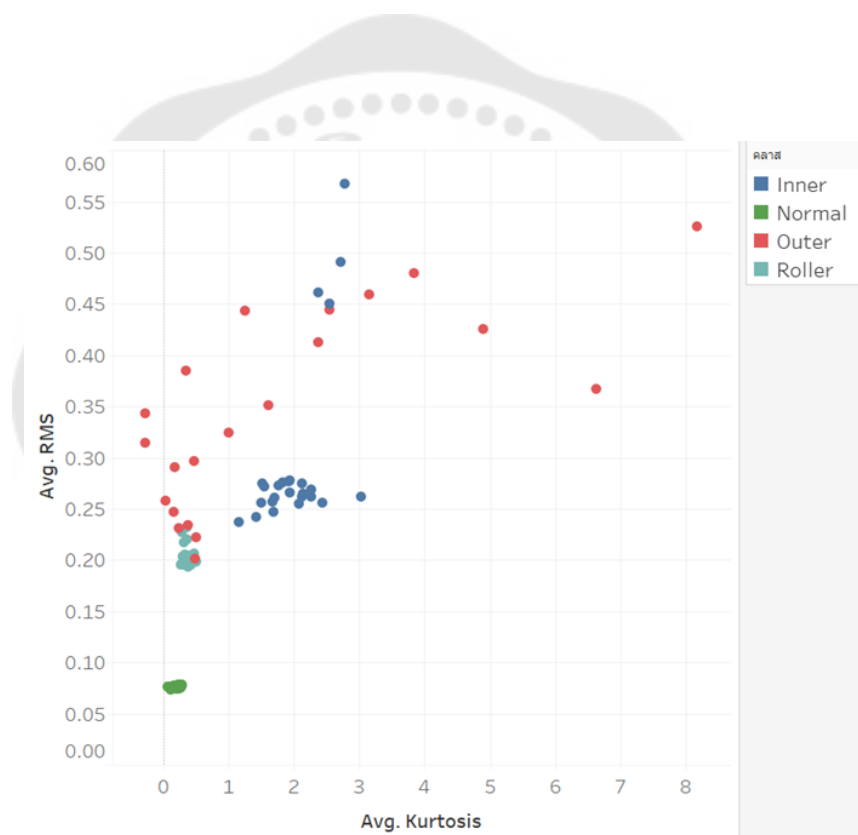


ภาพประกอบ 25 กราฟแสดงการกระจายตัว ระหว่าง Avg. RMS และ Avg. Peak-to-Peak

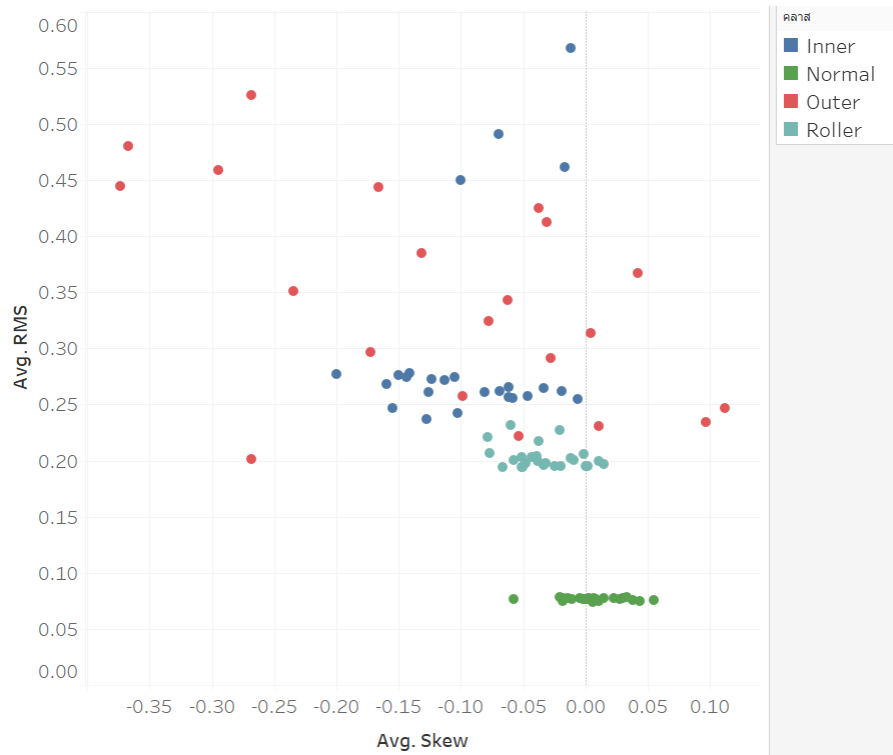


ภาพประกอบ 26 กราฟแสดงการกระจายตัวระหว่าง Avg. RMS และ Avg. Clearance

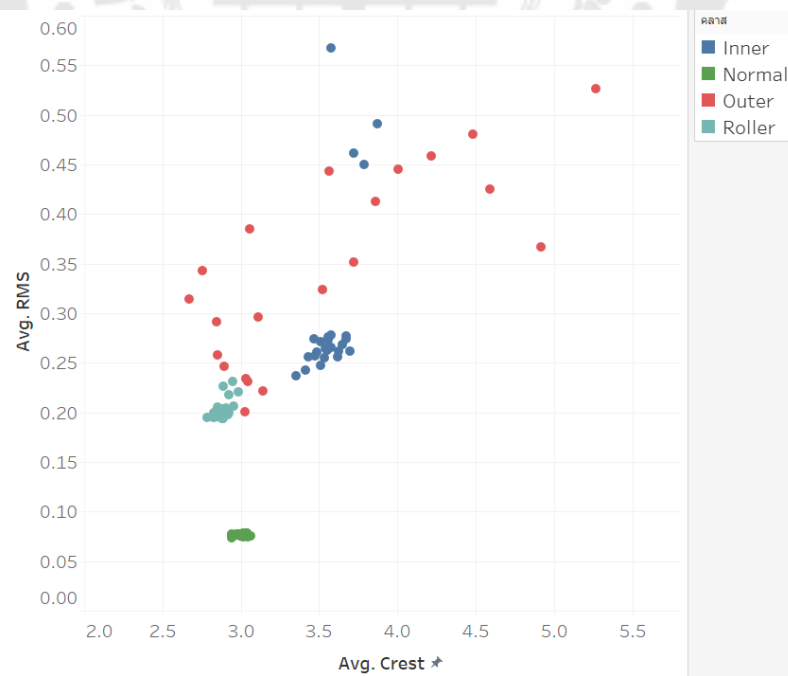
จากภาพประกอบ 27 , 28 และ 29 เมื่อพิจารณาข้อมูลทั้ง 3 ภาพ คลาส 1 (Normal) สีเขียว ข้อมูลแยกตัวจากกลุ่มอื่นอย่างชัดเจน จากภาพซ้ำกราฟการกระจายตัวระหว่าง Avg. RMS และ Avg. Kurtosis (ภาพซ้าย) จะพบว่าคลาส 0 (Inner) สีน้ำเงินแยกออกจากคลาสอื่นได้ชัดเจน แต่มีบางจุดแยกออกไปเป็นอีกกลุ่ม ซึ่งอาจจะเป็นสาเหตุให้คลาส 0 (Inner) มีความถูกต้องของการทำนายน้อยที่สุด คลาส 2 (Outer) สีแดง ทั้ง 3 ภาพ มีการกระจายตัวที่สูง สอดคล้องกับคลาสที่มีความแปรปรวนมากกว่าคลาสอื่นและคลาส 3 (Roller) กลุ่มข้อมูลเกาะกลุ่มกัน มีกลุ่มข้อมูลใกล้เคียงกับกลุ่มข้อมูลของคลาส 0 (Inner) สีน้ำเงิน



ภาพประกอบ 27 กราฟแสดงการกระจายตัวระหว่าง RMS และ Avg. Kurtosis



ภาพประกอบ 28 กราฟแสดงการกระจายตัวระหว่าง RMS และ Avg. Skew



ภาพประกอบ 29 กราฟแสดงการกระจายตัว ระหว่าง RMS และ Avg. Crest

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และ ข้อเสนอแนะ

สรุปผลโดยสามารถแบ่งหัวข้อในการสรุปได้ดังต่อไปนี้

1. สรุปผลการวิจัย
2. อภิปรายผลการวิจัย
3. ข้อเสนอแนะ

1. สรุปผลการวิจัย

งานวิจัยนี้นำเสนอการใช้เทคโนโลยีการเรียนรู้ของเครื่องในการวิเคราะห์สัญญาณความสั่นสะเทือนของตลับลูกปืนเพื่อทำนายความผิดปกติที่จะเกิดขึ้นกับตลับลูกปืนจำนวน 3 ประเภท อันได้แก่ Inner เป็นการชำรุดที่วงในของตลับลูกปืน Outer เป็นการชำรุดที่วงนอกของตลับลูกปืน และ Roller เป็นการชำรุดที่ลูกปืน ระบบที่นำเสนอจะนำสัญญาณความสั่นสะเทือนที่วัดจาก accelerometer ที่ติดตั้งบริเวณตลับลูกปืนมาคำนวณหาค่าทางสถิติพื้นฐานค่า Root Mean Square (RMS) , ค่า Peak-to-peak ค่าทั้งสองในทางกลศาสตร์การสั่นสะเทือนสามารถนำไปวิเคราะห์หาความผิดปกติของตลับลูกปืนได้ จึงเป็นพารามิเตอร์สำคัญในการวิเคราะห์การเรียนรู้ของเครื่อง และพารามิเตอร์อื่น ๆ รวมทั้งหมดมีจำนวน 12 ค่า และนำมาใช้เป็นคุณลักษณะเฉพาะที่จะป้อนให้กับโมเดลการเรียนรู้ของเครื่องใช้ในการทำนายความผิดปกติของตลับลูกปืน เราทดสอบประสิทธิภาพของระบบที่นำเสนอกับชุดข้อมูล Intelligent Maintenance Systems (IMS) จากผลการทดลอง ระบบที่นำเสนอที่ใช้โมเดลการเรียนรู้ของเครื่องทั้ง 4 ชนิด SVM , KNN , ป่าสุ่ม และ XG-Boost พบว่า XG-Boost มีค่าความแม่นยำในการทำนายความผิดปกติทั้ง 3 แบบที่ดีที่สุด โดยระบบที่นำเสนอสามารถตรวจจับความผิดปกติแบบ Inner, Outer และ Roller โดยมีค่าความแม่นยำ (Precision) เท่ากับ 79.66% , 90.99% และ 83.86% ค่าระลึก(Recall) เท่ากับ 75.84% , 92.28% และ 86.22% ค่า F1score 77.70% , 91.97% และ 85.03% ตามลำดับ จะพบว่าความผิดปกติแบบ Outer มีความแม่นยำมากที่สุด ตามด้วย Roller และ Inner โดยคุณลักษณะเฉพาะที่สำคัญที่สุด 3 ลำดับแรกประกอบไปด้วยค่า RMS, Peak-to-peak และ Clearance โดยมีค่าความสำคัญเท่ากับ 0.46 , 0.14 และ 0.13 ตามลำดับ

2. อภิปรายผลการวิจัย

ในการวิจัยนี้ได้ศึกษาการชำรุดของตลับลูกปืน โดยใช้การเรียนรู้ด้วยเครื่องทำนายหาคลาสของตลับลูกปืน พบว่าเราสามารถแบ่งแยกคลาสของตลับลูกปืนโดยทำการแปลงสัญญาณคลื่นสะท้อน ให้เป็นคุณลักษณะเฉพาะ ที่ใช้ในการวิเคราะห์ข้อมูล เราจะพบว่าคุณลักษณะเฉพาะสามารถแสดงถึงความแตกต่างระหว่างคลาสได้จากการแสดงในรูปแบบกราฟการแจกแจงแบบปกติ จากรูปประกอบ 17 , 18 และ 19 การแจกแจงแบบปกติ RMS Peak-to-Peak และ Clearance จะพบ Normal อยู่ในช่วง 0.05 ถึง 0.12 หากเปรียบเทียบคลาส Roller อยู่ในช่วง 0.14 ถึง 0.31 , Inner อยู่ในช่วง 0.18 ถึง 0.44 และ Outer อยู่ในช่วง 0.18 ถึง 0.53 จากการพิจารณาพบว่าคลาส Normal อยู่ในช่วงต่ำกว่า และมีความแปรปรวนน้อย เพราะในขณะที่ตลับลูกปืนทำงานอยู่ในช่วง Normal สัญญาณคลื่นสะท้อนอยู่ในช่วงความแปรปรวนที่ต่ำ เนื่องจากส่วนประกอบต่างๆในตลับลูกปืนอยู่ในสภาพพร้อมใช้งาน การสึกหรอนจากการใช้งานต่ำ ในคลาสอื่นจะพบว่า RMS มีค่าสูงขึ้นจากคลาส Normal และการแจกแจงแบบปกติมีความแปรปรวนที่สูงในคุณลักษณะเฉพาะ Peak-to-Peak และ Clearance มีลักษณะของการแจกแจงปกติ และความแปรปรวนคล้ายคลึงกับ RMS ทำให้คลาส 1 (Normal) แยกออกจากคลาสอื่นได้ชัดเจน ส่งผลให้คลาส 1 (Normal) มีค่าความถูกต้อง 100% หากจะพิจารณาค่าความถูกต้องของคลาสอื่นสามารถอธิบายได้จากจากรูปที่ 14 คุณลักษณะเฉพาะ RMS จะพบว่าคลาส 1 (Normal) แสดงอยู่ในช่วงค่า 0.05 ถึง 0.12 คลาส แยกตัวตัวได้ชัดเจน และ คลาส 3 (Roller) แสดงอยู่ในช่วงค่า 0.14 ถึง 0.31 มีการเปลี่ยนแปลงของค่าน้อย มีช่วงค่าถัดจาก คลาส 1 (Normal) ในคลาส 0 (Inner) แสดงอยู่ในช่วงค่า 0.18 ถึง 1.36 และมีค่าสูงขึ้นในช่วงท้ายของข้อมูล มีช่วงค่าที่ใกล้เคียงกับ คลาส 3 (Roller) ทำให้มีโอกาสที่คลาสนี้มีความถูกต้องน้อยกว่าคลาสอื่นและคลาส 2 (Outer) แสดงอยู่ในช่วงค่า 0 ถึง 0.91 ตลอดช่วงค่าความแปรปรวนเพิ่มขึ้น ค่ามีแนวโน้มสูงขึ้นตั้งแต่ช่วงต้นของข้อมูล ช่วงข้อมูลที่กว้างกว่าคลาสอื่น และความแปรปรวนที่มาก ทำให้การทำนายมีโอกาสแบ่งแยกออกจากคลาส 0 (Inner) และ คลาส 3 (Roller)

การอธิบายผลการทดลองการแบ่งคลาสของโมเดลโดยใช้กราฟ จากภาพประกอบ 25 ถึง 29 จะพบว่า คุณลักษณะเฉพาะสำคัญสามารถแบ่งคลาส 1 (Normal) ได้อย่างชัดเจน และพบว่า คลาส 0 (Inner) มีกลุ่มข้อมูลที่แตกต่างจากกลุ่มข้อมูลหลัก จึงเป็นสาเหตุทำให้คลาส 0 (Inner) มีความถูกต้องน้อยที่สุด หากทำการพิจารณาข้อมูลในการตัดข้อมูลกลุ่มข้อมูลที่แตกต่างออกมีความเป็นไปได้ที่ช่วยทำให้ความถูกต้องของคลาส 0 (Inner) เพิ่มสูงขึ้น

จากการพิจารณาชุดข้อมูลอื่นที่ไม่ใช่ช่วงข้อมูลที่นำไปเข้าการเรียนรู้ของเครื่อง โดยพิจารณากลุ่มตัวอย่าง ของคลาส 1 (Normal) ทั้งตลับลูกปืนทั้ง 4 ตลับลูกปืนในรูปแบบข้อมูลชุดเดียวกันและคนละชุดข้อมูล ตาราง 3 ถึง 14 ผลของการเปรียบเทียบพบว่า ค่าทางสถิติ คือ ค่าสูงสุด , ค่าต่ำสุด , ค่าเฉลี่ย , ค่ามัธยฐาน และ ค่าความแปรปรวน พบว่าค่ามีความใกล้เคียงกัน จึงได้มีการทดสอบเพิ่มเติมโดยการใช้ T-test เป็นการทดสอบสมมติฐานเพื่อใช้เปรียบเทียบค่าเฉลี่ยตัวอย่างกลุ่มข้อมูลสองกลุ่มมีความสัมพันธ์กันหรืออิสระต่อกันหรือไม่ ทั้งนี้จากการทดสอบของคลาส 1 (Normal) พบว่า P-value น้อยกว่า 0.05 เป็นการยอมรับสมมติฐานว่ากลุ่มข้อมูลทั้งสองมีความใกล้เคียงกัน เป็นการยืนยันได้ว่าสามารถนำชุดข้อมูลอื่น หรือ ช่วงเวลาอื่นๆ ในชุดเดียวกัน มา Test ด้วยการเรียนรู้ของเครื่องได้และเราคาดว่าความถูกต้องและความแม่นยำมีค่าสูงยอมรับผลการทำนายได้ ในส่วนของคลาส 2 (Outer) เป็นตัวแทนของคลาสชำรุด ทำการเปรียบเทียบชุดข้อมูล 2 ตลับลูกปืน 1 ดังตาราง 15 และ ชุดข้อมูล 3 ตลับลูกปืน 3 ดังตาราง 16 ค่าทางสถิติ คือ ค่าสูงสุด , ค่าต่ำสุด , ค่าเฉลี่ย , ค่ามัธยฐาน และค่าความแปรปรวน พบว่าค่ามีความใกล้เคียงกันบางค่า และมีบางค่ามีค่าแตกต่างกัน ทั้งนี้สาเหตุเป็นเพราะ ตำแหน่งติดตั้งของตลับลูกปืนอยู่คนละตำแหน่งการทำงานของเครื่องจักร มีความเป็นไปได้ที่ค่าการสั่นสะเทือนของตลับลูกปืนมีความต่างกัน จึงทำการทดสอบเพิ่มเติม โดยการใช้ T-test เป็นการทดสอบสมมติฐาน ทั้งนี้จากการทดสอบของคลาส 2 (Outer) พบว่า P-value น้อยกว่า 0.05 เป็นการยอมรับสมมติฐานว่ากลุ่มข้อมูลทั้งสองมีความใกล้เคียงกัน หากนำข้อมูลที่ได้ไปเข้าการเรียนรู้ของเครื่องเพื่อ Test ผลลัพธ์ คาดว่าสามารถทำนายคลาสชำรุดได้

ในการเปรียบเทียบโมเดลจะพบว่าโมเดล XG-Boost มีค่าความถูกต้องอยู่ที่ 88.68% หากจะเปรียบเทียบระหว่าง โมเดล SVM , KNN และ ป่าสุ่ม ค่าความถูกต้องโมเดลเท่ากับ 83.93% , 84.29% และ 87.93% หากพิจารณาค่าความถูกต้องของแต่ละคลาสในแต่ละโมเดล มีผลเหมือนกันคือ คลาส1 (Normal) ถูกต้องที่สุด และ คลาส 0 (Inner) ค่าความถูกต้องน้อยที่สุด

3. ข้อเสนอแนะ

1. ก่อนนำข้อมูลเข้าโมเดลควรทำการกระบวนเตรียมข้อมูล เพื่อแก้ไขข้อมูลไม่สมดุลระหว่างคลาส (Imbalance data) โดยการสุ่มตัวอย่างข้อมูลปริมาณใกล้เคียงกันในแต่ละคลาส

2. ในการวิจัยนี้ใช้กับชุดข้อมูล IMS หากเปลี่ยนชุดข้อมูลอาจต้องมีการ Retrain โมเดล เพราะ ชนิดของตลับลูกปืน ขนาดของตลับลูกปืน ตำแหน่งติดตั้ง และชนิดเครื่องมือวัดการสั่นสะเทือน ส่งผลต่อค่าวัดของการสั่นสะเทือน

3. การนำข้อมูลคุณลักษณะด้านความถี่มาพิจารณา สามารถช่วยอธิบายหาสัญญาณผิดปกติที่เกิดกับตลับลูกปืนได้ เป็นอีกคุณลักษณะเฉพาะที่ช่วยทำให้ช่วยตรวจสอบสัญญาณผิดปกติล่วงหน้า

4. ข้อมูลของคลาสที่ซ้ำชุดได้นำข้อมูลช่วงท้ายของชุดข้อมูลเป็นเวลา 4 ชั่วโมง หากต้องการให้สามารถทำนายให้ทราบล่วงหน้าควรทดลองเพิ่มเวลาของชุดข้อมูลเป็น 1 ถึง 2 วัน เพื่อพัฒนาโมเดลให้สามารถทราบการซ้ำชุดก่อนที่จะเกิดการเสียหายที่รุนแรงและมีเวลาในการวางแผนงานซ่อมบำรุงช่วยลดค่าใช้จ่ายในการปฏิบัติงาน

5. แนวทางในการพัฒนาเพื่อนำไปปรับใช้กับงานจริง เพื่อเป็นการพัฒนาประสิทธิภาพการตรวจจับเพิ่มขึ้นควรพัฒนาโมเดลเป็น 2 ขั้นตอน โดยขั้นตอน 1 ใช้สำหรับตรวจหาสิ่งผิดปกติแบ่งเป็น 2 คลาส เป็นคลาสปกติและผิดปกติ ขั้นตอน 2 ทำนายหาการซ้ำชุดของตลับลูกปืนเพื่อหาสาเหตุการซ้ำชุดของตลับลูกปืน

6. นอกจากเรื่องของตลับลูกปืน ยังสามารถนำการเรียนรู้ของเครื่องแบบ Supervise learning ไปขยายผลไปยังเรื่องการตรวจจับความผิดปกติของกระแสมอเตอร์ หรือเครื่องมือวัดทางอุตสาหกรรมได้ เพื่อช่วยจัดการงานซ่อมบำรุงให้มีประสิทธิภาพเพิ่มขึ้น

7. การสร้างระบบโดยการเรียนรู้ของเครื่อง สามารถสร้างระบบ Digital twin เพื่อเป็นการจำลองเครื่องจักร หรืออุปกรณ์ สำหรับทำการทดลองปรับพารามิเตอร์ ก่อนนำไปปรับใช้กับหน้างานจริง หรือเพื่อใช้สำหรับเฝ้าดูสุขภาพของการทำงานเครื่องจักรนั้น เป็นเครื่องมือช่วยหาสาเหตุของการซ้ำชุดของตลับลูกปืนได้

บรรณานุกรม

- Gousseau, W., Girardin, F., & Griffaton, J. (2016). Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati.
- Hanly, S. (2021). Top 12 Vibration Metrics to Monitor & How to Calculate Them.
- J. Lee, H. Q., Yu, G., Lin, J., IMS, R. T. S., & Cincinnati., U. o. (2007). Experiments on bearings. The data set was provided by the Center for Intelligent Maintenance Systems (IMS), University of Cincinnati. สืบค้นจาก <https://ti.arc.nasa.gov/tech/dash/groups/pcoc/prognostic-data-repository/#bearing>
- Kankar, P. K., Sharma, S. C., & Harsha, S. P. (2011). Fault diagnosis of ball bearings using machine learning methods. *Expert Systems with Applications*, 38(3), 1876-1886.
- Kim, H.-E., Tan, A. C. C., Mathew, J., Kim, E. Y. H., & Choi, B.-K. (2012). Machine Prognostics Based on Health State Estimation Using SVM *Asset Condition, Information Systems and Decision Models* (Chapter 9 pp. 169-186).
- LEGA Corporation. (2016). Measure vibration of machine in manufacturing process. สืบค้นจาก <https://legatool.com/wp/85/>
- Martinez-Rego, D., Fontenla-Romero, O., & Alonso-Betanzos, A. (2011). *Power wind mill fault detection via one-class & SVM vibration signal analysis*. Paper presented at the The 2011 International Joint Conference on Neural Networks.
- MathWorks, T. (2021). signal statistical feature
- Medium. (2021). Confusion matrix. สืบค้นจาก https://miro.medium.com/max/875/1*fxiTNlgOyvAombPJx5KGeA.png
- Miltadis, K. (2018). IMS Bearing Dataset Extracting Failure modes from vibration signals. สืบค้นจาก http://mkalikatzarakis.eu/wp-content/uploads/2018/12/IMS_dset.html
- Mosallam, A., Medjaher, K., & Zerhouni, N. (2013). Nonparametric time series modelling for industrial prognostics and health management. *The International Journal of Advanced Manufacturing Technology*, 69.
- Ren, L., Sun, Y., Wang, H., & Zhang, L. (2018). Prediction of Bearing Remaining Useful Life

- With Deep Convolution Neural Network. *IEEE Access*, 6, 13041-13049.
- scikit-learn. (2022). K-fold validation สืบค้นจาก https://scikit-learn.org/stable/images/grid_search_cross_validation.png
- Suh, S., Jang, J., Won, S., Jha, M. S., & Lee, Y. O. (2020). Supervised Health Stage Prediction Using Convolutional Neural Networks for Bearing Wear. *Sensors (Basel)*, 20(20).
- Suh, S., Lee, H., Jo, J., Lukowicz, P., & Lee, Y. (2019). Generative Oversampling Method for Imbalanced Data on Bearing Fault Detection and Diagnosis. *Applied Sciences*, 9(4).
- Wikipedia. (2021a). Diagram Random forest. https://en.wikipedia.org/wiki/Random_forest
- Wikipedia. (2021b). การเรียนรู้ของเครื่อง. สืบค้นจาก <https://th.wikipedia.org/wiki/การเรียนรู้ของเครื่อง>
- Wordpress. (2021a). classify of Support vector machine. สืบค้นจาก https://en.wikipedia.org/wiki/Random_forest
- Wordpress. (2021b). flow operate of gradient boosting algorithm. สืบค้นจาก <https://littleml.files.wordpress.com/2017/03/boosted-trees-process.png?w=810>
- Zhang, R., Peng, Z., Wu, L., Yao, B., & Guan, Y. (2017). Fault Diagnosis from Raw Sensor Data Using Deep Neural Networks Considering Temporal Coherence. *Sensors (Basel)*, 17(3).
- นที ไทยยาธรรม. (2019). K-Nearest Neighbors สืบค้นจาก <https://www.glurgeek.com/education/knn/>

ประวัติผู้เขียน

ชื่อ-สกุล	วีรชิต ตั้งศิริวัฒนวงศ์
วัน เดือน ปี เกิด	03 พฤษภาคม 2531
สถานที่เกิด	จ.นครราชสีมา
วุฒิการศึกษา	วิศวกรรมศาสตรบัณฑิต มหาวิทยาลัยธรรมศาสตร์
ที่อยู่ปัจจุบัน	200/301 หมู่ 8 ต.คูคต อ.ลำลูกกา จ.ปทุมธานี

