

การจำแนกสาเหตุความผิดปกติในระบบจำหน่าย 33 kV ของ การไฟฟ้าส่วนภูมิภาค โดยใช้การ  
เรียนรู้ของเครื่อง



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต  
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า  
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย  
ปีการศึกษา 2563  
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

FAULT CAUSE CLASSIFICATION ON 33 kV DISTRIBUTION SYSTEM OF THE PROVINCIAL  
ELECTRICITY AUTHORITY USING MACHINE LEARNING



A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Engineering in Electrical Engineering

Department of Electrical Engineering

FACULTY OF ENGINEERING

Chulalongkorn University

Academic Year 2020

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การจำแนกสาเหตุความผิดปกติพ่วงในระบบจำหน่าย 33 kV
โดย	ของ การไฟฟ้าส่วนภูมิภาค โดยใช้การเรียนรู้ของเครื่อง
สาขาวิชา	นายวิษณุ พรหมรัตน์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	วิศวกรรมไฟฟ้า
	ศาสตราจารย์ ดร.วาทีต เบญจพลกุล

---

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง  
ของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

.....	คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)	
คณะกรรมการสอบวิทยานิพนธ์	
.....	ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ชาญณรงค์ บาลมงคล)	
.....	อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ศาสตราจารย์ ดร.วาทีต เบญจพลกุล)	
.....	กรรมการ
(รองศาสตราจารย์ ดร.สุรชัย ชัยทัศน์ีย์)	
.....	กรรมการภายนอกมหาวิทยาลัย
(ดร.ชัยพร เขมะภาคะพันธ์)	

วิทยุ พรหมรัตน์ : การจำแนกสาเหตุความผิดปกติพ่วงในระบบจำหน่าย 33 kV ของ การไฟฟ้าส่วนภูมิภาค โดยใช้การเรียนรู้ของเครื่อง. ( FAULT CAUSE CLASSIFICATION ON 33 kV DISTRIBUTION SYSTEM OF THE PROVINCIAL ELECTRICITY AUTHORITY USING MACHINE LEARNING) อ.ที่ปรึกษาหลัก : ศ. ดร.วาทิต เภยจุพลกุล

วิทยานิพนธ์นี้นำเสนอกระบวนการจำแนกสาเหตุความผิดปกติพ่วงที่เกิดในระบบจำหน่ายของการไฟฟ้าส่วนภูมิภาค (กฟภ.) ได้แก่ กลุ่มของสัตว์ ต้นไม้ และอุปกรณ์ชำรุด โดยใช้ข้อมูลจากการวิเคราะห์รูปคลื่นแรงดันและกระแสที่บันทึกได้จากรีเลย์ที่สถานีไฟฟ้า เพื่อหาค่าแรงดันตกชั่วขณะ อัตราการเปลี่ยนแปลงของกระแสและแรงดัน ค่าสูงสุดของกระแสและแรงดันนิวัตร์ ความผิดปกติพ่วงแบบชั่วคราวหรือถาวร ชนิดของความผิดปกติพ่วง และความผิดปกติพ่วงที่มีการแปรเปลี่ยนประเภท จากนั้นให้การเรียนรู้ของเครื่อง เปรียบเทียบกับวิธีโครงข่ายประสาทเทียมจำแนกสาเหตุของความผิดปกติพ่วงเพื่อลดระยะเวลาในการตรวจสอบเคเบิลไลน์ระบบจำหน่ายเมื่อเกิดเหตุผิดปกติพ่วงและใช้ในการวางแผนการจัดการงานแก้ไขไฟฟ้าขัดข้อง ทดสอบความแม่นยำของอัลกอริทึมโดยการทดสอบจริงกับระบบจำหน่ายของ กฟภ. ในพื้นที่ภาคใต้ตอนกลางของประเทศไทย ผลการทดสอบกระบวนการที่นำเสนอแสดงว่า การเรียนรู้ของเครื่องสามารถจำแนกได้มีความถูกต้องมากกว่าโครงข่ายประสาทเทียมที่ข้อมูลฝึกฝน จำนวน 200 เหตุการณ์ โดยสามารถจำแนกว่าเหตุผิดปกติพ่วงอยู่ในกลุ่มใดที่ 72.72% และสามารถระบุถึงสาเหตุความผิดปกติพ่วง ด้วยความแม่นยำที่ 81.21% ถึง 88.89%

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

สาขาวิชา วิศวกรรมไฟฟ้า  
ปีการศึกษา 2563

ลายมือชื่อนิสิต .....  
ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

# # 6170406421 : MAJOR ELECTRICAL ENGINEERING

KEYWORD: Fault Cause Classification, Machine Learning, ANN, KNN, SVM

Witsanu Promrat : FAULT CAUSE CLASSIFICATION ON 33 kV DISTRIBUTION SYSTEM OF THE PROVINCIAL ELECTRICITY AUTHORITY USING MACHINE LEARNING. Advisor: Prof. Watit Benjapolakul, Ph.D.

This Thesis presents a process for classifying the cause of faults occurring in the Provincial Electricity Authority (PEA) distribution systems, consisting of a group of animals, tree contact, and equipment failure. By using data from the analysis of current and voltage waveforms obtained from relay installed at the substation to find the voltage dip, rate of change of current and voltage, neutral current and voltage, temporary or permanent fault, fault type, evolving fault, fed data to machine learning (ML) compared to the artificial neural network (ANN) to classify the cause of faults in order to reduce the duration to patrol the line of the distribution system after the fault occurs and to use in the management planning for resolving electrical interruption. Accuracy testing of the algorithm by real testing was performed with the recorded data from the distribution system of PEA in the central southern area of Thailand. The test result of the proposed process shows that the ML can classify more accuracy than the ANN with 200 training data, the group of fault events is correctly classified with the ML accuracy of 72.72%, the proposed method can provide the accuracy of identifying fault cause with 81.21% - 88.89% accuracy.

Field of Study: Electrical Engineering

Student's Signature .....

Academic Year: 2020

Advisor's Signature .....

## กิตติกรรมประกาศ

การที่วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีนั้นมาจากความช่วยเหลืออย่างมากของบุคคลหลายท่าน ซึ่งผู้วิจัยรู้สึกขอบคุณและซาบซึ้งเป็นอย่างมากจากหัวใจ

ขอขอบคุณศาสตราจารย์ ดร. วาทิต เบญจพลกุล อาจารย์ที่ปรึกษาที่ได้ให้คำแนะนำต่างๆอันเป็นประโยชน์ต่อวิทยานิพนธ์ฉบับนี้ ทำให้ผู้วิจัยเข้าใจถึงหลักการของการทำงานวิจัยเป็นอย่างดี

ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร.ชาญณรงค์ บาลมงคล รองศาสตราจารย์ ดร.สุรชัย ชัยทัศนีย์ และ อาจารย์ ดร.ชัยพร เขมมะภาตะพันธ์ ที่ให้เกียรติมาเป็นกรรมการวิทยานิพนธ์ และชี้แนะแนวทางอันเป็นประโยชน์ต่องานวิจัยนี้

ขอขอบคุณผู้บริหารของ กฟผ. และคณะกรรมการทุกท่านที่คัดเลือกและให้โอกาสผู้วิจัยเป็นผู้ได้รับทุนสมาร์ตกริด ของ กฟผ. รวมทั้งผู้บริหารของ กฟจ. ระเบียบ ผู้อำนวยการเขต กฟต.2 ที่อนุญาตให้เข้ามาศึกษา ความรู้ที่ได้จากหลักสูตรและงานวิจัยนี้ผู้วิจัยจะนำไปใช้ให้เกิดประโยชน์กับองค์กรให้มากที่สุดและสัญญาว่าจะให้โอกาสกับผู้อื่นเช่นเดียวกับที่ผู้วิจัยได้รับมา

สุดท้ายขอขอบพระคุณ แม่ ป้า และลุงของผู้วิจัย ผู้ที่เป็นครอบครัว เป็นหัวใจ เป็นกำลังใจ สนับสนุนผู้วิจัยในทุกๆด้าน เป็นทุกเบื้องหลังในทุกๆความสำเร็จของผู้วิจัย

วิษณุ พรหมรัตน์

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ .....	ง
กิตติกรรมประกาศ .....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ช
สารบัญรูป .....	ฎ
บทที่ 1 บทนำ .....	1
1.1 แนวเหตุผลในการทำวิทยานิพนธ์.....	1
1.2 วัตถุประสงค์ของวิทยานิพนธ์.....	3
1.3 ขอบเขตของวิทยานิพนธ์.....	3
1.4 ขั้นตอนการดำเนินงาน.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	4
บทที่ 2 ความรู้พื้นฐานและหลักการที่เกี่ยวข้อง.....	5
2.1 กระบวนการหาสาเหตุความผิดพลาดของ กฟภ. ในปัจจุบัน .....	5
2.2 สาเหตุการเกิดความผิดพลาดในระบบจำหน่ายไฟฟ้า.....	6
2.3 ข้อมูลที่บันทึกได้เมื่อเกิดความผิดพลาด .....	7
2.4 การสกัดคุณลักษณะเด่น .....	11
2.5 การเลือกคุณลักษณะเด่น (Feature Selection).....	18
2.6 เครื่องมือที่ใช้ในการจำแนก .....	20
2.7 การประเมินความถูกต้องของโมเดล .....	33
บทที่ 3 การออกแบบระบบและพัฒนาระบบ.....	39

3.1 ภาพรวมของระบบ.....	40
3.2 รายละเอียดของระบบการทำงาน .....	41
3.3 ข้อมูลที่นำมาใช้ในการวิเคราะห์.....	42
3.4 การสกัดคุณลักษณะเด่น .....	43
3.5 การเลือกคุณลักษณะเด่น .....	46
3.6 Training Model .....	47
3.7 การประเมินความถูกต้องของโมเดล .....	49
3.8 การนำไปใช้งาน .....	50
บทที่ 4 การพัฒนาและผลการทดสอบ .....	51
4.1 การพัฒนาขั้นตอนวิธีการจำแนกด้วย Machine Learning .....	51
4.2 การพัฒนาขั้นตอนวิธีการจำแนกด้วย Artificial Neural Network.....	65
4.3 ผลการคัดเลือกคุณลักษณะเด่นและโมเดลที่เหมาะสม.....	79
4.4 ผลการทดสอบกับเหตุการณ์จริง .....	80
4.5 การวิเคราะห์ผล .....	86
บทที่ 5 สรุปผล.....	88
5.1 สรุปผลการวิจัย.....	88
5.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต.....	89
บรรณานุกรม.....	90
ประวัติผู้เขียน .....	94



## สารบัญตาราง

	หน้า
ตารางที่ 2-1 ผลการจำแนกประเภทด้วยค่า K ต่างๆ [16].....	30
ตารางที่ 2-2 ตัวอย่าง Confusion Matrix [17] .....	34
ตารางที่ 4-1 การคัดเลือกคุณลักษณะเด่นของการจำแนกเป็นกลุ่มด้วยวิธี Decision Tree .....	52
ตารางที่ 4-2 การคัดเลือกคุณลักษณะเด่นของการจำแนกเป็นกลุ่มด้วยวิธี Linear Discriminant ...	52
ตารางที่ 4-3 การคัดเลือกคุณลักษณะเด่นของการจำแนกเป็นกลุ่มด้วยวิธี SVM .....	53
ตารางที่ 4-4 การคัดเลือกคุณลักษณะเด่นของการจำแนกเป็นกลุ่มด้วยวิธี KNN.....	53
ตารางที่ 4-5 การคัดเลือกคุณลักษณะเด่นของการจำแนกเป็นกลุ่มด้วยวิธี Ensemble .....	54
ตารางที่ 4-6 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี Decision Tree .....	55
ตารางที่ 4-7 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี Linear Discriminant.....	55
ตารางที่ 4-8 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี SVM .....	56
ตารางที่ 4-9 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี KNN .....	56
ตารางที่ 4-10 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี Ensemble	57
ตารางที่ 4-11 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มสัตว์ด้วยวิธี Decision Tree	58
ตารางที่ 4-12 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มสัตว์ด้วยวิธี Linear Discriminant.....	58
ตารางที่ 4-13 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มสัตว์ด้วยวิธี SVM.....	59
ตารางที่ 4-14 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มสัตว์ด้วยวิธี KNN.....	59
ตารางที่ 4-15 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มสัตว์ด้วยวิธี Ensemble .....	60
ตารางที่ 4-16 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มต้นไม้ด้วยวิธี Decision Tree .....	61



ตารางที่ 4-40 การคัดเลือกคุณลักษณะเด่นที่ 8 โหนดซ่อน ของการจำแนกภายในกลุ่มสัตว์ .....	73
ตารางที่ 4-41 การคัดเลือกคุณลักษณะเด่นที่ 9 โหนดซ่อน ของการจำแนกภายในกลุ่มสัตว์ .....	74
ตารางที่ 4-42 การคัดเลือกคุณลักษณะเด่นที่ 10 โหนดซ่อน ของการจำแนกภายในกลุ่มสัตว์ .....	74
ตารางที่ 4-43 การคัดเลือกคุณลักษณะเด่นที่ 4 โหนดซ่อน ของการจำแนกภายในกลุ่ม ต้นไม้.....	75
ตารางที่ 4-44 การคัดเลือกคุณลักษณะเด่นที่ 5 โหนดซ่อน ของการจำแนกภายในกลุ่ม ต้นไม้.....	76
ตารางที่ 4-45 การคัดเลือกคุณลักษณะเด่นที่ 6 โหนดซ่อน ของการจำแนกภายในกลุ่ม ต้นไม้.....	76
ตารางที่ 4-46 การคัดเลือกคุณลักษณะเด่นที่ 7 โหนดซ่อน ของการจำแนกภายในกลุ่ม ต้นไม้.....	76
ตารางที่ 4-47 การคัดเลือกคุณลักษณะเด่นที่ 8 โหนดซ่อน ของการจำแนกภายในกลุ่ม ต้นไม้.....	77
ตารางที่ 4-48 การคัดเลือกคุณลักษณะเด่นที่ 9 โหนดซ่อน ของการจำแนกภายในกลุ่ม ต้นไม้.....	77
ตารางที่ 4-49 การคัดเลือกคุณลักษณะเด่นที่ 10 โหนดซ่อน ของการจำแนกภายในกลุ่ม ต้นไม้.....	77
ตารางที่ 4-50 ผลความแม่นยำของแต่ละจำนวนเหตุการณ์ที่ใช้จำแนกในแต่ละโมเดล .....	78
ตารางที่ 4-51 ผลการเลือกคุณลักษณะเด่นและอัลกอริทึมที่เหมาะสมของ Machine Learning.....	80
ตารางที่ 4-52 ผลการเลือกคุณลักษณะเด่นและจำนวนโหนดซ่อนเหมาะสมของ ANN .....	80
ตารางที่ 4-53 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกสาเหตุเป็นกลุ่มๆด้วยวิธี ML.....	82
ตารางที่ 4-54 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี ML.	83
ตารางที่ 4-55 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มสัตว์ด้วยวิธี ML .....	83
ตารางที่ 4-56 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มต้นไม้ด้วยวิธี ML ....	83
ตารางที่ 4-57 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกสาเหตุเป็นกลุ่มๆด้วยวิธี ANN ..	84
ตารางที่ 4-58 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี ANN .....	84
ตารางที่ 4-59 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มสัตว์ด้วยวิธี ANN ....	85
ตารางที่ 4-60 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มต้นไม้ด้วยวิธี ANN ..	85

## สารบัญรูป

	หน้า
รูปที่ 2-1 ขั้นตอนการหาสาเหตุของเหตุการณ์ไฟฟ้าขัดข้อง .....	5
รูปที่ 2-2 แผนภูมิแสดงสาเหตุหลักของความผิดปกติในพื้นที่ กฟภ. เขต 2 ภาคใต้ ปี 2560-2562 ..	7
รูปที่ 2-3 ตัวอย่างรายละเอียดการบันทึกรายละเอียดข้อมูลการเกิดความผิดปกติ .....	8
รูปที่ 2-4 ตัวอย่างรูปคลื่นกระแสขณะเกิดความผิดปกติ .....	9
รูปที่ 2-5 ตัวอย่างรูปคลื่นแรงดันขณะเกิดความผิดปกติ .....	10
รูปที่ 2-6 ตัวอย่างแสดงการทำงานของรีเลย์ป้องกัน .....	10
รูปที่ 2-7 ตัวอย่างไฟล์ข้อมูลที่ได้จากการบันทึกของรีเลย์ .....	11
รูปที่ 2-8 ตัวอย่างจำนวนเฟสที่ได้รับผลกระทบ (Number of Phases affected) [15].....	15
รูปที่ 2-9 ระยะเวลาจัดความผิดปกติ (Fault Clearing Time) [15].....	16
รูปที่ 2-10 การแปลงเวฟเลตของรูปคลื่นแรงดันจากสาเหตุฟ้าผ่า [15].....	17
รูปที่ 2-11 การแปลงเวฟเลตของรูปคลื่นแรงดันจากสาเหตุต้นไม้ [15].....	17
รูปที่ 2-12 การเลือกคุณลักษณะเด่นด้วยวิธี Sequential Forward Selection.....	19
รูปที่ 2-13 การเลือกคุณลักษณะเด่นด้วยวิธี Backward Elimination .....	19
รูปที่ 2-14 โครงข่ายประสาทเทียมแบบชั้นเดียว (Single – Layer Neural Networks).....	22
รูปที่ 2-15 โครงข่ายประสาทเทียมแบบหลายชั้น (Multi – Layer Neural Networks).....	23
รูปที่ 2-16 โครงสร้างของระบบผู้เชี่ยวชาญ.....	24
รูปที่ 2-17 ตัวอย่างต้นไม้ตัดสินใจ [16].....	25
รูปที่ 2-18 ซัพพอร์ตเวกเตอร์ (Support Vector) [16] .....	26
รูปที่ 2-19 ความสัมพันธ์ระหว่างระยะขอบกับเวกเตอร์ปกติของระนาบเกิน [16].....	27
รูปที่ 2-20 พื้นที่แสดงอาณาเขตแต่ละคลาสในกรณีไบนารี [16].....	28
รูปที่ 2-21 ปริภูมิตัวอย่าง [16] .....	29

รูปที่ 2-22 ระยะทางจากตัวอย่างที่ค่า K เท่ากับ 1, 3 และ 6 [16].....	31
รูปที่ 2-23 การเรียนรู้แบบรวมกลุ่มด้วยเทคนิค Voting Method.....	32
รูปที่ 2-24 การเรียนรู้แบบรวมกลุ่มด้วยเทคนิค Bagging Method.....	33
รูปที่ 2-25 การประเมินความถูกต้องด้วยวิธี Holdout [17].....	37
รูปที่ 2-26 การประเมินความถูกต้องด้วยวิธี 5-fold cross validation .....	38
รูปที่ 3-1 ภาพรวมแนวคิดของการจำแนกสาเหตุความผิดปกติในระบบจำหน่าย .....	40
รูปที่ 3-2 ไดอะแกรมแสดงรายละเอียดของระบบการทำงาน .....	41
รูปที่ 3-3 ไดอะแกรมขั้นตอนการจำแนกสาเหตุความผิดปกติในระบบจำหน่ายของ กฟภ.....	42
รูปที่ 3-4 ตัวอย่างข้อมูลรูปคลื่นกระแส และแรงดันเมื่อแปลงเป็นไฟล์นามสกุล .csv .....	43
รูปที่ 3-5 ตัวอย่างความผิดปกติชนิดต่าง ๆ.....	45
รูปที่ 3-6 กระบวนการของ Machine Learning (The Process of Machine Learning) .....	48
รูปที่ 3-7 ตัวอย่างการจำแนกสาเหตุด้วย Artificial Neural Network โดยโปรแกรม Rapid Miner .....	49
รูปที่ 4-1 ค่าความแม่นยำในการจำแนกเป็นกลุ่มของแต่ละอัลกอริทึมที่คุณลักษณะเด่นต่างกัน .....	54
รูปที่ 4-2 ค่าความแม่นยำในการจำแนกภายในกลุ่มอุปกรณ์ของแต่ละอัลกอริทึม .....	57
รูปที่ 4-3 ค่าความแม่นยำในการจำแนกภายในกลุ่มสัตว์ของแต่ละอัลกอริทึม .....	60
รูปที่ 4-4 ค่าความแม่นยำในการจำแนกภายในกลุ่มต้นไม้ของแต่ละอัลกอริทึม .....	63
รูปที่ 4-5 ค่าความแม่นยำในการจำแนกของแต่ละจำนวนข้อมูลฝึกฝนในแต่ละโมเดล .....	64
รูปที่ 4-6 ค่าความแม่นยำในการจำแนกเป็นกลุ่มของแต่ละจำนวนโนดซ่อน .....	68
รูปที่ 4-7 ค่าความแม่นยำในการจำแนกภายในกลุ่มอุปกรณ์ ของแต่ละจำนวนโนดซ่อน .....	71
รูปที่ 4-8 ค่าความแม่นยำในการจำแนกภายในกลุ่มของสัตว์ ของแต่ละจำนวนโนดซ่อน.....	75
รูปที่ 4-9 ค่าความแม่นยำในการจำแนกภายในกลุ่มของต้นไม้ ของแต่ละจำนวนโนดซ่อน.....	78
รูปที่ 4-10 ค่าความแม่นยำในการจำแนกของแต่ละจำนวนข้อมูลฝึกฝนในแต่ละโมเดล.....	79
รูปที่ 4-11 ไดอะแกรมขั้นตอนการจำแนกด้วยโมเดลที่สร้างจาก Machine Learning.....	81

รูปที่ 4-12 ไดอะแกรมขั้นตอนการจำแนกด้วยโมเดลที่สร้างจาก Artificial Neural Network ..... 82



## บทที่ 1

### บทนำ

#### 1.1 แนวเหตุผลในการทำวิทยานิพนธ์

ในระบบจำหน่ายของการไฟฟ้าส่วนภูมิภาค (กฟภ.) การส่งพลังงานไฟฟ้าจากสถานีไฟฟ้าไปยังโหลดหรือผู้ใช้ไฟฟ้าจำเป็นต้องจ่ายกระแสไฟฟ้าอย่างต่อเนื่องเพื่อให้ระบบมีเสถียรภาพและความเชื่อถือได้สูง และหากเกิดเหตุขัดข้องกับระบบจำหน่ายแล้ว นอกจากทำให้ กฟภ. ต้องสูญเสียรายได้จากโอกาสในการขายไฟแล้ว สถานประกอบการ โรงงานอุตสาหกรรมต่างๆ ยังได้รับความเสียหายด้วย เช่น การที่เครื่องจักรต้องหยุดทำงานเนื่องจากมีไฟดับทำให้สินค้าที่กำลังผลิตเกิดการเสียหายหรือในย่านเมืองท่องเที่ยวที่มีโรงแรม ถนนคนเดิน สถานบันเทิงยามค่ำคึกคัก พลุกพล่านไปด้วยนักท่องเที่ยว ซึ่งเมื่อเกิดเหตุการณ์ไฟดับอาจทำให้นักท่องเที่ยวรู้สึกถึงความไม่ปลอดภัยทั้งในชีวิตและทรัพย์สินจึงเลือกที่จะหลีกเลี่ยงการท่องเที่ยวในพื้นที่นั้น ซึ่งความเสียหายดังกล่าวอาจไม่สามารถประเมินค่าออกมาเป็นตัวเลขได้

เมื่อมีเหตุการณ์ไฟดับเกิดขึ้นจำเป็นต้องมีการตรวจสอบหาสาเหตุ ดำเนินการแก้ไขและจ่ายกระแสไฟฟ้าคืนให้รวดเร็วที่สุด ซึ่งกระบวนการในการตรวจสอบหาสาเหตุ นั้น นอกจากการทราบถึงตำแหน่งที่เกิดความผิดปกติ (Fault Location) ที่ทำให้ผู้ปฏิบัติงานสามารถเข้าตรวจสอบไปยังตำแหน่งที่เกิดความผิดปกติได้รวดเร็วขึ้นแล้ว การทราบถึงสาเหตุของความผิดปกติ (Fault Cause) ได้ล่วงหน้า นั้น จะทำให้ผู้ปฏิบัติงานที่ตรวจสอบเคเบิลไลน์ (Patrol Man) นั้นสามารถพกสเพื่อหาสาเหตุได้รวดเร็วขึ้น ทำให้สามารถดำเนินการแก้ไขและจ่ายกระแสไฟฟ้าคืนได้รวดเร็วยิ่งขึ้น [1] ส่งผลให้ระยะเวลาของการเกิดไฟดับน้อยลง และนอกจากนั้นยังเป็นข้อมูลให้กับพนักงานประจำศูนย์ควบคุมการจ่ายไฟ (Dispatching Center) ใช้ในการประเมินและวางแผนจัดลำดับความสำคัญของเหตุการณ์ในการสั่งการตรวจสอบและแก้ไขได้เช่น ในกรณีเกิดความผิดปกติพร้อมกัน 2 จุดและเป็นไลน์ที่มีความสำคัญเท่ากัน แต่ขณะนั้นมีเพียง 1 ทีม ที่สามารถเข้าดำเนินการแก้ไขได้ หากทราบถึงสาเหตุล่วงหน้าทำให้สามารถประเมินระยะเวลาในการแก้ไขได้ ซึ่งพนักงานของศูนย์ควบคุมการจ่ายไฟอาจพิจารณาความสำคัญโดยให้เข้าดำเนินการแก้ไขกับเหตุการณ์ที่ใช้เวลาในการแก้ไขน้อยก่อน และจากประสบการณ์การที่ผู้วิจัยเคยปฏิบัติงานประจำศูนย์สั่งการแก้ไฟฟ้าขัดข้อง (Outage Management Center) ผู้วิจัยเคยพบเหตุการณ์ที่มีการทริปล็อกเอาต์ (Trip LockOut) ของเซอร์กิตเบรกเกอร์ป้องกันไลน์ (Circuit Breaker) และเมื่อรถแก้ไฟฟ้าขัดข้อง (Patrol Car) ตรวจสอบเคเบิลไลน์กลับไม่พบสาเหตุจึงดำเนินการทดลองจ่ายเซอร์กิตเบรกเกอร์ของวงจรดังกล่าว ผลปรากฏว่าไม่

สามารถจ่ายได้เนื่องจากยังคงมีความผิดปกติอยู่ในระบบจำหน่ายจึงต้องตรวจสอบซ้ำอีกครั้ง และยังไม่พบสาเหตุเหมือนเดิม จึงต้องทดลองจ่ายเซอร์กิตเบรกเกอร์เป็นครั้งที่ 2 ซึ่งยังจ่ายไม่ได้เหมือนเดิม จนกระทั่งมีผู้แจ้งว่าได้ยินเสียงระเบิดบริเวณลูกถ้วย ตรวจสอบจึงพบว่ามี การรวบไฟ (Flash Over) บริเวณผิวลูกถ้วยทำให้ลูกถ้วยชำรุดเกิดการลัดวงจรลงดิน จากเหตุการณ์ในครั้งนั้นจึงทำให้ผู้วิจัยได้ตระหนักว่านอกจากการทราบระยะทางของตำแหน่งที่เกิดความผิดปกติที่ได้จากการรายงานของรีเลย์แล้ว หากเราสามารถทราบถึงสาเหตุของความผิดปกตินั้นล่วงหน้าด้วยแล้ว ก็จะสามารถช่วยลดระยะเวลาที่เกิดไฟฟ้าดับลงได้

ระบบจำหน่ายของ กฟภ. มีการตั้งค่าการสับต่อวงจรกลับอัตโนมัติ (Automatic Reclosing) เนื่องจากความผิดปกติส่วนใหญ่ประมาณมากกว่า 70% เป็นการลัดวงจรไฟฟ้าแบบชั่วคราว (Temporary Faults) [2] ซึ่งสามารถสับต่ออัตโนมัติหลังจากเคลียร์ความผิดปกติไปแล้ว และจากข้อมูลเหตุการณ์ไฟฟ้าขัดข้องของ กฟภ. พบว่าเมื่อมีเหตุการณ์ Trip Reclose ส่วนใหญ่จะไม่ทราบสาเหตุของความผิดปกติที่เข้ามาในระบบจำหน่าย [3] เมื่อไม่ทราบถึงสาเหตุที่ชัดเจนจึงไม่ได้รับการแก้ไขหรือปรับปรุง ซึ่งบางสาเหตุนั้นแม้จะเป็นการลัดวงจรแบบชั่วคราวแต่ยังอาจมีโอกาสกลับมาอีกครั้งหากไม่ได้รับการแก้ไขเช่น กิ่งไม้โยกเอนจากแรงลมมาสัมผัสกับสายระบบจำหน่าย หากไม่ตัดกิ่งไม้ก็มีโอกาสเกิดซ้ำอีก เมื่อเกิดซ้ำก็เกิดเหตุการณ์ Trip Reclose ซ้ำอีกครั้ง ซึ่งทำให้เกิดไฟกระพริบ เป็นการจ่ายไฟไม่ต่อเนื่องเชิงคุณภาพไฟฟ้าส่งผลถึงความเชื่อถือได้ของระบบจำหน่ายของ กฟภ. ซึ่งหากสามารถทราบถึงสาเหตุของความผิดปกติแบบชั่วคราวที่ทำให้เกิดการ Trip Reclose ได้ ก็สามารถหาแนวทางป้องกันไม่ให้เกิดเหตุการณ์ไฟกระพริบจากการ Trip Reclose ซ้ำได้อีก

จากงานวิจัยในอดีตมีการใช้วิธีจำแนกสาเหตุความผิดปกติทั้งในระบบส่งและระบบจำหน่าย โดยสามารถแบ่งในส่วนของข้อมูลนำเข้าออกได้เป็น 2 กลุ่ม 1. กลุ่มที่ใช้ข้อมูลจากสถานะแวดล้อมภายนอกขณะเกิดความผิดปกติมาเป็นตัวแปรในการวิเคราะห์เช่น ข้อมูลสภาพอากาศ เวลาขณะเกิดความผิดปกติ สภาพพื้นที่ที่เกิดความผิดปกติ [4-7] 2. กลุ่มที่ใช้ข้อมูลจากสัญญาณทางไฟฟ้าหรือรูปคลื่นกระแสและแรงดันขณะเกิดความผิดปกติมาสกัดคุณลักษณะเด่น [8-10] จากนั้นใช้อัลกอริทึมประเภทที่สามารถใช้แก้ปัญหา Classification เป็นเครื่องมือในการจำแนกสาเหตุ

จากที่ได้กล่าวมาทั้งหมดนั้นจึงทำให้ผู้วิจัยต้องการหาวิธีการที่จะทำให้ผู้วิจัยสามารถทราบถึงสาเหตุของความผิดปกติได้ล่วงหน้าก่อนที่จะมีการเข้าตรวจสอบหรือทันทีที่มีการทริปของเซอร์กิตเบรกเกอร์ และในปัจจุบันรีเลย์ที่ใช้ในระบบจำหน่ายของ กฟภ. มีความสามารถในการบันทึกค่าสัญญาณของกระแสและสัญญาณของแรงดันเมื่อมีการกวนระบบ (Disturbance Record) ในระบบจำหน่าย ประกอบกับผู้วิจัยต้องการที่จะพัฒนาอัลกอริทึมจากข้อมูลและทรัพยากรที่ กฟภ. มี จึงเป็นที่มาของการพัฒนาอัลกอริทึมเพื่อจำแนกสาเหตุของความผิดปกติในระบบจำหน่ายของ กฟภ. โดยใช้ข้อมูลของรูปคลื่น



ดังนั้นวิทยานิพนธ์ฉบับนี้จึงนำเสนอแนวคิดในการพัฒนาอัลกอริทึมเพื่อจำแนกสาเหตุของความผิดปกติในระบบจำหน่ายของ กฟภ. ด้วยการวิเคราะห์รูปคลื่น เพื่อช่วยให้ผู้ปฏิบัติงานสามารถประเมินสถานการณ์และลดระยะเวลาในการตรวจสอบความผิดปกติในระบบไฟฟ้า ส่งผลให้สามารถแก้ไขความผิดปกติได้รวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น

## 1.2 วัตถุประสงค์ของวิทยานิพนธ์

- 1) เพื่อศึกษาการจำแนกสาเหตุของความผิดปกติในระบบจำหน่ายไฟฟ้าด้วยการวิเคราะห์คุณสมบัติของรูปคลื่น
- 2) เพื่อสามารถพัฒนาอัลกอริทึมในการจำแนกสาเหตุของความผิดปกติในระบบจำหน่ายไฟฟ้าของ กฟภ. ด้วยการวิเคราะห์คุณสมบัติของรูปคลื่น
- 3) เพื่อลดระยะเวลาในการตรวจสอบความผิดปกติในระบบไฟฟ้า
- 4) เพื่อเป็นข้อมูลเพิ่มเติมให้กับผู้ปฏิบัติงานในการประเมินและตัดสินใจในการแก้ไขสถานการณ์ได้อย่างมีประสิทธิภาพยิ่งขึ้น

## 1.3 ขอบเขตของวิทยานิพนธ์

- 1) พิจารณาความผิดปกติจากสาเหตุในกลุ่มของ อุปกรณ์ชำรุด สัตว์ และต้นไม้ เนื่องจากเป็น 3 สาเหตุหลักที่ทำให้เกิดความผิดปกติในพื้นที่การไฟฟ้าส่วนภูมิภาค เขต 2 ภาคใต้
- 2) ศึกษาในระบบจำหน่ายของการไฟฟ้าส่วนภูมิภาค ที่ระดับแรงดันขนาด 33 kV. โดยการบันทึกของ Disturbance Record และ รายงานไฟฟ้าขัดข้องพื้นที่ เขต 2 ภาคใต้ ปี 2560-2562 ประกอบการพิจารณา
- 3) คัดเลือกคุณลักษณะเด่นด้วยวิธี Sequential Forward Selection เนื่องจากเป็นการวัดประสิทธิภาพการทำงานของโมเดลโดยเลือกค่าที่มีความแม่นยำสูงสุด
- 4) ใช้โปรแกรม Matlab และ Rapid Miner เป็นเครื่องมือหลักในการวิเคราะห์และออกแบบอัลกอริทึม

## 1.4 ขั้นตอนการดำเนินงาน

- 1) กำหนดวัตถุประสงค์และขอบเขตของการวิจัย
- 2) ศึกษางานวิจัยในอดีตที่เกี่ยวข้องกับการจำแนกและทำนายสาเหตุของความผิดปกติในระบบไฟฟ้ากำลัง
- 3) ศึกษารายละเอียดและข้อจำกัดของข้อมูลรูปคลื่นขณะเกิดความผิดปกติและการบันทึกรายละเอียดสาเหตุของความผิดปกติในอดีตจนถึงปัจจุบันของ กฟภ.

- 4) ศึกษาขั้นตอน วิธีการ กระบวนการ การจำแนกสาเหตุของความผิดปกติวิธีต่างๆเพื่อนำมาออกแบบอัลกอริทึมและปรับใช้ให้เหมาะสมกับข้อมูลและทรัพยากรที่ กฟภ. มีอยู่
- 5) ศึกษาการใช้งานโปรแกรม Matlab และ Rapid Miner เพื่อนำมาเป็นเครื่องมือที่ใช้ในการวิเคราะห์
- 6) ออกแบบและเขียนโปรแกรมเพื่อสร้างอัลกอริทึมที่ใช้สกัดคุณลักษณะเด่นของแต่ละสาเหตุของความผิดปกติ
- 7) นำเอาคุณลักษณะเด่นมาวิเคราะห์และพัฒนาอัลกอริทึมในการจำแนกสาเหตุของความผิดปกติในระบบจำหน่าย 33 kV ของ กฟภ. เขต 2 ภาคใต้
- 8) ทดสอบผล วิเคราะห์จุดบกพร่อง และปรับปรุงเพื่อเพิ่มความแม่นยำให้มากขึ้น
- 9) วิเคราะห์และสรุปผลการทดสอบของงานวิจัย
- 10) เรียบเรียงผลงานวิจัยและนำเสนอคณะกรรมการต่อไป

### 1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1) อัลกอริทึมที่พัฒนาสามารถจำแนกสาเหตุของความผิดปกติในระบบจำหน่ายของ กฟภ. โดยใช้ข้อมูลจากรูปคลื่นได้
- 2) เมื่อเกิดเหตุการณ์ไฟฟ้าขัดข้องผู้ปฏิบัติงานสามารถทราบถึงสาเหตุของความผิดปกติทันที ช่วยให้สามารถประเมินสถานการณ์และเพื่อช่วยลดระยะเวลาในการตรวจสอบได้
- 3) สามารถป้องกันความผิดปกติจากสาเหตุต่างๆได้
- 4) สามารถลดการสูญเสียรายได้จากโอกาสในการขายไฟของ กฟภ. และผู้ใช้ไฟฟ้าจากการที่สามารถลดระยะเวลาที่เกิดไฟดับได้
- 5) สามารถลดรายจ่ายของ กฟภ. ที่ใช้ในการตรวจสอบระบบจำหน่ายเช่น ต้นทุนค่าน้ำมันของยานพาหนะที่ใช้ในการตรวจสอบ ค่าล่วงเวลาพนักงาน เป็นต้น

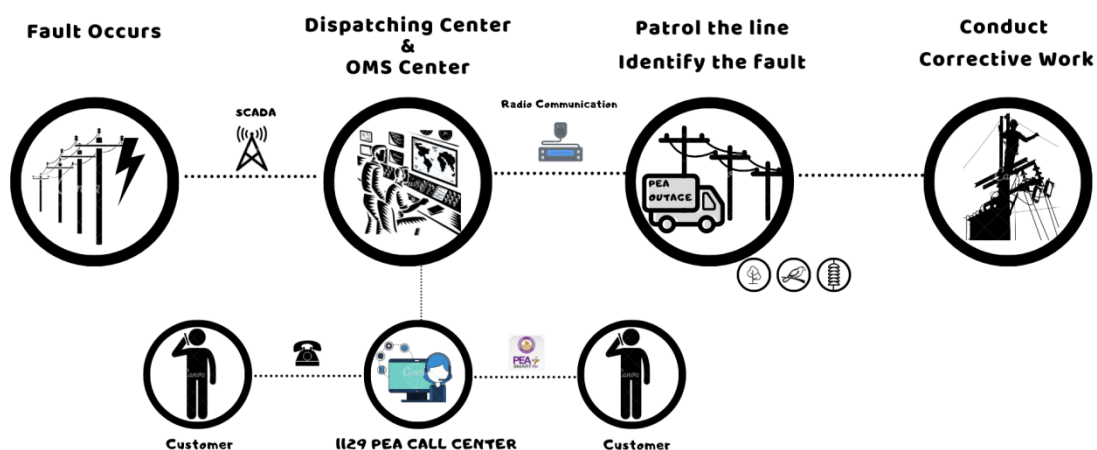
## บทที่ 2

### ความรู้พื้นฐานและหลักการที่เกี่ยวข้อง

แนวเหตุผลและทฤษฎีที่เกี่ยวข้องประกอบไปด้วยหลักการ แนวคิด ข้อมูลพื้นฐาน งานวิจัยในอดีต ที่เกี่ยวข้องกับการจำแนกสาเหตุของความผิดปกติในระบบจำหน่าย และสามารถให้คำตอบเมื่อเกิดความผิดปกติในอนาคตได้ว่าเกิดจากสาเหตุอะไร ซึ่งจะประกอบไปด้วย กระบวนการหาสาเหตุความผิดปกติที่ กฟภ. ปฏิบัติอยู่ในปัจจุบัน, สาเหตุการเกิดความผิดปกติในระบบจำหน่ายของ กฟภ., ข้อมูลที่บันทึกได้เมื่อเกิดความผิดปกติในระบบจำหน่ายของ กฟภ., การสกัดคุณลักษณะเด่นจากข้อมูลเพื่อบ่งชี้ถึงเอกลักษณ์ของสาเหตุ, การคัดเลือกคุณลักษณะเด่น, การประยุกต์เครื่องมือที่ใช้ในการจำแนกสาเหตุความผิดปกติ, และการทดสอบความแม่นยำของอัลกอริทึม

#### 2.1 กระบวนการหาสาเหตุความผิดปกติของ กฟภ. ในปัจจุบัน

กระบวนการในการหาสาเหตุของความผิดปกติที่ กฟภ. ปฏิบัติอยู่ในปัจจุบัน สามารถอธิบายภาพรวมได้ดังรูปที่ 2-1 และสามารถแบ่งออกเป็นข้อๆ ได้ดังต่อไปนี้ [1]



รูปที่ 2-1 ขั้นตอนการหาสาเหตุของเหตุการณ์ไฟฟ้าขัดข้อง

1) เมื่อเกิดความผิดปกติในระบบจำหน่ายของ กฟภ. อุปกรณ์ป้องกันที่ถูกติดตั้งที่สถานีไฟฟ้าจะปลด (Trip) เพื่อตัดวงจรที่เกิดความผิดปกติดังกล่าวออกและจะส่งข้อมูลรายละเอียด เช่น เวลาที่เกิดการ Trip, เฟสที่เกิดความผิดปกติ, ระยะทางระหว่างตำแหน่งที่เกิดความผิดปกติกับสถานีไฟฟ้า ไปยังศูนย์ควบคุมการจ่ายไฟ (ศจฟ.) ผ่านระบบ Supervisory Control And Data Acquisition (SCADA)

2) ผู้ใช้ไฟฟ้าในพื้นที่ที่ได้รับผลกระทบจะโทรศัพท์แจ้งไปยังศูนย์บริการผู้ใช้ไฟฟ้า (1129 PEA CALL CENTER) หรือผ่าน PEA Smart Plus Application ซึ่งข้อมูลที่รับแจ้งจะประกอบไปด้วย บริเวณที่เกิดไฟฟ้าดับ และอาจรวมถึงข้อมูลสาเหตุที่เกิดไฟดับ (กรณีที่ผู้ใช้ไฟฟ้าอยู่ในเหตุการณ์ขณะเกิดเหตุ) จากนั้น 1129 PEA CALL CENTER จะประสานไปยังศูนย์ปฏิบัติการระบบไฟฟ้าในพื้นที่ และศูนย์ปฏิบัติการฯ จะแจ้งไปยังศูนย์ควบคุมการจ่ายไฟ เพื่อรวบรวมข้อมูลและวางแผนการตรวจสอบแก้ไขต่อไป

3) ศูนย์ควบคุมการจ่ายไฟจะวางแผนและสั่งการให้รถแก่กระแสไฟฟ้าขัดข้องในพื้นที่ ตรวจสอบเคเบิลไลน์เพื่อหาสาเหตุ

4) เมื่อรถแก่กระแสไฟฟ้าขัดข้องพบสาเหตุและจุดที่เกิดความผิดปกติ จะรายงานกลับไปยังศูนย์ควบคุมการจ่ายไฟ เพื่อบันทึกข้อมูล และ รถแก่กระแสไฟฟ้าขัดข้องดำเนินการแก้ไขตามสาเหตุความผิดปกติที่เกิดขึ้นและจ่ายกระแสไฟฟ้าคืนสู่สภาวะปกติให้รวดเร็วที่สุด

จะเห็นได้ว่าเมื่อเกิดเหตุการณ์ไฟฟ้าขัดข้องในปัจจุบันกระบวนการหาสาเหตุของความผิดปกติจะได้ออกจากการแจ้งเหตุจากผู้ใช้ไฟฟ้าซึ่งเป็นเพียงการสันนิษฐานคร่าวๆเท่านั้น ทำให้กระบวนการแก่กระแสไฟฟ้าขัดข้องยังไม่สามารถปฏิบัติงานได้อย่างเต็มประสิทธิภาพ ซึ่งหากมีการพัฒนาอัลกอริทึมเพื่อนำมาหาสาเหตุเมื่อเกิดความผิดปกติจะช่วยให้อาจลดระยะเวลาในการตรวจสอบหาสาเหตุ ทำให้สามารถจ่ายไฟคืนสู่สภาวะปกติได้รวดเร็วขึ้น

## 2.2 สาเหตุการเกิดความผิดปกติในระบบจำหน่ายไฟฟ้า

ระบบจำหน่ายไฟฟ้าเหนือพื้นดินของ กฟภ. เขต 2 ภาคใต้ มีสาเหตุที่ทำให้เกิดความผิดปกติได้หลายสาเหตุโดยแบ่งเป็นกลุ่มได้ดังนี้ [3]

1) สาเหตุเกิดจากการล้มเหลวของการฉนวนของอุปกรณ์ เช่น ลูกถ้วยแตกร้าว กับดักเสิร์จ ชำรุด เป็นต้น

2) สาเหตุจากต้นไม้สัมผัสสายไฟฟ้า เช่น ยางพารา ต้นปาล์ม ทางมะพร้าว เป็นต้น

3) สาเหตุจากสัตว์ เช่น นก งู ลิงลม กระรอก เป็นต้น

4) สาเหตุจากภัยธรรมชาติ เช่น พายุ ลมพายุพัดทำให้สายไฟฟ้าแกว่งไปสัมผัสกับแผ่นป้ายโฆษณา เป็นต้น

5) สาเหตุจากยานพาหนะ หรือรถยนต์ชนเสาไฟฟ้า

6) สาเหตุอื่น ๆ

โดยตั้งแต่ปี 2560-2562 พบสาเหตุความผิดปกติดังแสดงในรูปที่ 2-2



รูปที่ 2-2 แผนภูมิแสดงสาเหตุหลักของความผิดปกติในพื้นที่ กฟภ. เขต 2 ภาคใต้ ปี 2560-2562

### 2.3 ข้อมูลที่บันทึกได้เมื่อเกิดความผิดปกติ

ข้อมูลที่สามารถบันทึกได้เมื่อเกิดเหตุการณ์ความผิดปกติของ กฟภ. สามารถแบ่งออกเป็น 2 กลุ่ม

#### 2.3.1 ข้อมูลที่บันทึกโดยผู้ปฏิบัติงาน (Operator) ประจำศูนย์ควบคุมการจ่ายไฟ

เมื่อเกิดความผิดปกติในระบบจำหน่ายของ กฟภ. จะมีการบันทึกรายละเอียดต่าง ๆ ของแต่ละเหตุการณ์ผ่านโปรแกรม e-Respond (โปรแกรมจัดการงานแก้ไขไฟฟ้าขัดข้องของ กฟภ.) เพื่อเก็บเป็นสถิติและนำข้อมูลมาใช้วิเคราะห์เพื่อวางแผนป้องกันไม่ให้เกิดซ้ำในอนาคตโดยข้อมูลสามารถออกรายงานเป็นรูปแบบไฟล์นามสกุล .xlsx ดังรูปที่ 2-3

29 พฤษภาคม 2562, 8:06:47																
50.รายงานเหตุการณ์ไฟฟ้าขัดข้อง																
การไฟฟ้าเขต: การไฟฟ้าเขต2(นครศรีธรรมราช) ภาค 4																
จาก 28 พฤษภาคม 2562, 0:00 น. ถึง 28 พฤษภาคม 2562, 23:59 น.																
หมายเลข เหตุการณ์	ลำดับ	วันที่/เวลา ไฟดับ	เวลา	วันที่/เวลา ที่ จ่ายไฟคืน ระบบได้ครั้งแรก	วันที่/เวลา ที่ จ่ายไฟคืน ครบทั้งหมด	รวม เวลา ไฟดับ (น.)	รหัส อุปกรณ์ที่ ทำงาน	พีดเดอร์	ประเภท การ ทำงาน	สาเหตุ/ รายละเอียด	ทราบสาเหตุ	กฟผ.รับผิดชอบ	สภาพอากาศ	мыш. ถูก กระทบ (ราย)	สถานที่จุด เกิดเหตุ	ประเภท เหตุการณ์
3425758582	1	28-May-62	1:49:00	28 พ.ค. 2562, 2:23:00	28 พ.ค. 2562, 2:23:00	34	TSA09R-01	TSA09	T/L	ลิ่ง/ค้าง	สันนิษฐาน	กฟผ.ทุ่งใหญ่	อากาศปกติ	11,614	ส.ลิ่งและสาย	ไฟฟ้าขัดข้อง
3410691823	2	28-May-62	9:09:00	28 พ.ค. 2562, 17:16:00	28 พ.ค. 2562, 17:16:00	487	TYI04YB-01	-	Operate	ปรับปรองอื่นๆ	ทราบสาเหตุ	กฟผ. นานอน	อากาศปกติ	0	ปรับปรอง ระบบจำหน่าย	แผนดับไฟ
3421079125	3	28-May-62	9:14:00	28 พ.ค. 2562, 14:23:00	28 พ.ค. 2562, 14:23:00	309	TYI07R-01	TYI07	Operate	ปรับปรองอื่นๆ	ทราบสาเหตุ	กฟผ. นานอน	อากาศปกติ	0	ปรับปรอง ระบบจำหน่าย	แผนดับไฟ
3421090856	4	28-May-62	9:18:00	28 พ.ค. 2562, 16:07:00	28 พ.ค. 2562, 16:07:00	409	TRA01R-03	TRA01	Operate	ปรับปรองอื่นๆ	ทราบสาเหตุ	กฟผ. สีดา	อากาศปกติ	2,114	ปรับปรอง ระบบจำหน่าย	แผนดับไฟ
3425946345	5	28-May-62	9:21:00	28 พ.ค. 2562, 9:21:00	28 พ.ค. 2562, 9:21:00	1	KCD02VB-01	KCD02	T/R	นก	สันนิษฐาน	กฟผ.กาญจนบุรี	อากาศปกติ	0	ส.นกและสาย	ไฟฟ้าขัดข้อง
3425946443	6	28-May-62	9:36:00	28 พ.ค. 2562, 10:09:32	28 พ.ค. 2562, 10:09:32	34	BKO03VB-01	BKO03	T/L	งู	ทราบสาเหตุ	กฟผ.บ้านเขาเหล็ก	อากาศปกติ	0	มือครูชั้น และสายแรง สูงทำ D/F ใส่ช้อน หมอมใส่กาจิ้ง	ไฟฟ้าขัดข้อง
3426199954	7	28-May-62	10:30:00	28 พ.ค. 2562, 10:58:00	28 พ.ค. 2562, 10:58:00	28	LRA02R-01	LRA02	T/L	งู	สันนิษฐาน	กฟผ.ห้วยยอด	อากาศปกติ	10,720	ส.งูและสาย	ไฟฟ้าขัดข้อง
3426214172	8	28-May-62	12:30:00	28 พ.ค. 2562, 12:50:00	28 พ.ค. 2562, 12:50:00	20	TSA09R-03	TSA09	T/L	งู	สันนิษฐาน	กฟผ.บางขัน	อากาศปกติ	5,299	ส.งูและสาย	ไฟฟ้าขัดข้อง

### รูปที่ 2-3 ตัวอย่างรายละเอียดการบันทึกรายละเอียดข้อมูลการเกิดความผิดปกติ

หมายเลขเหตุการณ์ เพื่อระบุ (Identification) และเป็นเลขอ้างอิงแต่ละเหตุการณ์ไม่ให้ซ้ำกัน

ลำดับ เพื่อลำดับเหตุการณ์ความผิดปกติในแต่ละวัน

วันที่ไฟดับ เพื่อระบุ วัน/เดือน/ปี ที่เกิดเหตุการณ์ความผิดปกติ

เวลา คือ เวลาที่เกิดเหตุการณ์ความผิดปกติ

วันที่/เวลา ที่จ่ายไฟคืนระบบได้ครั้งแรก คือ เวลาครั้งแรกที่สามารถจ่ายไฟคืนระบบได้บางส่วน

วันที่/เวลา ที่จ่ายไฟคืนครบทั้งหมด คือ เวลาที่สามารถจ่ายไฟคืนได้ครบทุกพื้นที่

รหัสอุปกรณ์ที่ทำงาน คือ อุปกรณ์ที่มีการทำงาน ซึ่งภายในรหัสอุปกรณ์จะบ่งบอกถึง พื้นที่ที่เกิดไฟดับ, อุปกรณ์อะไร เช่น เซอร์คิตเบรกเกอร์ที่สถานีไฟฟ้าหรือรีโคลสเซอร์ที่ติดตั้งในไลน์ระบบจำหน่าย, สถานีไฟฟ้าไหน, พีดเดอร์อะไร

พีดเดอร์ เพื่อระบุว่า สถานีไฟฟ้าอะไร พีดเดอร์อะไร ที่เกิดความผิดปกติ

ประเภทการทำงาน จะบอกว่าอุปกรณ์ที่ทำงานนั้น ทำงานแบบ Trip Lock Out (T/L), Trip Reclose (T/R) หรือ การสั่งการทำงานจาก ศจพ. (Operate)

สาเหตุ/รายละเอียด คือสาเหตุที่ทำให้เกิดความผิดปกติเช่น งู นก ต้นไม้แตะสาย

**ทราบสาเหตุ** คือ การบ่งบอกว่า ทราบสาเหตุของความผิดปกติของแบบชัดเจน หรือเป็นการสันนิษฐานสาเหตุเนื่องจากเมื่อตรวจสอบเคลียร์ไลน์แต่ไม่พบสาเหตุ

**กฟพ. รับผิดชอบ** คือ พื้นที่ที่เกิดความผิดปกติอยู่ในพื้นที่รับผิดชอบของการไฟฟ้าอะไร

**สภาพอากาศ** คือ ณะที่เกิดความผิดปกติมีสภาพอากาศเป็นอย่างไร ประกอบไปด้วย อากาศปกติ, ฝนตก, ลมแรง, ฝนตกลมแรง, ฝนตกฟ้าคะนอง (พายุ)

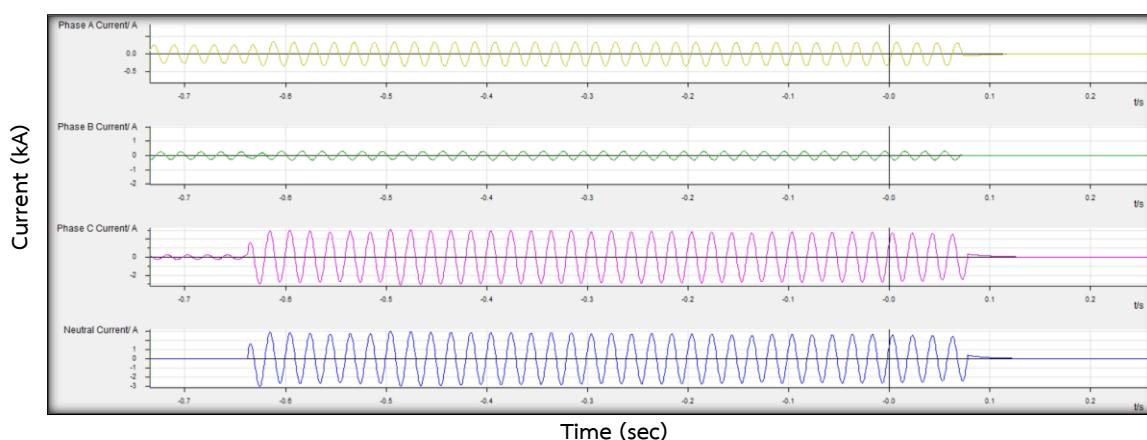
**ผพพ.ถูกกระทบ (ราย)** คือ จำนวนผู้ใช้ไฟฟ้าที่ได้รับผลกระทบจากเหตุการณ์ความผิดปกตินี้ๆ

**สถานที่จุดเกิดเหตุ** คือ การระบุพื้นที่ที่เกิดความผิดปกติเพิ่มเติมเช่น บ้านอะไร ใกล้กับอะไร

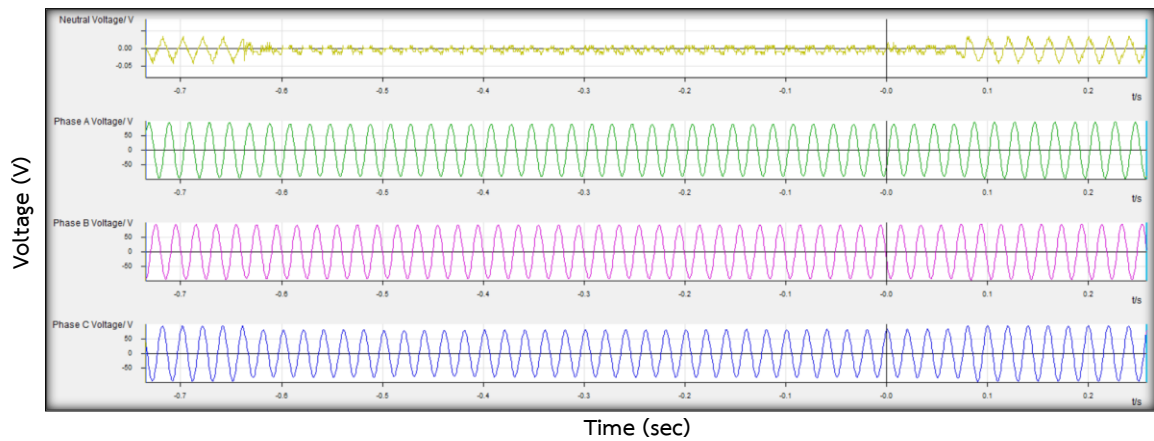
**ประเภทเหตุการณ์** คือ ในการเกิดไฟฟ้าดับแต่ละครั้งก็จะมีทั้งการดับจากแผนดับไฟ และดับจากเหตุการณ์ไฟฟ้าขัดข้อง (ความผิดปกติ)

### 2.3.2 ข้อมูลที่ถูกบันทึกเป็นสัญญาณทางไฟฟ้าโดยรีเลย์ที่ติดตั้งที่สถานีไฟฟ้า (Disturbance Record)

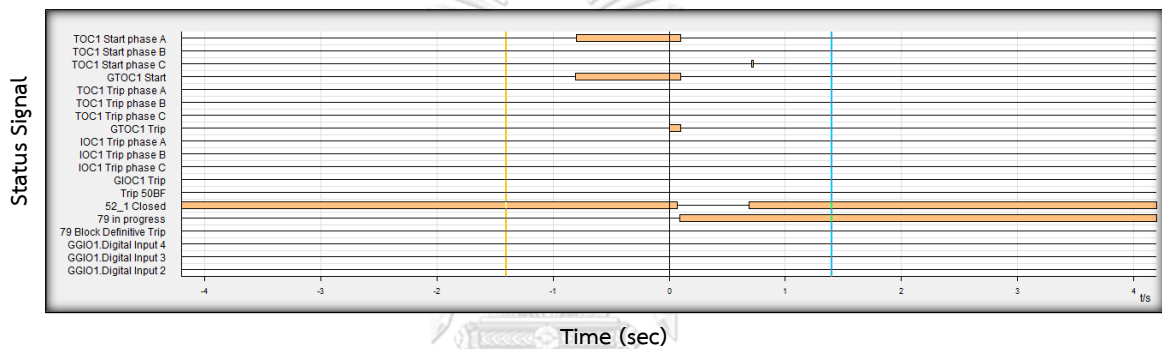
เมื่อมีการรบกวนระบบ (System Disturbance) รีเลย์เป็นอุปกรณ์ป้องกันระบบไฟฟ้าซึ่งถูกติดตั้งที่สถานีไฟฟ้า ทำหน้าที่สั่ง Trip ไปยังเซอร์กิตเบรกเกอร์เมื่อระบบไฟฟ้ามีความผิดปกติเกิดขึ้นตามที่ได้ตั้งค่าไว้ (setting) ซึ่งแต่ละรุ่นจะมีฟังก์ชันหลักและฟังก์ชันเสริมที่ต่างกัน ซึ่งรีเลย์รุ่น INGEPAC EF-MD ยี่ห้อ Ingeteam [11] ที่ กฟพ. เขต 2 ภาคใต้นำมาใช้ติดตั้งที่สถานีไฟฟ้า ได้แก่ สถานีไฟฟ้า เกาะสมุย 1 (KMA), นครศรีฯ 2 (NTB), ปากพั่น (PPN), กาญจนดิษฐ์ (KCD), อ่าวลึก (ALA), ท้ายเหมือง (TMG), และ ป่าตอง (PAT) รวม 7 สถานีไฟฟ้า เป็นรุ่นที่สามารถบันทึกสัญญาณรูปคลื่นกระแสและแรงดันขณะเกิดความผิดปกติได้ โดยจะบันทึกเมื่อมีการ Trip ของเซอร์กิตเบรกเกอร์ ตัวอย่างข้อมูลรูปคลื่นที่บันทึกได้ ดังแสดงในรูปที่ 2-4, 2-5 และ 2-6



รูปที่ 2-4 ตัวอย่างรูปคลื่นกระแสขณะเกิดความผิดปกติ



รูปที่ 2-5 ตัวอย่างรูปคลื่นแรงดันขณะเกิดความผิดปกติ



รูปที่ 2-6 ตัวอย่างแสดงการทำงานของรีเลย์ป้องกัน

โดยค่าที่บันทึกมี 37 ช่องดังรูปที่ 2-7 ที่ความถี่ 50 Hz อัตราสุ่ม 144 ที่ 10,800 จุด ตลอดช่วงสัญญาณ ประกอบด้วย 2 ส่วนหลักคือ

- 1) ส่วนที่เป็นสัญญาณแอนะล็อก 18 ช่อง จะถูกเก็บค่าในรูปคลื่น sinusoidal ประกอบด้วยขนาดของ กระแส แรงดัน ทั้ง 3 เฟส กระแสนิวทรัล แรงดันนิวทรัล ฯลฯ
- 2) ส่วนที่เป็นสัญญาณดิจิทัล 19 ช่อง จะอยู่ในรูปสัญญาณดิจิทัล มีเพียงค่า 0 และ 1 ประกอบด้วยการทำงานของอุปกรณ์ป้องกันในฟังก์ชันต่างๆ เช่น การเริ่มหน่วงเวลาทำงาน (TOC1 Start) การสั่งทริปหลังจากหน่วงเวลามาแล้ว (TOC1 Trip) การทริปทันทีทันใด (IOC1 Trip) ฯลฯ



รูปที่ 2-7 ตัวอย่างไฟล์ข้อมูลที่ได้จากการบันทึกของรีเลย์

## 2.4 การสกัดคุณลักษณะเด่น

การจะทราบถึงสาเหตุของความผิดปกติจำเป็นต้องนำข้อมูลที่บันทึกได้มาหาคุณลักษณะเด่นที่สามารถบ่งชี้ถึงสาเหตุของความผิดปกติที่สนใจได้ และข้อมูลที่มีบางข้อมูลอาจไม่มีความจำเป็นต่อการบ่งชี้ถึงสาเหตุของความผิดปกติ ทำให้การประมวลผลล่าช้าและอาจส่งผลให้การบ่งชี้สาเหตุนั้นผิดพลาดได้

ค่าที่สามารถบ่งชี้ถึงเอกลักษณ์ของสาเหตุความผิดปกติที่สนใจเรียกว่าคุณลักษณะเด่น (Feature) และการดึงค่าคุณลักษณะเด่นจากข้อมูลที่บันทึกได้เรียกว่า การสกัดคุณลักษณะเด่น (Feature Extraction) โดยหลายงานวิจัยในอดีตได้มีกระบวนการสกัดคุณลักษณะเด่นเพื่อใช้ในการจำแนกสาเหตุของความผิดปกติที่แตกต่างกัน ซึ่งสามารถแบ่งค่าการสกัดคุณลักษณะเด่นนั้นออกได้เป็น 2 กลุ่ม คือ ข้อมูลจากสภาวะแวดล้อมภายนอก และข้อมูลจากสัญญาณทางไฟฟ้า

### 2.4.1 ข้อมูลจากสภาวะแวดล้อมภายนอก

#### 1) วงจรที่เกิดความผิดปกติ (Circuit ID)

ในแต่ละวงจรการจ่ายไฟจะจ่ายโหลดไปยังพื้นที่ต่างๆ ต่างกัน ซึ่งในแต่ละพื้นที่นั้นอาจมีสาเหตุของความผิดปกติมาจากสาเหตุและตำแหน่งเดิม ใน [6, 7, 12] ได้ใช้วงจรที่เกิดความผิดปกติ (Circuit ID) เป็นคุณลักษณะเด่นในการจำแนกสาเหตุความผิดปกติในระบบจำหน่ายที่สนใจ ข้อดีของคุณลักษณะเด่นของค่านี้อาจสามารถนำมาใช้ได้โดยตรงไม่ต้องผ่านกระบวนการที่ซับซ้อน แต่ข้อเสียคือหากมีการเปลี่ยนแปลงวงจรการจ่ายไฟก็ต้องเก็บข้อมูลใหม่

## 2) สภาพอากาศ (Weather)

บางสาเหตุของความผิดปกติของอากาศมีแนวโน้มเกิดมากในบางสภาพอากาศ เช่น สาเหตุจาก ภูเขาไฟ ส่วนใหญ่จะเกิดขึ้นในวันที่มีสภาพอากาศปกติหรือค่อนข้างไปทางร้อน แต่วันก่อนหน้านั้นมีสภาพอากาศ ฝนตก หรือ สาเหตุจากฟ้าผ่า อาจมีแนวโน้มเกิดในสภาพอากาศ ฝนตก ฟ้าคะนอง [3]

งานวิจัยในอดีต [6, 7, 12] ได้แบ่งสภาพอากาศออกเป็น 9 แบบ ดังนี้ 1. ปกติ (Fair) 2. หนาวเย็น (Cold) 3. ฝนตก (Rain) 4. ลมแรง (Wind) 5. ฟ้าคะนองและลมแรง (Wind and Lightning) 6. ฟ้าคะนอง (Lightning) 7. ลูกเห็บตก (Hail) 8. ทิมะตก (Snow) 9. หนาว (Cold) ซึ่งคุณลักษณะเด่นของค่านี้สามารถนำมาใช้ได้โดยไม่ต้องผ่านกระบวนการที่ซับซ้อนเช่นเดียวกับวงจรที่เกิดความผิดปกติ แต่ในบางครั้งอาจมีการคลาดเคลื่อนในการบันทึกสาเหตุ เช่น ขณะเกิดความผิดปกติ บริเวณพื้นที่ที่เกิดความผิดปกติอาจมีลมแรง แต่เมื่อผู้ปฏิบัติงานที่ตรวจสอบเคลียร์ไลน์ (Patrol Man) เข้าตรวจสอบพบว่าไม่มีลมแล้วจึงรายงานว่าสภาพอากาศปกติ

## 3) ฤดูกาล (Season)

ลักษณะเด่นอีกอย่างหนึ่งที่มีส่วนบ่งชี้ได้ว่าคุณลักษณะเด่นของความผิดปกติมีแนวโน้มมาจากสาเหตุอะไรมักจะขึ้นอยู่กับฤดูกาลด้วย เช่น สาเหตุจากสัตว์มักเกิดในฤดูร้อนมากที่สุด โดยงานวิจัยใน [6, 7, 12] ได้แบ่ง ฤดูกาล ออกเป็น 4 ฤดูกาลดังนี้ 1. ฤดูใบไม้ผลิ (Spring) 2. ฤดูร้อน (Summer) 3. ฤดูใบไม้ร่วง (Fall) 4. ฤดูหนาว (Winter) แต่ทั้งนี้อาจปรับเปลี่ยนตามลักษณะของฤดูกาลแต่ละพื้นที่ เช่น ประเทศไทย อาจสามารถแบ่งเป็น 3 ฤดู 1. ฤดูร้อน (Summer) 2. ฤดูฝน (Rainy) 3. ฤดูหนาว (Winter)

## 4) วันของสัปดาห์ (Day of Week)

ตามหลักสากลสามารถแบ่งวันของสัปดาห์ได้เป็น 7 วัน และงานวิจัยใน [6, 7, 12] กำหนดเป็นหมายเลขแทนวันต่างๆ ดังนี้ หมายเลข 1. แทนวันอาทิตย์, 2. แทนวันจันทร์, 3. แทนวันอังคาร, 4. แทนวันพุธ, 5. แทนวันพฤหัสบดี, 6. แทนวันศุกร์, 7. แทนวันเสาร์ แต่คุณลักษณะเด่นดังกล่าวนี้ไม่เป็นที่ยอมรับเนื่องจากในหลายงานวิจัยพบว่าไม่มีผลที่เชื่อมโยงกับสาเหตุของความผิดปกติ และผลการทดลองคุณลักษณะเด่นดังกล่าวไม่ได้ส่งผลให้การจำแนกสาเหตุมีความแม่นยำมากขึ้น

## 5) ช่วงเวลาของวัน (Time of Day)

จากงานวิจัย [6, 7, 12] ได้นำช่วงเวลาของวันมาเป็นคุณลักษณะเด่นค่าหนึ่ง ในการจำแนกสาเหตุของความผิดปกติโดยแบ่งเป็น 2 ลักษณะ คือ

1. แบ่งเป็น 2 ช่วงเวลา คือ ช่วงเวลากลางวัน, กลางคืน
2. แบ่งเป็น 4 ช่วงเวลา คือ ช่วงเวลาเช้า, บ่าย, ค่ำ, ดึก

## 2.4.2 ข้อมูลจากสัญญาณทางไฟฟ้า

ข้อมูลสัญญาณทางไฟฟ้าจะได้รับมาจากเครื่องมือวัดที่สามารถบันทึกสัญญาณต่างๆ ทางไฟฟ้าเช่น กระแสไฟฟ้า แรงดันไฟฟ้า สถานะของอุปกรณ์ป้องกัน ขณะที่เกิดความผิดปกติได้ ตัวอย่างเช่น ในประเทศมาเลเซีย [13] ได้มีการติดตั้ง Digital Fault Recorder (DFR) ที่ระบบส่งเพื่อนำข้อมูลสัญญาณทางไฟฟ้ามาวิเคราะห์ และ ประเทศสหรัฐอเมริกา [9, 14] ได้ใช้เครื่องมือวัดคุณภาพไฟฟ้า (PQ-Meter) ติดตั้งที่ระบบจำหน่ายเพื่อวัดค่าสัญญาณทางไฟฟ้า โดยข้อมูลสัญญาณทางไฟฟ้าที่ได้จำเป็นต้องผ่านกระบวนการบางอย่างเพื่อสกัดเอาคุณลักษณะเด่นที่สามารถบ่งชี้ถึงสาเหตุของความผิดปกติที่สนใจได้ และเพื่อเป็นการลดขนาดของข้อมูลทำให้ระยะเวลาในการประมวลผลมีความรวดเร็วยิ่งขึ้น โดยงานวิจัยในอดีตได้มีการสกัดลักษณะเด่นทางไฟฟ้าดังนี้

### 1) อัตราการเปลี่ยนแปลงกระแสเทียบกับเวลา (Rate of Change of Current Curve)

[13] ใช้อัตราการเปลี่ยนแปลงของกระแสเทียบกับเวลา เป็นคุณลักษณะเด่นในการจำแนกความผิดปกติจากฟ้าผ่าและต้นไม้ ฟังก์ชันของสัญญาณรูปคลื่นกระแสอยู่ในรูปค่าเฉลี่ยกำลังสอง (Root Mean Square, RMS) โดยเสนออัตราการเปลี่ยนแปลงของกระแส RMS ดังสมการที่ (2-1)

$$f_c(t, \lambda) = 1 - e^{-\lambda t} \Big|_{t=t_0-t_n} \quad (2-1)$$

เมื่อ  $f_c(t, \lambda)$  คือ ฟังก์ชันสัญญาณกระแสในรูป RMS

$t$  คือ เวลาที่เกิดความผิดปกติ

$t_0$  คือ เวลาเริ่มต้นของการเกิดความผิดปกติ

$t_n$  คือ เวลาสิ้นสุดของการเกิดความผิดปกติ

$\lambda$  คือ อัตราการเปลี่ยนแปลงของกระแสเทียบกับเวลา

โดยเมื่อวัดความคลาดเคลื่อนของสัญญาณกระแสที่ได้จากการคำนวณเทียบกับสัญญาณกระแสที่วัดได้จาก DFR จะได้ดังสมการที่ (2-2)

$$\Delta f_d = f_m(t) - f_c(t, \lambda) \Big|_{t=t_0-t_n} \quad (2-2)$$

เมื่อ  $\Delta f_d$  คือ ค่าความคลาดเคลื่อน

$f_m(t)$  คือ ค่าที่วัดได้จาก DFR ในรูป RMS

$f_c(t, \lambda)$  คือ ค่าที่ได้จากการคำนวณ

## 2) ค่าแรงดันตกชั่วขณะ (Voltage Dip)

ค่าแรงดันตกชั่วขณะคำนวณจากค่าความแตกต่างของแรงดันบัสในเวลาปกติกับค่าแรงดันบัสขณะเกิดความผิดปกติสามารถคำนวณดังสมการที่ (2-3) ใน [13] ใช้ค่าแรงดันตกชั่วขณะเป็นคุณลักษณะเด่นร่วมกับอัตราการเปลี่ยนแปลงกระแสเทียบกับเวลา ในการจำแนกความผิดปกติจากฟ้าผ่าและต้นไม้

$$\text{Voltage dip} = \frac{V_p - V_{f,\min}}{V_p} \times 100 \quad (2-3)$$

เมื่อ  $V_p$  = ค่าแรงดันก่อนความผิดปกติ

$V_{f,\min}$  = ค่าแรงดันต่ำสุดระหว่างเกิดความผิดปกติ

## 3) ความผิดปกติแบบชั่วคราวและความผิดปกติแบบถาวร (Temporary Fault or Permanent Fault)

การทำงานของอุปกรณ์ตัดต่อกลับอัตโนมัติหรือรีโคลสเซอร์ (Recloser) สามารถแบ่งออกได้เป็น 2 กลุ่มคือ

1. Recloser สั่งการให้เซอร์กิตเบรกเกอร์ Trip Reclose สำเร็จ เรียกว่า ความผิดปกติแบบชั่วคราว (ระบบสามารถกำจัดสาเหตุของความผิดปกติได้)

2. Recloser สั่งการให้เซอร์กิตเบรกเกอร์ Trip Reclose ไม่สำเร็จ เรียกว่าความผิดปกติแบบถาวร (ระบบไม่สามารถกำจัดสาเหตุความผิดปกติได้)

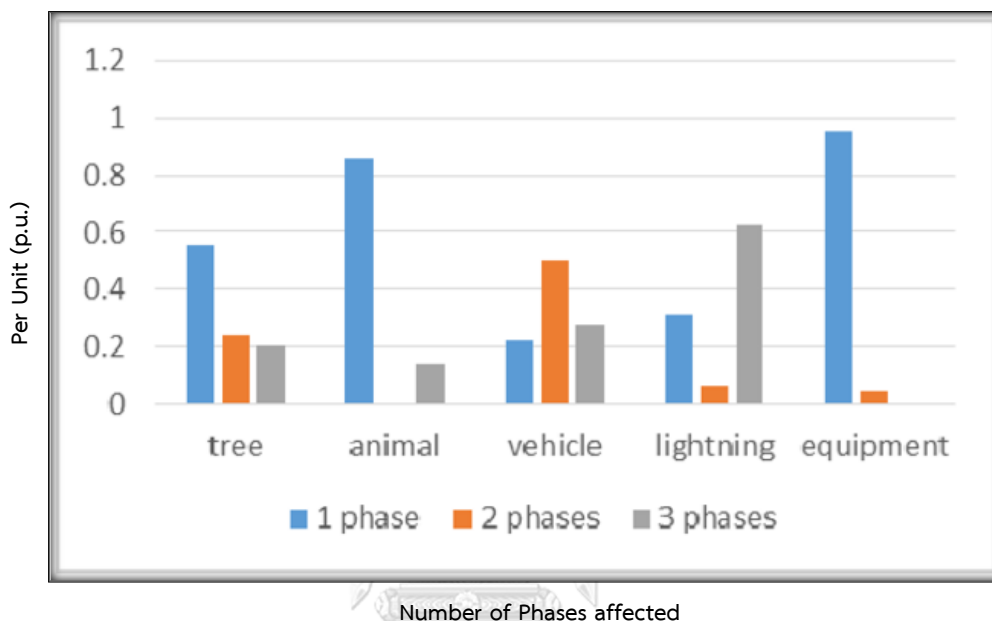
ความผิดปกติแบบชั่วคราวและความผิดปกติแบบถาวรเป็นคุณลักษณะเด่นหนึ่งที่สามารถบ่งชี้ถึงสาเหตุของความผิดปกติได้ ใน [13] ใช้คุณลักษณะเด่นดังกล่าวนี้จำแนกสาเหตุ รถครน ออกจากสาเหตุอื่นในระบบส่งจ่ายของประเทศมาเลเซีย โดยทุกสาเหตุที่เกิดจากรถครนจะเป็นความผิดปกติแบบถาวร เนื่องจากเมื่อเกิดความผิดปกติ Recloser จะสั่งการให้เซอร์กิตเบรกเกอร์ Trip Reclose ภายในระยะเวลาเพียงมิลลิวินาที ซึ่งเมื่อรถครนไปสัมผัสกับสายส่งแล้วก็ไม่สามารถเอาตัวครนออกจากสายส่งได้ภายในระยะเวลาเพียงเท่านั้น

## 4. ตำแหน่งมุมเริ่มต้นที่เกิดความผิดปกติ (Fault Insertion Angle)

ตำแหน่งที่เริ่มเกิดความผิดปกติสามารถระบุได้ด้วยค่ามุมที่เริ่มตกบนคลื่นแรงดันซึ่งสามารถเชื่อมโยงกับค่ากระแสที่เพิ่มขึ้นในเฟสเดียวกัน ใน [9, 14] นำคุณลักษณะเด่นค่านี้มาใช้จำแนกสาเหตุความผิดปกติที่เกิดจากสัตว์ ฟ้าผ่า และต้นไม้ แต่คุณลักษณะเด่นค่านี้สามารถนำมาใช้กับความผิดปกติแบบ 1 เฟสเท่านั้น หากมากกว่า 1 เฟส มุมที่นำมาพิจารณานั้นจะไม่แน่นอน

### 5. จำนวนเฟสที่ได้รับผลกระทบ (Number of Phases affected)

ชนิดของความผิดพลาดเป็นคุณลักษณะเด่นค่าหนึ่งที่สามารถบ่งชี้ถึงสาเหตุของความผิดพลาดที่เกิดจากสาเหตุใด เช่น ใน [15] สาเหตุอุปกรณ์ชำรุด และสัตว์ ส่วนใหญ่จะเกิดความผิดพลาดแบบ 1 เฟส ดังรูปที่ 2-8

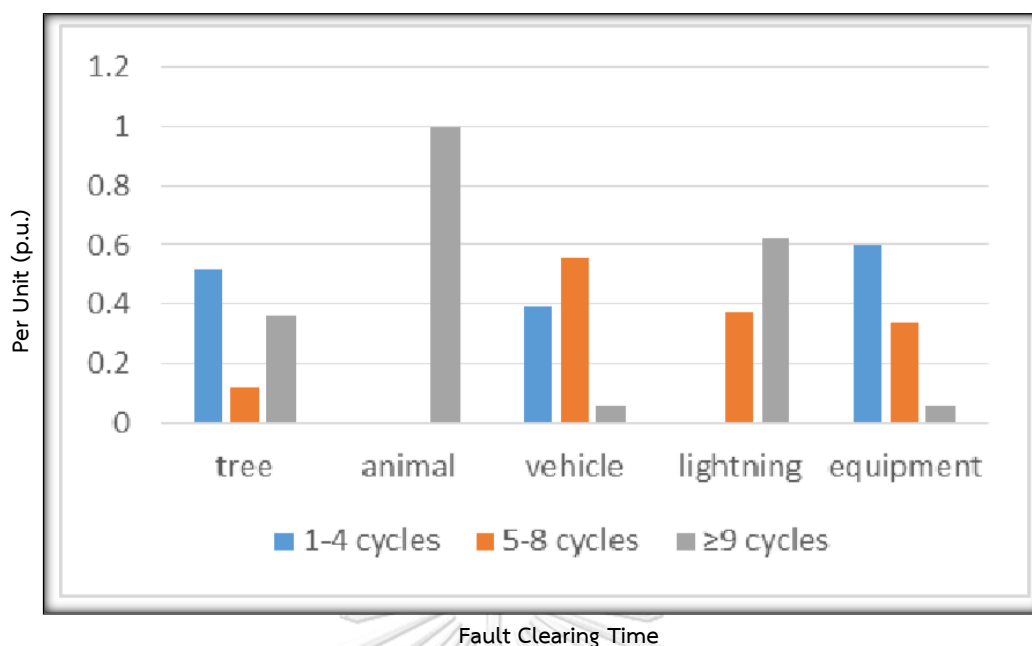


รูปที่ 2-8 ตัวอย่างจำนวนเฟสที่ได้รับผลกระทบ (Number of Phases affected) [15]

- จากงานวิจัยในอดีต จำนวนของเฟสที่ได้รับผลกระทบถูกนำเสนอในรูปแบบที่ต่างกันดังนี้
- ความผิดพลาดชนิดเฟสเดียวหรือหลายเฟส (Single Phase or Multi Phase) [14]
  - แบบ 1 เฟส, 2 เฟส หรือ 3 เฟส [15]
  - ชนิดของความผิดพลาด (Fault Type) [7]

### 6. ระยะเวลาขจัดความผิดพลาด (Fault Clearing Time)

ระยะเวลาขจัดความผิดพลาดอาจไม่ได้ขึ้นอยู่กับสาเหตุของความผิดพลาดเพียงอย่างเดียว อาจขึ้นอยู่กับค่าการตั้งค่า (Setting) อุปกรณ์ป้องกัน [9] ระยะเวลาที่เกิดความผิดพลาด แต่ทั้งนี้ก็มี การนำคุณลักษณะเด่นค่านี้มาเป็นตัวบ่งชี้ถึงสาเหตุของความผิดพลาด เช่นในงานวิจัย [15] พบว่า ระยะเวลาในการขจัดความผิดพลาดที่เท่ากับหรือมากกว่า 9 วัฏคลื่น (cycles) สามารถบ่งชี้สาเหตุ ความผิดพลาดที่เกิดจากสัตว์ได้ ดังรูปที่ 2-9



รูปที่ 2-9 ระยะเวลาขจัดความผิดปกติ (Fault Clearing Time) [15]

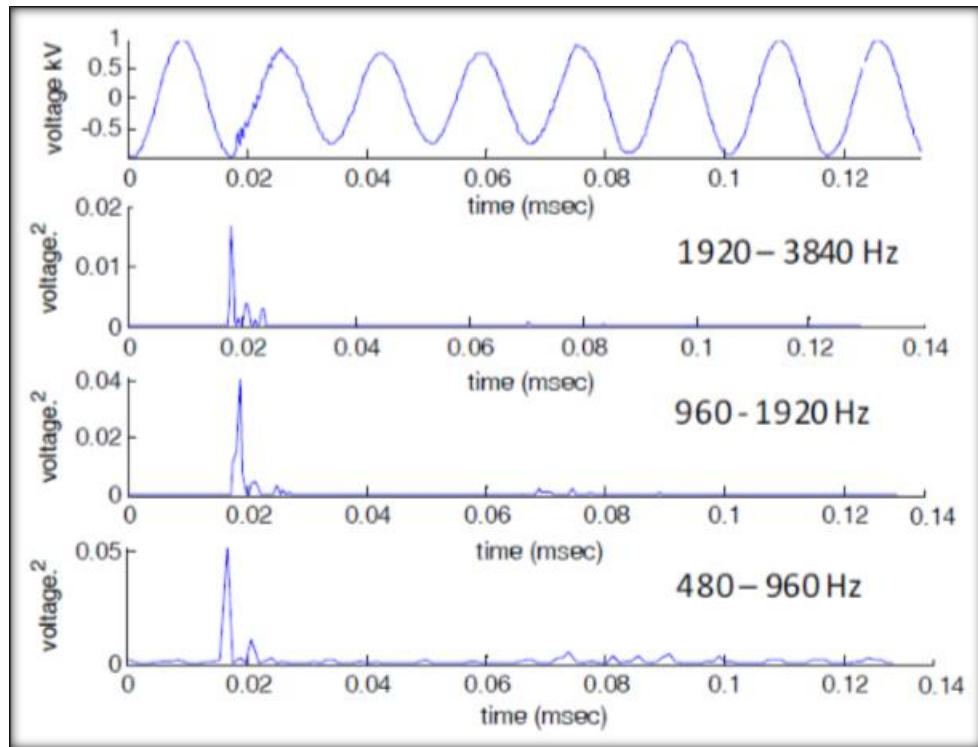
#### 7. วิเคราะห์คุณลักษณะของความถี่ (Frequency-Domain Characteristics)

การแปลงเวฟเลต (Wavelet Transform) เพื่อวิเคราะห์คุณลักษณะของความถี่ที่เปลี่ยนแปลงตามเวลาของรูปคลื่นแรงดัน โดยในการวิเคราะห์ต้องทำให้อยู่ในรูปมาตรฐานและแบ่งช่วงความถี่ออกเป็น 3 ช่วง [9, 14] ได้แก่

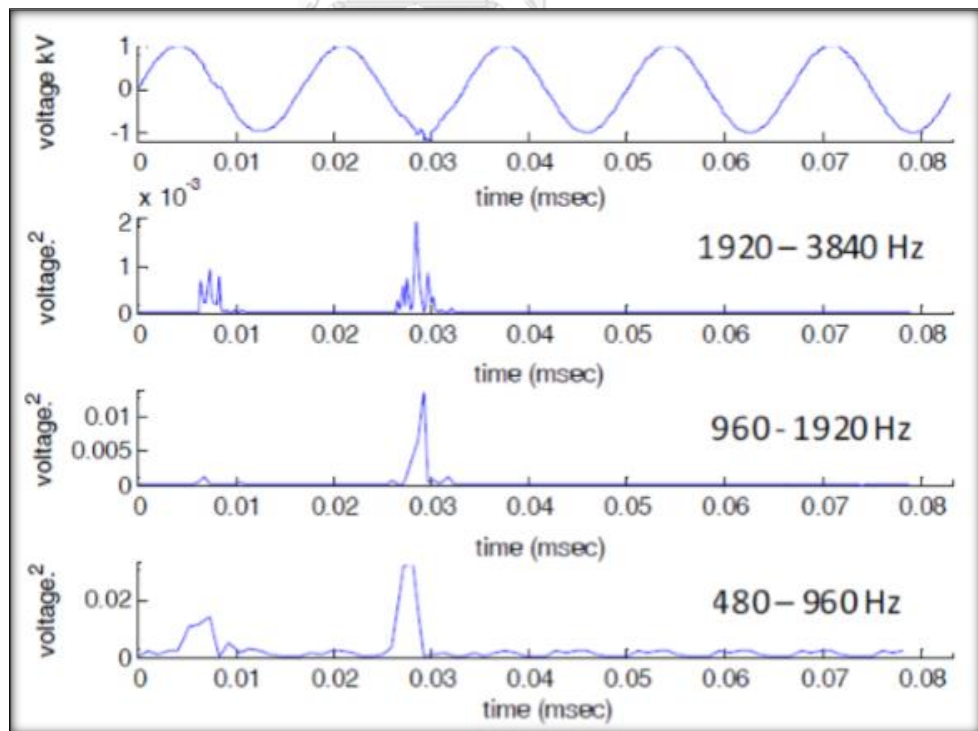
1. พิจารณาช่วงความถี่ 1920-3840 Hz.
2. พิจารณาช่วงความถี่ 960-1920 Hz.
3. พิจารณาช่วงความถี่ 480-960 Hz.

โดยใน [15] สามารถใช้คุณลักษณะเด่นดังกล่าวนี้จำแนกสาเหตุของความผิดปกติที่เกิดจากฟ้าผ่า ดังรูปที่ 2-10 และ ต้นไม้ ดังรูปที่ 2-11





รูปที่ 2-10 การแปลงเวฟเลตของรูปคลื่นแรงดันจากสาเหตุฟ้าผ่า [15]



รูปที่ 2-11 การแปลงเวฟเลตของรูปคลื่นแรงดันจากสาเหตุต้นไม้ [15]

## 8. ค่าแรงดันอาร์ค (Voltage Arc)

ค่าแรงดันอาร์ค (Voltage Arc) เป็นคุณลักษณะเด่นค่าหนึ่งที่สามารถบ่งชี้ถึงเอกลักษณ์ของสาเหตุของความผิดปกติพร้อมได้ ขณะที่เกิดความผิดปกติเมื่อวัตถุไปสัมผัสกับสายส่งจะทำให้เกิดการอาร์ค (Arc) และความต่างศักย์ระหว่างอากาศเรียกว่า (Voltage Arc) [9, 14] โดยสามารถคำนวณได้จากสมการที่ (2-4)

$$V_{arc} = V_f - \left[ R.I_f + L \left( \frac{dI_f}{dt} \right) \right] \quad (2-4)$$

- เมื่อ  $V_{arc}$  คือ ค่าแรงดันอาร์คสูงสุด ณ จุดเกิดความผิดปกติ  
 $V_f$  คือ แรงดันที่วัดจากหม้อแปลงแรงดัน (VT)  
 $R$  คือ ความต้านทานของสายส่ง  
 $L$  คือ ค่าความเหนี่ยวนำของสายส่ง  
 $I_f$  คือ ค่าของกระแสลัดวงจรที่วัดได้จาก DFR หรือ PQM

## 2.5 การเลือกคุณลักษณะเด่น (Feature Selection)

เมื่อมีการสกัดคุณลักษณะเด่น (Feature Extraction) จากข้อมูลที่มีแล้วนั้น อาจมีบางคุณลักษณะที่ไม่มีนัยสำคัญต่อสาเหตุของความผิดปกติที่สนใจซึ่งสามารถละทิ้งได้ และเป็นการลดการใช้งานทรัพยากรด้านการคำนวณและหน่วยความจำใช้งานโดยไม่จำเป็น วิธีการลดมิติ (Dimensionality Reduction) ของเวกเตอร์คุณลักษณะนี้เรียกว่า การเลือกคุณลักษณะ (Feature Selection)

วิธีการเลือกคุณลักษณะเด่นทำได้โดยสุ่มเลือกคุณลักษณะเด่นมาจำนวนหนึ่ง จากนั้นทำการทดสอบ (Validation) ประสิทธิภาพแล้วเลือกใช้ชุดคุณลักษณะเด่นที่ได้ประสิทธิภาพสูงสุด หรืออาจเลือกทั้งหมดพร้อมกันแล้วสรุปผลลัพธ์ด้วยวิธีเสียงข้างมาก [16] ซึ่งสามารถแบ่งเป็น 2 กลุ่มดังนี้

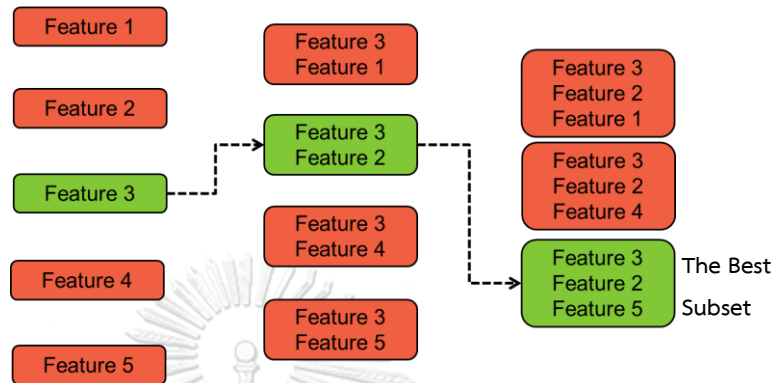
### 2.5.1 วิธีแรปเปอร์ (Wrapper Method)

เป็นการคัดเลือกค่าคุณลักษณะเด่น (Feature) ด้วยการสร้างโมเดล (classification model) ขึ้นมาจากเซตของคุณลักษณะเด่นที่กำหนดไว้และวัดประสิทธิภาพการทำงานของโมเดลโดยเลือกค่าที่มีความแม่นยำสูงสุด แต่จะใช้เวลานานเนื่องจากต้องคัดเลือกจากทุกคุณลักษณะเด่นโดยวิธีนี้จะประกอบไปด้วยวิธีย่อย 2 แบบ คือ วิธี Sequential Forward Selection และ Backward Elimination [17]

- Sequential Forward Selection เป็นการสร้างโมเดลโดยการเพิ่มทีละ 1 คุณลักษณะเด่น ถ้าคุณลักษณะเด่นที่ใส่เพิ่มให้ประสิทธิภาพที่ดีก็จะเก็บไว้และเลือกคุณลักษณะเด่น

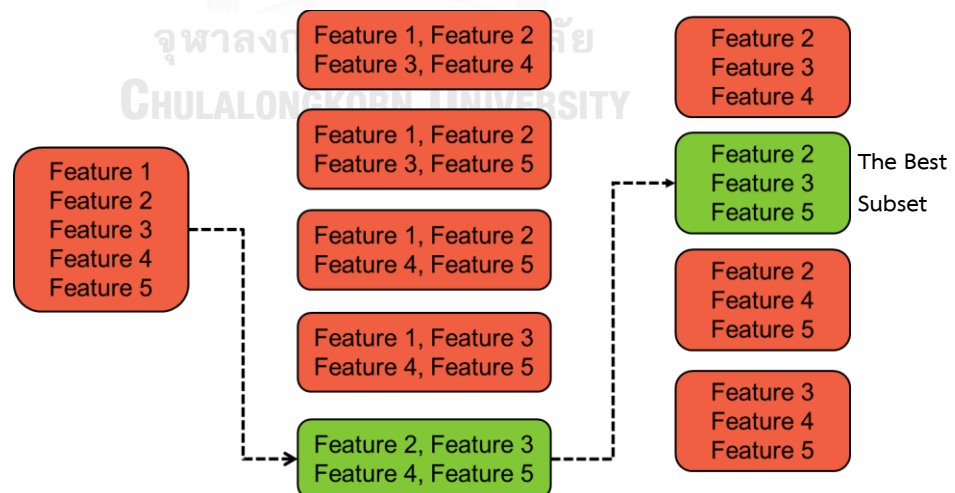


อื่นๆ มาเพิ่มต่อไปจนประสิทธิภาพของโมเดลไม่ได้ดีที่สุดก็จะหยุดทำงานดังแสดงในรูป 2-12 ใน [18] ประยุกต์ใช้วิธี Sequential Forward Selection ในการคัดเลือกคุณลักษณะเด่น



รูปที่ 2-12 การเลือกคุณลักษณะเด่นด้วยวิธี Sequential Forward Selection

- Backward Elimination เป็นการสร้างโมเดลที่เริ่มจากการใช้คุณลักษณะเด่นทั้งหมดก่อนและตัด (eliminate) คุณลักษณะเด่นที่ไม่สำคัญทิ้งไปทีละ 1 คุณลักษณะเด่น ถ้าประสิทธิภาพดีขึ้นก็ตัดคุณลักษณะเด่นอื่นๆ ต่อไป จนกว่าจะได้โมเดลที่ดีที่สุด ดังแสดงในรูปที่ 2-13



รูปที่ 2-13 การเลือกคุณลักษณะเด่นด้วยวิธี Backward Elimination

### 2.5.2 วิธีฟิลเตอร์ (Filter Method)

เป็นการคัดเลือกคุณลักษณะเด่นโดยใช้การคำนวณหาค่าน้ำหนักซึ่งอาจจะเป็นค่าความสัมพันธ์ระหว่างแต่ละคุณลักษณะเด่น กับ class ต่างๆ โดยมีเทคนิคในการคำนวณค่าน้ำหนักหลายวิธีต่างกัน เช่น Information Gain, Chi-Square [19]

- Information Gain (IG) คือ การประเมินค่าด้วยการแบ่งข้อมูลเป็นกลุ่มแล้วคำนวณค่า Gain โดยจะมีค่า Entropy เป็นค่าถ่วงน้ำหนัก จากนั้นมาหาว่าคุณลักษณะใดมีค่า Gain สูงสุด จะถูกเลือกเป็นกลุ่มที่มีอำนาจในการจำแนก ดังสมการที่ (2-5) และ (2-6)

$$Entropy(c_1) = -P(c_1) \log_2 P(c_1) \quad (2-5)$$

$$IG = Entropy(initial) - [P(c_1) \times Entropy(c_1) + P(c_2) \times Entropy(c_2) + \dots] \quad (2-6)$$

โดยที่  $P(c_1)$  = ความน่าจะเป็นของเหตุการณ์  $c_1$

- Chi-Square คือ การประเมินค่าของ Attribute โดยคำนวณทางสถิติ ดังสมการที่ (2-7)

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \quad (2-7)$$

โดยที่  $\chi^2$  คือ ค่าสถิติ Chi-Square

$O_i$  คือ ความถี่ที่ได้จากการสังเกต (Observed Frequency)

$E_i$  คือ ความถี่ที่คาดหวัง (Expected Frequency) เท่ากับ

จำนวนข้อมูลคุณสมบัติส่วนที่คาดหวัง หรือ  $E_i = np_i$

$k$  คือ จำนวนกลุ่มหรือจำนวนระดับ

$n$  คือ จำนวนตัวอย่าง

$p_i$  คือ สัดส่วนของระดับที่  $i$

## 2.6 เครื่องมือที่ใช้ในการจำแนก

### 2.6.1 โครงข่ายประสาทเทียม (Artificial Neural Network)

งานวิจัยในอดีต [4, 6, 7, 12, 20] นำ Artificial Neural Network มาใช้เป็นเครื่องมือในการจำแนกสาเหตุของความผิดปกติโดยใช้ข้อมูล 1. วงจรที่เกิดความผิดปกติ (Circuit ID), 2. สภาพอากาศ (Weather), 3. ฤดูกาล (Season), 4. วันของสัปดาห์ (Day of Week), 5. ช่วงเวลาของวัน

(Time of Day), 6. จำนวนเฟสที่ได้รับผลกระทบ (Number of Phases affected), 7. เกิดที่อุปกรณ์ป้องกันตัวใด (Protective Device) เป็นข้อมูลนำเข้า (Input) ให้วิธี Artificial Neural Network จำแนกความผิดปกติที่เกิดจาก สัตว์ ต้นไม้ และฟ้าผ่า

Artificial Neural Network [16] ใช้หลักการเบื้องต้นมาจากการจำลองการทำงานของเซลล์ประสาท (Neural) ของมนุษย์ โดยแต่ละเซลล์ประสาทจะมีหลักการสื่อสารกันผ่านการกระตุ้นด้วยศักย์ไฟฟ้า กล่าวคือ กระแสประสาทเริ่มต้นส่งสัญญาณจากตัวเซลล์ (Cell body) ผ่านแกนประสาทนำออก (Axon) ไปยัง จุดประสานประสาทหรือไซแนปส์ (Synapse) ซึ่งจะสื่อสารเชื่อมต่อกับเซลล์ประสาทอื่นผ่านใยประสาทนำเข้า (Dendrite) ณ จุดที่เซลล์ประสาททำการสื่อสารกันนี้เอง ที่เป็นหลักการสำคัญที่นำไปประยุกต์ใช้ใน Artificial Neural Network หรือเป็นการจำลองกระบวนการคิดของมนุษย์ โดยการนำข้อมูลนำเข้า เชื่อมโยงผลรวมของข้อมูลเข้าด้วยกัน เพื่อสร้างตัวแบบหรือโมเดลในการพยากรณ์ข้อมูลในอนาคต และในกระบวนการนำเข้าข้อมูลจะพยายามปรับปรุงและพัฒนาข้อมูลให้มีความเหมาะสมกับเงื่อนไขที่กำหนด เพื่อลดความผิดพลาดในการพยากรณ์ข้อมูลให้ต่ำที่สุด โดยสามารถแบ่งเป็น 2 ประเภทคือ 1. วิธีโครงข่ายประสาทเทียมแบบชั้นเดียว (Single – Layer Neural Networks) 2. วิธีโครงข่ายประสาทเทียมแบบหลายชั้น (Multi – Layer Neural Networks)

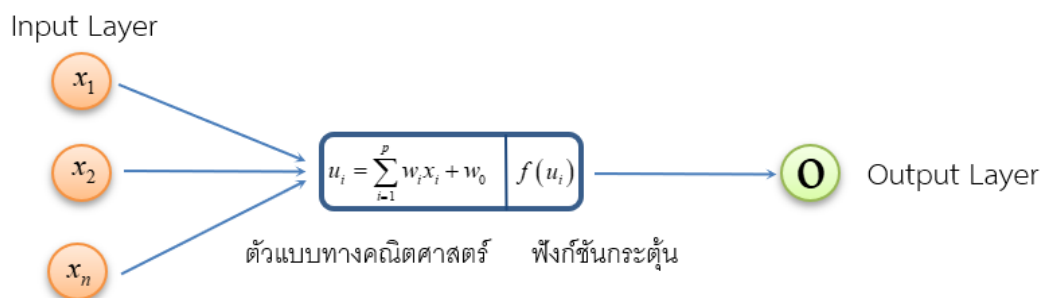
โครงสร้างของวิธี Artificial Neural Network ประกอบด้วย 3 ส่วน

1. สถาปัตยกรรมของโครงข่าย (Network Architecture) ประกอบด้วย 3 ชั้นหลัก
  - 1.1 ชั้นรับข้อมูล (Input Layer) คือ ข้อมูลหรือตัวแปรที่จะนำเข้าสู่ระบบที่ถูกเก็บในรูปแบบของเวกเตอร์ โดยจะต้องเป็นข้อมูลที่เป็นตัวเลขเท่านั้น Artificial Neural Network ถึงจะสามารถประมวลผลได้
  - 1.2 ชั้นแฝง (Hidden Layer) คือ ชั้นที่อยู่ตรงกลางระหว่างชั้นรับข้อมูลและชั้นแสดงผล เป็นชั้นที่ช่วยสร้างรูปแบบความสัมพันธ์และแปลงความสัมพันธ์ใหม่โดยใช้ฟังก์ชันกระตุ้น
  - 1.3 ชั้นแสดงผล (Output Layer) คือ ผลลัพธ์ที่เกิดขึ้นจริง (Actual Outputs) จากกระบวนการสอนของ Artificial Neural Network

2. การปรับค่าน้ำหนัก (Adjusting Weight) คือ สิ่งที่ได้จากการสอนของวิธี Artificial Neural Network หรือความรู้ (Knowledge)
3. ฟังก์ชันกระตุ้น (Activation Function) เป็นฟังก์ชันการแปลงแสดงสมการทางคณิตศาสตร์

#### 2.6.1.1 วิธีโครงข่ายประสาทเทียมแบบชั้นเดียว (Single-Layer Neural Networks)

Single-Layer Neural Networks ประกอบด้วยชั้นรับข้อมูล (Input Layer) และชั้นแสดงผล (Output Layer) เท่านั้น ซึ่งในชั้นรับข้อมูลจะส่งข้อมูลเข้าสู่กระบวนการหาค่าน้ำหนักตามเส้นเชื่อมโยงต่างๆ แล้วนำมาคำนวณด้วยฟังก์ชันทางคณิตศาสตร์ จากนั้นส่งค่ามาที่ฟังก์ชันกระตุ้นเพื่อทำการพิจารณาตามเงื่อนไขของฟังก์ชันว่าชั้นแสดงผลควรให้ผลลัพธ์ใด ซึ่ง Single-Layer Neural Networks เหมาะสำหรับการแบ่งข้อมูลด้วยเส้นตรงเท่านั้น ตัวอย่างวิธีโครงข่ายแบบชั้นเดียว ลักษณะโครงสร้างของโครงข่ายแบบชั้นเดียวแสดงดังรูปที่ 2-14



รูปที่ 2-14 โครงข่ายประสาทเทียมแบบชั้นเดียว (Single – Layer Neural Networks)

ซึ่งสามารถเขียนสมการได้ดังนี้

$$u_i = \sum_{i=1}^p w_i x_i + w_0 ; i = 1, 2, \dots, p \quad (2-8)$$

เมื่อ  $x_i$  คือ เวกเตอร์ของข้อมูลนำเข้า มีขนาด  $p$  มิติ

$w_i$  คือ ค่าน้ำหนักที่เชื่อมต่อระหว่างชั้นรับข้อมูลกับชั้นแสดงผล

$w_0$  คือ ค่าน้ำหนักคงที่ ในชั้นรับข้อมูลกับชั้นแสดงผล

โดยค่าเริ่มต้นของ  $w_0$  และ  $w_i$  เป็นค่าที่ได้จากการสุ่มข้อมูลในช่วง  $[-1, 1]$  ซึ่งค่าที่

โปรแกรมของการคำนวณได้กำหนดเอาไว้

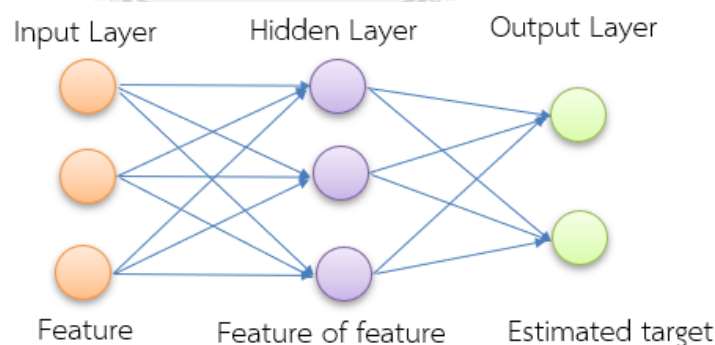
$w_0$  คือ เวกเตอร์ผลรวมของค่าถ่วงน้ำหนักในชั้นรับข้อมูลกับชั้นแสดงผล  
นำผลจากสมการ (2-8) มาทำการปรับค่าตามฟังก์ชันกระตุ้นที่ได้กำหนด คือ ฟังก์ชัน  
มอยด์ (Sigmoid Function) ดังแสดงในสมการ (2-9) ซึ่งให้ผลลัพธ์มีค่าอยู่ในช่วง  $[0,1]$

$$y = f(u_i) = \frac{1}{1 + \exp^{-u_i}} \quad (2-9)$$

เมื่อ  $y$  คือ เวกเตอร์ของผลลัพธ์ที่ได้จากกระบวนการของ Artificial Neural Network

#### 2.6.1.2 วิธีโครงข่ายประสาทเทียมแบบหลายชั้น (Multi-Layer Neural Networks)

Multi-Layer Neural Networks เป็นวิธีโครงข่ายที่พัฒนาจาก Single – Layer Neural Networks โดยการเพิ่มให้มีชั้นซ่อน (Hidden Layer) ตั้งแต่ 1 ชั้นขึ้นไป โดยชั้นซ่อนจะทำหน้าที่คล้ายการสกัดคุณลักษณะเด่น (Feature Extraction) ขึ้นใหม่จากชั้นนำเข้าไปป้อนชั้นซ่อนถัดไป หรือชั้นนำออก วิธีโครงข่ายแบบหลายชั้นจะใช้ในกรณีที่มีปัญหาซับซ้อนสูงซึ่งวิธีโครงข่ายแบบชั้นเดียวไม่สามารถใช้ในการแก้ไขปัญหาได้ ซึ่งปัญหาของข้อมูลโดยส่วนใหญ่ไม่สามารถทำการแบ่งประเภทได้ด้วยเส้นตรงที่ใช้ Single – Layer Neural Networks ดังนั้นวิธีการนี้จึงเป็นที่นิยมใช้ในการแก้ไขปัญหาในการทำงาน ลักษณะโครงสร้างของวิธีโครงข่ายของ Multi-Layer Neural Networks แสดงดังรูปที่ 2-15



รูปที่ 2-15 โครงข่ายประสาทเทียมแบบหลายชั้น (Multi – Layer Neural Networks)

#### 2.6.2 การถดถอยโลจิสติกส์ (Logistic Regression)

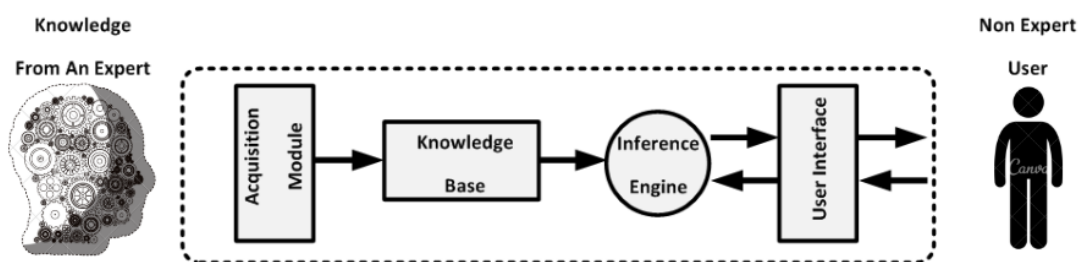
การถดถอยโลจิสติกส์ (Logistic Regression) เป็นอัลกอริทึมที่มีเอาไว้เพื่อศึกษาความน่าจะเป็นที่มีผลลัพธ์ได้ 2 คำตอบ หรือเป็น Binary เช่น 0/1, Yes/No, เสี่ยง/ไม่เสี่ยง ถือว่าอยู่กึ่งกลางระหว่างจะเป็น Classification หรือ Regression งานวิจัยในอดีต [20, 21] นำ Logistic Regression

มาจำแนกสาเหตุความผิดปกติที่เกิดจากต้นไม้ออกจากสาเหตุอื่น โดยใช้ข้อมูล 1. วงจรที่เกิดความผิดปกติ (Circuit ID), 2. สภาพอากาศ (Weather), 3. ฤดูกาล (Season), 4. วันของสัปดาห์ (Day of Week), 5. ช่วงเวลาของวัน (Time of Day), 6. จำนวนเฟสที่ได้รับผลกระทบ (Number of Phases affected), 7. เกิดที่อุปกรณ์ป้องกันตัวใด (Protective Device) เป็นข้อมูลนำเข้า (Input) เพื่อทำการจำแนกสาเหตุต้นไม้ออกจากสาเหตุอื่นๆ

### 2.6.3 ระบบผู้เชี่ยวชาญ (Expert System)

ระบบผู้เชี่ยวชาญ คือระบบคอมพิวเตอร์ที่รวบรวมและจำลองความรู้ของมนุษย์ วิธีการแก้ปัญหา ที่มีความยุ่งยาก ซับซ้อน จำเป็นต้องใช้ความรู้ความเชี่ยวชาญพิเศษ ประสบการณ์ ความรู้ของมนุษย์ในการแก้ปัญหา เพื่อให้ปัญหานั้นได้รับการแก้ไขอย่างมีประสิทธิภาพ [22]

โครงสร้างพื้นฐานของระบบผู้เชี่ยวชาญประกอบด้วย 4 ส่วนหลัก คือ 1. ส่วนจัดหาความรู้ (Acquisition Module) 2. ฐานความรู้ (Knowledge Base) 3. ส่วนควบคุมการอนุมาน (Inference Engine) 4. ส่วนติดต่อกับผู้ใช้งาน (User Interface Module) ดังรูปที่ 2-16



รูปที่ 2-16 โครงสร้างของระบบผู้เชี่ยวชาญ

ผู้ใช้จะติดต่อกับระบบผู้เชี่ยวชาญที่ส่วนติดต่อกับผู้ใช้งาน (User Interface Module) โดยจะป้อนข้อมูล และสามารถได้รับคำปรึกษาหรือคำตอบผ่านส่วนนี้ และยังสามารถป้อนผลคำตอบเพื่อเป็นฐานความรู้ต่อไปได้

ในบทความ [13] ประยุกต์ระบบผู้เชี่ยวชาญโดยใช้อัตราการเปลี่ยนแปลงของกระแสและแรงดันเทียบกับเวลา เป็นข้อมูลนำเข้าและจำแนกสาเหตุความผิดปกติที่เกิดจาก ฟิวส์ มลภาวะ รถเครน ต้นไม้ และหม้อแปลงกระแสระเบิด

#### 2.6.4 การตัดสินใจต้นไม้ (Decision Tree)

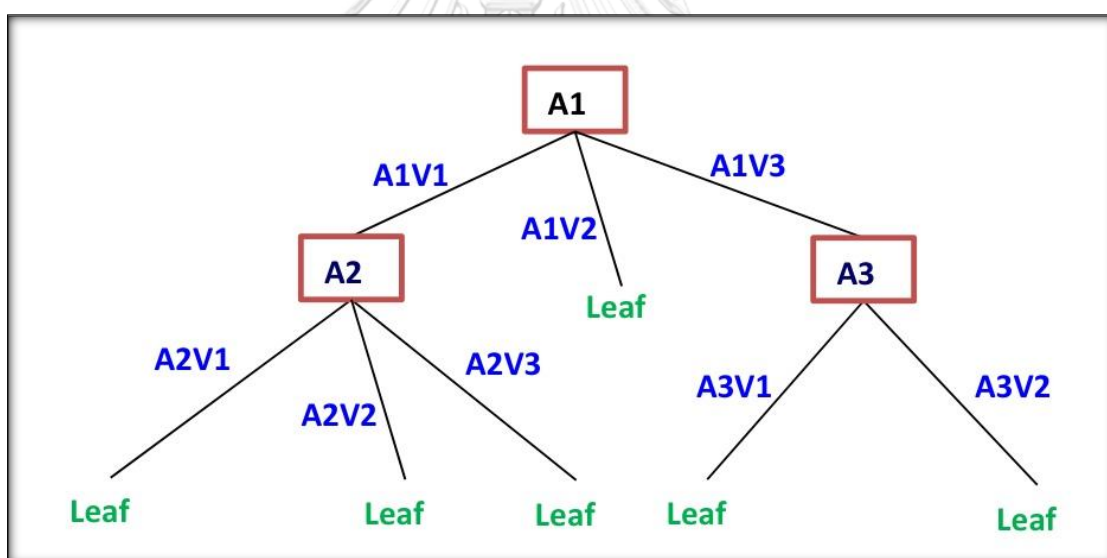
ต้นไม้ตัดสินใจ (Decision Tree) [16] เป็นเครื่องมือที่ใช้ในการช่วยตัดสินใจ โดยการใช้โครงสร้างกราฟที่มีลักษณะคล้ายต้นไม้ดังรูปที่ 2-17 ทำให้มนุษย์สามารถแปลผลได้ง่ายขึ้น เพื่อใช้ประโยชน์ในการจำแนกประเภท (Classification) ทั้งในปัญหาแบบไบนารี (Binary Classes) และปัญหาแบบหลายคลาส (Multiple Classes) รวมทั้งยังสามารถนำไปใช้แสดงโครงสร้างข้อมูลในฐานข้อมูล (Database) เพื่อความรวดเร็วในการค้นหา [16]

Decision Tree ประกอบด้วย 3 องค์ประกอบหลักได้แก่

**ปม (Node)** ใช้สำหรับระบุชื่อของลักษณะประจำ (Attribute)

**กิ่ง (Branch)** ใช้สำหรับระบุค่าของแต่ละ Attribute ของปมที่เหนือกิ่งกึ่งนั้น บางครั้งอาจเรียกว่า เส้นเชื่อม (Edge) หรือการโยง (Link)

**ใบ (Leaf)** ใช้สำหรับระบุค่าเป้าหมาย (Target) หรือคลาส (Class) ของปัญหา



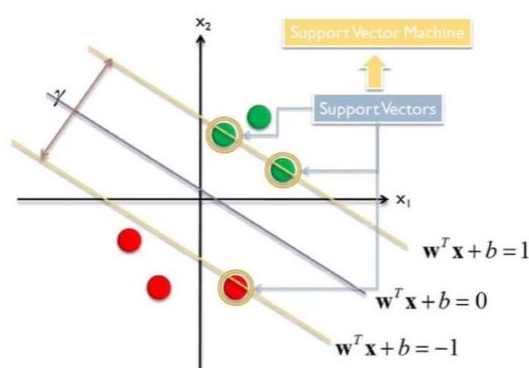
รูปที่ 2-17 ตัวอย่างต้นไม้ตัดสินใจ [16]

ใน [10] ใช้ต้นไม้ตัดสินใจ (Decision Tree) ประยุกต์ใช้ในการจำแนกสาเหตุความผิดปกติที่เกิดจาก ต้นไม้ ฟาร์ม และสัตว์ ในระบบส่งของการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย

### 2.6.5 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

Support Vector Machine (SVM) [16] เป็นเครื่องมือที่ใช้ในการจำแนกประเภทที่นิยมมาก มีประสิทธิภาพดี ทั้งนี้เนื่องจาก SVM มีการจดจำข้อมูลที่อยู่บริเวณขอบของแต่ละคลาส หรือที่เรียกว่า ซัพพอร์ตเวกเตอร์ (Support Vector) อีกทั้งยังรองรับปริภูมิที่มีข้อมูลเป็นแบบเชิงเส้นและแบบไม่เชิงเส้นก็ได้ โดยการอาศัยฟังก์ชันเคอร์เนล (Kernel) ในการลดความซับซ้อนของข้อมูลโดยการส่ง (Mapping) ไปยังปริภูมิพิเศษที่มีความเป็นเชิงเส้น ทั้งนี้ปริภูมิพิเศษดังกล่าวอาจสร้างได้จริง หรือเป็นเพียงนามธรรมก็ได้ด้วยการอาศัยกลวิธีเชิงเคอร์เนล (Kernel Trick) ภายในปริภูมิผลคูณภายใน

ใน SVM ระยะขอบ (Margin) ที่มากที่สุดระหว่างคลาสหนึ่งกับอีกคลาสหนึ่ง ถูกใช้เป็นบรรทัดฐานในการตัดสินใจเลือกระนาบเกินที่เหมาะสมที่สุด (Optimal Hyperplane) และเวกเตอร์ข้อมูลที่อยู่บริเวณขอบของแต่ละคลาสจะถูกรเรียกว่า ซัพพอร์ตเวกเตอร์ (Support Vector) ดังรูปที่ 2-18



รูปที่ 2-18 ซัพพอร์ตเวกเตอร์ (Support Vector) [16]

โดย  $\gamma$  คือ ระยะขอบ (Margin) และข้อมูลบริเวณขอบของแต่ละคลาสที่อยู่ภายในเครื่องหมายวงกลมทั้งหมด คือ ซัพพอร์ตเวกเตอร์ และเป็นที่มาของชื่อ ซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์เหล่านี้จะถูกบันทึกไว้ในตัวจำแนกประเภท (Classifier) ซึ่งหากจำนวนซัพพอร์ตเวกเตอร์ยิ่งมาก ขนาดหน่วยความจำที่ใช้เก็บตัวจำแนกประเภทก็ยิ่งมากตามไปด้วย ทั้งนี้ หากจำนวนเวกเตอร์ซัพพอร์ตมีสัดส่วนมากเกินไปเมื่อเทียบจำนวนข้อมูลที่น่าไปใช้ฝึกฝนทั้งหมด จะบอกเราได้ถึงปัญหา Overfitting ที่เกิดขึ้น



เมื่อทราบเกณฑ์ที่ใช้ในการเลือกระนาบเกินที่เหมาะสมซึ่งเกี่ยวข้องกับระยะขอบ ดังนั้นความสัมพันธ์ระหว่างระยะขอบกับระนาบเกินต้องถูกนำมาพิจารณา โดยเริ่มต้นจากซัพพอร์ตเวกเตอร์ของแต่ละคลาสอย่างละเวกเตอร์  $(\mathbf{x}_+, \mathbf{x}_-)$  ซึ่งสอดคล้องกับสมการระนาบเกินในรูปที่ 2-18 ดังนี้

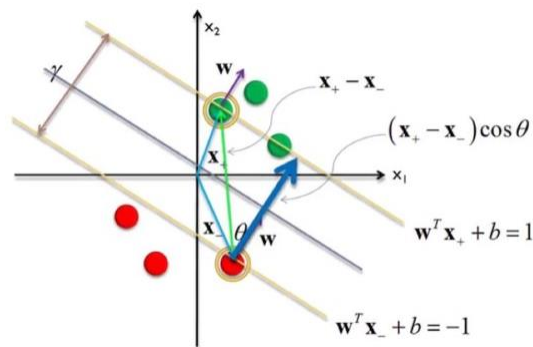
$$\mathbf{w}^T \mathbf{x}_+ + b = 1 \quad (2-10)$$

$$\mathbf{w}^T \mathbf{x}_- + b = -1 \quad (2-11)$$

นำสมการที่ (2-10) ลบออกจากสมการที่ (2-11) จะได้

$$\mathbf{w}^T (\mathbf{x}_+ - \mathbf{x}_-) = 2 \quad (2-12)$$

โดยจะปรากฏเทอม  $(\mathbf{x}_+, \mathbf{x}_-)$  ซึ่งเป็นเวกเตอร์ลัพธ์ที่เริ่มจากระนาบเกินคลาสหนึ่งไปอีกคลาสหนึ่ง และกระทำมุม  $\theta$  กับเวกเตอร์ปกติ  $\mathbf{w}$  ดังรูปที่ 2-19



รูปที่ 2-19 ความสัมพันธ์ระหว่างระยะขอบกับเวกเตอร์ปกติของระนาบเกิน [16]

จากรูปที่ 4-19 เวกเตอร์ที่มีทิศทางเดียวกับเวกเตอร์ปกติ  $\mathbf{w}$  และมีขนาดเท่ากับระยะขอบ  $\gamma$  สามารถหาได้จากสมการที่ (2-13)

$$(\mathbf{x}_+ - \mathbf{x}_-) \cos \theta = (\mathbf{x}_+ - \mathbf{x}_-) \left( \frac{\mathbf{w}^T (\mathbf{x}_+ - \mathbf{x}_-)}{\|\mathbf{w}\| \|\mathbf{x}_+ - \mathbf{x}_-\|} \right) \quad (2-13)$$

แทนค่าจากสมการที่ (2-12) ลงในสมการที่ (2-13) จะได้

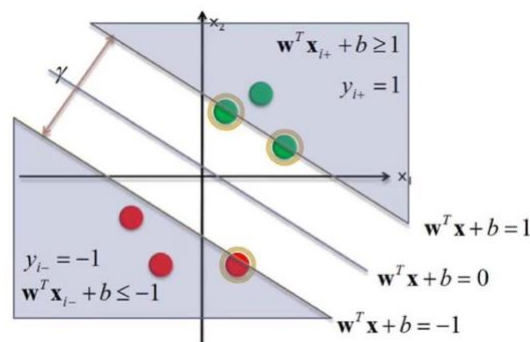
$$(\mathbf{x}_+ - \mathbf{x}_-) \frac{\mathbf{w}^T (\mathbf{x}_+ - \mathbf{x}_-)}{\|\mathbf{w}\| \|\mathbf{x}_+ - \mathbf{x}_-\|} = \frac{2}{\|\mathbf{w}\|} \left( \frac{\mathbf{w}^T (\mathbf{x}_+ - \mathbf{x}_-)}{\|\mathbf{w}\| \|\mathbf{x}_+ - \mathbf{x}_-\|} \right) \quad (2-14)$$

โดยเทอม  $\frac{(\mathbf{x}_+ - \mathbf{x}_-)}{\|\mathbf{x}_+ - \mathbf{x}_-\|}$  คือ เวกเตอร์หนึ่งหน่วย (Unit Vector) ที่มีทิศทางเดียวกับเวกเตอร์ปกติ  $\mathbf{w}$

ดังนั้นระยะขอบ  $\gamma$  จึงมีค่าเท่ากับขนาดของเวกเตอร์ในสมการที่ (2-14) กล่าวคือ

$$\gamma = \frac{2}{\|\mathbf{w}\|} \quad (2-15)$$

เมื่อได้ความสัมพันธ์ระหว่างระยะขอบกับขนาดของเวกเตอร์ปกติของระนาบเกินแล้ว จึงสามารถสร้างปัญหาการหาค่าเหมาะที่สุดจากวัตถุประสงค์ที่ต้องการให้ระยะขอบมีความกว้างมากที่สุดในการจำแนกประเภทแบบ 2 คลาส (Binary Classification) อันได้แก่ คลาสที่มีค่าเป็นบวก (+) และคลาสที่มีค่าเป็นลบ (-) โดยการพิจารณาเงื่อนไขเป็นพื้นที่ในบริเวณของแต่ละคลาสดังรูปที่ 2-20



รูปที่ 2-20 พื้นที่แสดงอาณาเขตแต่ละคลาสในกรณีไบนารี [16]

โดยพื้นที่ในบริเวณทั้งสองสามารถนำมารวมเป็นอสมการเดียวกันได้โดยการนำค่าคำตอบ  $y$  ของข้อมูลแต่ละข้อมูลไปคูณกับอสมการทั้งสองบริเวณดังสมการที่ (2-16)

$$y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad \forall i \quad (2-16)$$

จากความสัมพันธ์ของระยะขอบกับขนาดของเวกเตอร์ในสมการที่ (2-15) ร่วมกับเงื่อนไขในสมการที่ (2-16) ดังนั้น ปัญหาเพื่อการหาค่าเหมาะที่สุดจะเป็นดังนี้

$$\text{Maximize: } \gamma = \frac{2}{\|\mathbf{w}\|} \quad (\text{ให้ระยะขอบมีค่ากว้างที่สุด})$$

$$\text{Subject to: } y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad \forall i \quad (\text{ให้ข้อมูลอยู่ในบริเวณของคลาสที่ถูกต้อง})$$

ใน [23] ประยุกต์ใช้ SVM ในการจำแนกสาเหตุความผิดปกติพ่วงจากสาเหตุ ฟิวส์, ไฟไหม้ป่า, มลค้ำควา, มลภาวะ, วัตถุแปลกปลอม, รถชนเสา

### 2.6.6 วิธีการค้นหาเพื่อนบ้านใกล้ที่สุด K ตัว (K-Nearest-Neighbor Classifiers)

ขั้นตอนวิธี K-Nearest Neighbors (KNN) [16] เป็นขั้นตอนวิธีหนึ่งที่ไม่ซับซ้อน และมีประสิทธิภาพมากอีกวิธีหนึ่งซึ่งนิยมใช้เป็นอันดับแรกๆ ในการแก้ปัญหาการจำแนกประเภท (Classification) หลักการสำคัญของขั้นตอนวิธีนี้คือ การค้นหาตัวอย่างจำนวนหนึ่ง (จำนวน K ตัวอย่าง) โดยขั้นตอนจากปริภูมิตัวอย่างที่ปรากฏใกล้เคียงที่สุดกับสิ่งที่ต้องการจำแนกประเภทดังนี้

กำหนดให้ชุดข้อมูลหนึ่งประกอบด้วยตัวอย่างที่ทราบประเภทจำนวน 9 ตัวอย่าง (1-9) และมีตัวอย่างที่ไม่ทราบประเภท (X) จำนวน 1 ตัวอย่าง มีการกระจายตัวในปริภูมิตัวอย่างดังรูปที่ 2-21



โดยตัวอย่างที่ 1-3 เป็นประเภท B, ตัวอย่างที่ 4-6 เป็นประเภท G, และตัวอย่างที่ 7-9 เป็นประเภท R เมื่อต้องการประเภทของ X ด้วยขั้นตอนวิธี KNN ทำได้ดังนี้

1. กำหนดค่า K ซึ่งหมายถึงจำนวนตัวอย่างที่ใกล้เคียงที่สุด K ตัว โดยมีค่าตั้งแต่ 1 ถึงจำนวนตัวอย่างทั้งหมดในปริภูมิตัวอย่าง (ในตัวอย่างนี้คือ 9)

2. กำหนดวิธีการวัดความใกล้เคียงหรือความเหมือน (Similarity) ซึ่งโดยทั่วไปจะวัดจากค่าผกผันของระยะทาง (Distance) ระหว่างตัวอย่าง

ค่าระยะทางที่นิยมใช้มากแบบหนึ่ง คือ ระยะทางแบบยูคลิด (Euclidean Distance) ซึ่งระยะทางแบบยูคลิดจากจุด  $p(p_1, p_2, \dots, p_N)$  ไปจุด  $q(q_1, q_2, \dots, q_N)$  ในปริภูมิ N มิติสามารถหาได้จาก

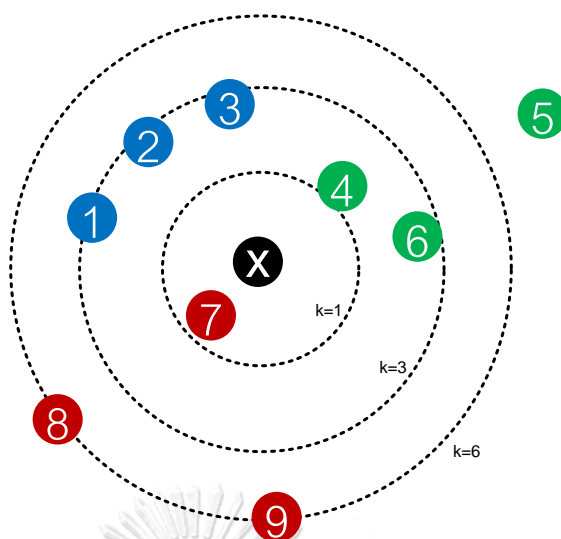
$$d(p,q) = \sqrt{\sum_{i=1}^N (p_i - q_i)^2} \quad (2-17)$$

3. นับจำนวนตัวอย่างที่ใกล้เคียงที่สุด K ตัว โดยหากจำนวนตัวอย่างประเภทใดมีมากที่สุด ใน k ตัว ให้กำหนดประเภทของตัวอย่างที่ต้องการทราบประเภทเป็นประเภทที่มีมากที่สุดนั้น ซึ่งคำตอบจะแตกต่างกันไปตามค่า K ที่เลือกดังตารางที่ 2-1

ตารางที่ 2-1 ผลการจำแนกประเภทด้วยค่า K ต่างๆ [16]

k	ผลการจำแนกประเภท
1	7-9
2	7-9, 4-6
3	4-6
4	4-6
5	4-6, 1-3
6	1-3
7	1-3
8	1-3, 7-9
9	1-3, 4-6, 7-9

จากผลในตารางที่ 2-1 พบว่าค่า K บางค่ามีผลทำให้การจำแนกประเภทกำกวม เนื่องจากมีจำนวนตัวอย่างที่ใกล้เคียงที่สุดเท่ากันหลายประเภท หากพิจารณาค่า K บางค่าดังในรูปที่ 2-22 จะพบว่าค่า K =6 อาจไม่เหมาะสมสำหรับปัญหานี้



รูปที่ 2-22 ระยะทางจากตัวอย่างที่ค่า K เท่ากับ 1, 3 และ 6 [16]

ในความเป็นจริง ปริภูมิตัวอย่างมีจำนวนและมีมิติมากจนไม่สามารถมองเห็นได้ครบแบบในตัวอย่างนี้ การเลือกค่า K ที่เหมาะสม จึงเป็นปัญหาสำคัญโดยตรงที่ส่งผลถึงความถูกต้องในการจำแนกประเภท

ในกรณีที่ไม่มีข้อมูลประกอบอื่นใด ค่า K ที่นิยมใช้มากที่สุดคือ 1 ซึ่งบางครั้งเรียกขั้นตอนวิธีนี้ว่า The Nearest Neighbor

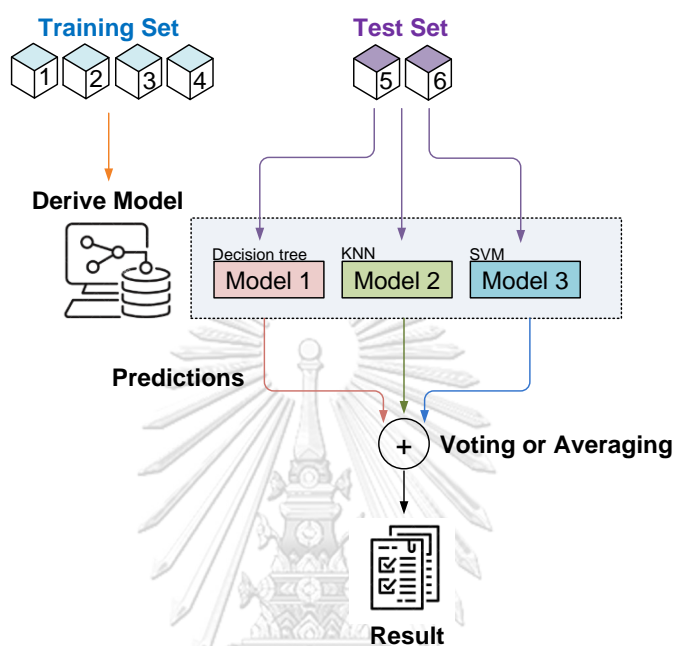
จากขั้นตอนวิธี K-Nearest Neighbors จะพบปัญหาในการใช้เวลาในการประมวลผลมากในกรณีที่ชุดข้อมูลมีจำนวนตัวอย่างมาก เพราะต้องคำนวณระยะห่างและความคล้าย เพื่อเปรียบเทียบกับตัวอย่างที่มีในชุดข้อมูลนั้น วิธีการแก้ไขปัญหานี้ทำได้หลายวิธี เช่น การลดจำนวนตัวอย่างลงด้วยขั้นตอนวิธี Condensed Nearest Neighbor (CNN) หรือ การลดจำนวนครั้งในการเปรียบเทียบลงด้วยขั้นตอนวิธี K-Dimension Tree

### 2.6.7 การเรียนรู้แบบรวมกลุ่ม (Ensemble Learning)

การเรียนรู้แบบรวมกลุ่ม [24] คือ การรวมตัวจำแนกประเภทข้อมูลหลายๆ วิธีที่มีความแตกต่างกันและอิสระต่อกันเข้าด้วยกัน และนำมาร่วมกันพิจารณาว่าผลลัพธ์อันไหนดีที่สุดเพื่อเพิ่มประสิทธิภาพของโมเดล เป้าหมายของการเรียนรู้แบบรวมกลุ่มคือการทำให้โมเดลมีความเหมาะสมกับชุดข้อมูลลักษณะต่างๆ (Generalizability) และมีความทนทาน (Robustness) โดยเทคนิคของการเรียนรู้แบบรวมกลุ่มมีอยู่หลากหลายวิธี ซึ่งวิธีที่ได้รับความนิยมได้แก่ วิธี Voting และ Bagging

## 1. Voting Method

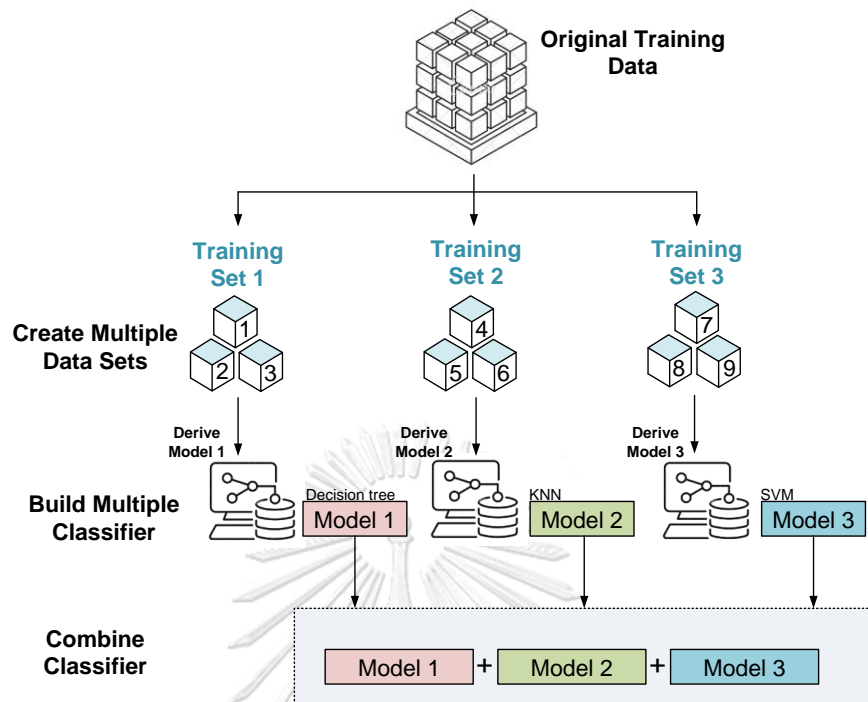
Voting Method [25] คือ การสร้างโมเดลจากหลากหลายวิธี เช่น Decision Tree, KNN, SVM จากนั้นรวบรวมคำตอบ (ผลทำนาย) ของแต่ละโมเดลและโหวตดูคำตอบที่เหมือนกันมากที่สุดหรือโหวตจากความน่าจะเป็นของค่าเฉลี่ย ดังแสดงในรูปที่ 2-23



รูปที่ 2-23 การเรียนรู้แบบรวมกลุ่มด้วยเทคนิค Voting Method

## 2. Bagging Method

Bagging Method [26] คือ การใช้ตัวจำแนกประเภทข้อมูลหลายวิธี เช่น Decision Tree, SVM, KNN มาเรียนรู้จากชุดข้อมูลที่ต่างกัน โดยแต่ละชุดข้อมูลจะถูกสุ่มมาจากข้อมูลตั้งต้นเดียวกัน และนำผลลัพธ์ที่ได้ในแต่ละโมเดลมารวมกันและทำนายผลลัพธ์สุดท้ายด้วยวิธีโหวตเสียงข้างมากดังแสดงในรูปที่ 2-24 เนื่องจากข้อมูลที่เรียนรู้มีความแตกต่างกันทำให้เกิดความหลากหลายในแบบจำลองจึงทำให้วิธีนี้จะช่วยลดความแปรปรวนและลดการเกิดปัญหาความเหมาะสมกับข้อมูลฝึกฝนมากเกินไป (overfitting)



รูปที่ 2-24 การเรียนรู้แบบรวมกลุ่มด้วยเทคนิค Bagging Method

## 2.7 การประเมินความถูกต้องของโมเดล

ในการสร้างและใช้งานตัวจำแนกข้อมูลเราจะต้องประเมินความถูกต้องและความผิดพลาดของตัวจำแนกข้อมูล ที่จะทำให้ทราบถึงประสิทธิภาพของตัวจำแนกข้อมูลที่เราสร้างขึ้น โดยสามารถประยุกต์วิธีต่างๆ ดังต่อไปนี้

### 2.7.1 ตัวชี้วัดความผิดพลาดจากการจำแนกข้อมูล [17]

หลังจากสร้างตัวจำแนกข้อมูล เราต้องตรวจสอบความถูกต้องในการจำแนกข้อมูลเพื่อประเมินว่าค่าความถูกต้องได้รับจากตัวจำแนกข้อมูลที่สร้างขึ้นสามารถยอมรับได้หรือไม่ โดยในการตรวจสอบความถูกต้องในการจำแนกข้อมูล เราใช้ข้อมูลทดสอบอีกชุดที่อยู่นอกเหนือจากชุดข้อมูลฝึกฝน เมื่อทดสอบแล้วจะทราบถึงจำนวนหรือเปอร์เซ็นต์ความถูกต้อง และทราบอัตราความผิดพลาดหรืออัตราการจำแนกข้อมูลผิด (Misclassification Rate) ของตัวจำแนก  $M$  ใดๆ ที่สร้างขึ้น ซึ่งสามารถคำนวณได้จาก  $1 - acc(M)$  โดยที่  $acc(M)$  คือ ค่าความถูกต้องในการจำแนกข้อมูลของตัวจำแนกข้อมูล  $M$

ในการวิเคราะห์หาความถูกต้องและความผิดพลาดในการจำแนกข้อมูล เราสามารถประยุกต์ใช้ Confusion Matrix ดังตารางที่ 2-2 ซึ่งเป็นตารางที่มีขนาด  $m \times n$  โดยข้อมูลแต่ละเซลล์

$cm_{i,j}$  หมายถึง จำนวนข้อมูลที่มีหมวดหมู่  $i$  แต่ถูกตัวจำแนกข้อมูลจำแนกหมวดหมู่เป็นหมวดหมู่  $j$  โดยที่ถ้าตัวจำแนกข้อมูลที่เราทดสอบเป็นตัวจำแนกข้อมูลที่มีความถูกต้องสูง จะทำให้เซลล์  $cm_{1,1}$ ,  $cm_{2,2}$  ไปจนถึง เซลล์  $cm_{m,m}$  มีจำนวนข้อมูลเป็นจำนวนมากและเซลล์  $cm_{i,j}$  อื่นๆ มีค่าเท่ากับ 0 หรือมีค่าใกล้ๆ 0

ตารางที่ 2-2 ตัวอย่าง Confusion Matrix [17]

		Predicted class	
		$C_1$	$C_2$
Actual class	$C_1$	True positives	False negatives
	$C_2$	false positives	true negatives

การประยุกต์ใช้เปอร์เซ็นต์ความถูกต้องในการจำแนกข้อมูลได้รับการประยุกต์ใช้อย่างแพร่หลาย อย่างไรก็ตาม การใช้เพียงเปอร์เซ็นต์ความถูกต้องเพียงอย่างเดียวอาจประเมินความถูกต้องของตัวจำแนกข้อมูลได้ไม่ دقيق ตัวอย่างเช่น เมื่อเราพิจารณาข้อมูลฝึกฝนที่เกี่ยวข้องกับหมวดหมู่ของผู้ป่วยที่เป็นมะเร็งและไม่เป็นมะเร็ง เมื่อทดสอบปรากฏว่าเปอร์เซ็นต์ความถูกต้องมีค่าเท่ากับ 90% ซึ่งมีความสูง อย่างไรก็ตาม เมื่อพิจารณาลึกลงไปอีกเราพบว่ามีเพียง 3-4% ของข้อมูลทั้งหมดที่พิจารณาเท่านั้นที่อยู่ในหมวดหมู่ของผู้ป่วยที่เป็นมะเร็ง ดังนั้นเราจะเห็นว่าค่าเปอร์เซ็นต์ความถูกต้อง 90 เปอร์เซ็นต์ ไม่อาจถูกใช้เพื่อวัดความถูกต้องของตัวจำแนกข้อมูลได้อย่างแม่นยำ โดยค่าความถูกต้อง 90 เปอร์เซ็นต์อาจเกิดจากการที่ตัวจำแนกข้อมูลสามารถจำแนกหรือทำนายหมวดหมู่ของข้อมูลของผู้ป่วยที่ไม่เป็นโรคมะเร็งได้อย่างถูกต้อง แต่นั่นจะไม่ใช้สิ่งสำคัญเท่าไรนัก เนื่องจากวัตถุประสงค์หลักของเราคือต้องการระบุ/จำแนกผู้ป่วยที่เป็นโรคมะเร็งด้วยเหตุนี้เราอาจพิจารณาค่า “Sensitivity” และ “Specificity” ซึ่งค่า Sensitivity จะอ้างอิงอัตราส่วนระหว่างจำนวน positive tuples ที่ถูกจำแนกได้อย่างถูกต้อง แต่สำหรับค่า specificity จะถูกคำนวณจากอัตราส่วนระหว่างจำนวน negative tuples ที่ถูกจำแนกได้อย่างถูกต้อง นอกจากนี้ อาจจะใช้ค่า “precision” เพื่อใช้ในการคำนวณเปอร์เซ็นต์ของข้อมูลของหมวดหมู่ที่เป็น positive ซึ่งถูกจำแนกเป็นหมวดหมู่ positive โดยค่าทั้ง 3 สามารถคำนวณได้ดังสมการที่ (2-18), (2-19) และ (2-20)

$$sensitivity = \frac{|true\ positives|}{|positive\ tuples|} \quad (2-18)$$

$$specificity = \frac{|true\ positive|}{|negative\ tuples|} \quad (2-19)$$



$$\text{precision} = \frac{|true\ positive|}{(|true\ positive| + |false\ positive|)} \quad (2-20)$$

จากการคำนวณค่าต่างๆ ข้างต้น เราสามารถคำนวณค่าความถูกต้องของตัวจำแนกข้อมูลได้ดังสมการที่ (2-21)

$$\begin{aligned} \text{accuracy} = & \text{sensitivity} \frac{|positive\ tuples|}{(|positive\ tuples| + |negative\ tuples|)} \\ & + \text{specificity} \frac{|negative\ tuples|}{(|positive\ tuples| + |negative\ tuples|)} \end{aligned} \quad (2-21)$$

### 2.7.2 ตัวชี้วัดความผิดพลาดจากการทำนายข้อมูล [17, 27]

กำหนดให้  $D^T$  คือ ชุดข้อมูลทดสอบที่อยู่ในรูปแบบ  $(X_1, y_1), (X_2, y_2), \dots, (X_d, y_d)$  ซึ่ง  $X_i$  เป็นเซตของข้อมูลที่มีหมวดหมู่  $y_i$  และ  $d$  คือ จำนวนข้อมูลทั้งหมดในชุดข้อมูลทดสอบ  $D^T$  โดยในการทำนายข้อมูลจะคืนค่าเป็นเปอร์เซ็นต์ของความเป็นหมวดหมู่ข้อมูล  $y_i$  สำหรับ  $X_i$  จะสามารถทำได้ค่อนข้างยาก ดังนั้นในการประเมินความถูกต้อง เราจะไม่ทำนายหมวดหมู่  $y_i$  สำหรับ  $X_i$  แล้วนำไปเทียบกับหมวดหมู่  $y_i$  ที่แท้จริงของ  $X_i$  แต่เราจะพิจารณาหาความแตกต่างระหว่างผลลัพธ์ที่ได้จากการทำนายกับค่าที่แท้จริง โดยเราสามารถใช้ “Loss functions” ในการตรวจวัดความผิดพลาดระหว่างหมวดหมู่ที่แท้จริง  $y_i$  เทียบกับหมวดหมู่ที่ได้จากการทำนาย  $y_i'$  ซึ่งจะมีฟังก์ชันการทำงานดังสมการที่ (2-22) และ (2-24)

$$\text{Absolute Error} : |y_i - y_i'| \quad (2-22)$$

$$\text{Squared Error} : (y_i - y_i')^2 \quad (2-23)$$

จากฟังก์ชันข้างต้น เราสามารถคำนวณหาข้อผิดพลาดในการทดสอบ (error rate) ซึ่งเป็นค่าเฉลี่ยของความผิดพลาดของการทำนายหมวดหมู่ของข้อมูลในชุดข้อมูลทดสอบได้ดังสมการที่ (2-24) และ (2-25)

$$\text{Mean Absolute Error} : \frac{\sum_{i=1}^d |y_i - y_i'|}{d} \quad (2-24)$$

$$\text{Mean Squared Error} : \frac{\sum_{i=1}^d (y_i - y_i')^2}{d} \quad (2-25)$$

อย่างไรก็ตาม ในบางครั้งเราอาจต้องการคำนวณหาค่าความผิดพลาดที่มีความสัมพันธ์กับสิ่งที่จะเป็นเมื่อเราทำนายหมวดหมู่ข้อมูล ซึ่งจะเป็นการคำนวณอัตราส่วนระหว่างผลรวมของความแตกต่างของผลลัพธ์จากการทำนายกับค่าจริงเทียบกับการสูญเสียที่เกิดจากการทำนายข้อมูล โดยค่าความสัมพันธ์กับค่าความผิดพลาดสามารถคำนวณได้ดังสมการที่ (2-26) และ (2-27)

$$\text{Relative Absolute Error} : \frac{\sum_{i=1}^d |y_i - y_i'|}{\sum_{i=1}^d |y_i - \bar{y}|} \quad (2-26)$$

$$\text{Relative Squared Error} : \frac{\sum_{i=1}^d (y_i - y_i')^2}{\sum_{i=1}^d (y_i - \bar{y})^2} \quad (2-27)$$

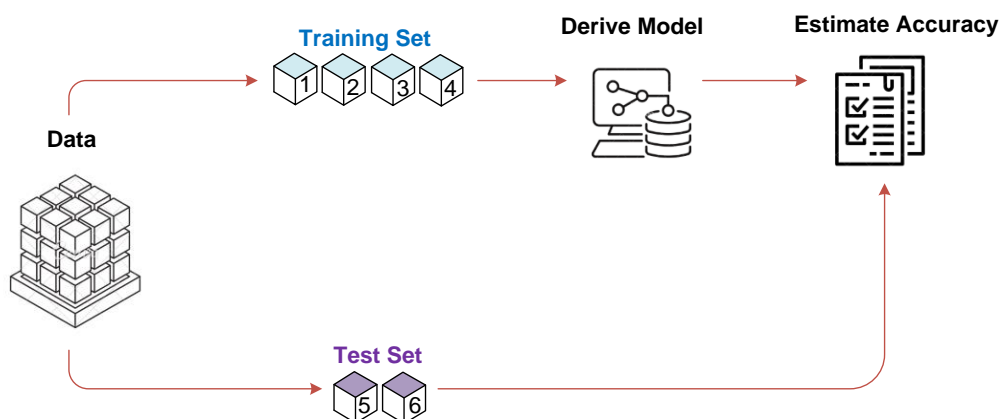
เมื่อ  $\bar{y}$  คือ ค่าเฉลี่ยของค่าหมวดหมู่  $y_i$  ในชุดข้อมูลฝึกฝน ที่ซึ่งสามารถคำนวณได้จาก  $\bar{y} = \frac{\sum_{i=1}^d y_i}{d}$

### 2.7.3 การประเมินความถูกต้องของตัวจำแนกหรือตัวทำนายข้อมูล [17]

จากตัวชี้วัดความถูกต้องของการจำแนก/ทำนายข้อมูลข้างต้น เราสามารถประยุกต์ใช้ตัวชี้วัดเหล่านั้นเพื่อประเมินความถูกต้องของตัวจำแนกข้อมูลหรือตัวทำนายข้อมูลได้อย่างเชื่อถือได้อย่างไร ดังนั้นในการจะประเมินความถูกต้องของตัวจำแนกข้อมูลหรือตัวทำนายข้อมูลได้อย่างเชื่อถือได้สามารถใช้เทคนิคต่างๆ ที่ได้รับความนิยมดังนี้

#### วิธีการ Holdout จุฬาลงกรณ์มหาวิทยาลัย

Holdout แบ่งชุดข้อมูลออกเป็น 2 ชุด ด้วยวิธีการสุ่มโดยชุดข้อมูลที่ได้จะเป็นชุดข้อมูลฝึกฝนและชุดข้อมูลทดสอบ ซึ่งโดยปกติ ชุดข้อมูลฝึกฝนจะมีปริมาณข้อมูลเท่ากับ 2 ใน 3 ของชุดข้อมูลทั้งหมดและ ชุดข้อมูลทดสอบจะมีปริมาณข้อมูลเท่ากับ 1 ใน 3 ของชุดข้อมูล หลังจากแบ่งชุดข้อมูลแล้วชุดข้อมูลฝึกฝนจะถูกใช้ในการสร้างตัวจำแนกข้อมูล และชุดข้อมูลทดสอบจะถูกใช้ในการทำสอบตัวจำแนกที่สร้างขึ้นดังรูปที่ 2-25

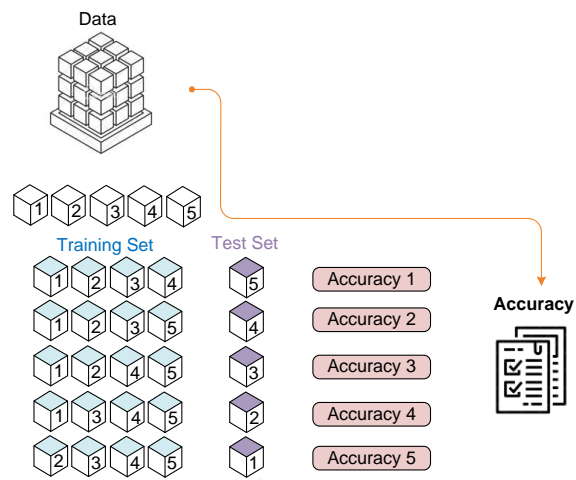


รูปที่ 2-25 การประเมินความถูกต้องด้วยวิธี Holdout [17]

### วิธี Cross-Validation

วิธีการ Cross-validation หรือที่นิยมเรียกว่า “k-fold cross validation” เริ่มจากการสุ่มแบ่งข้อมูลออกเป็น  $k$  ส่วน ซึ่งมีข้อมูลไม่ซ้ำกัน โดยชุดข้อมูลย่อย  $D_1, D_2, \dots, D_k$  มีจำนวนข้อมูลเท่ากันจากนั้นจะเรียนรู้และทดสอบตัวจำแนกข้อมูลทั้งสิ้น  $k$  ครั้ง โดยในการทำงานรอบที่  $i$  เรากำหนดให้ชุดข้อมูลย่อย  $D_i$  เป็นชุดข้อมูลทดสอบ และชุดข้อมูลอื่นๆ เป็นชุดข้อมูลการสอน

การทำงานของ k-fold cross validation พิจารณาข้อมูลเป็นจำนวนเท่าๆกัน ซึ่งทำหน้าที่เป็นข้อมูลฝึกฝนทั้งสิ้น  $k-1$  และทำหน้าที่เป็นข้อมูลทดสอบ 1 ครั้ง โดยการประเมินค่าความถูกต้องสามารถทำได้ในรอบที่  $k$  ของการทำงาน โดยในการประเมินค่าความถูกต้องจากการจำแนกข้อมูลในการทำงานรอบที่  $k$  สามารถคำนวณได้จากจำนวนข้อมูลที่ถูกต้องหารด้วยจำนวนข้อมูลทั้งหมดในชุดข้อมูล แต่สำหรับการทำนายข้อมูล เราสามารถประเมินค่าความผิดพลาดได้จากจำนวนค่าความสูญเสียในรอบที่  $k$  หารด้วยจำนวนข้อมูลทั้งหมดในชุดข้อมูล ซึ่งหากกำหนดค่า  $k = 5$  จะเรียกว่า 5-fold cross-validation ดังรูปที่ 2-26



รูปที่ 2-26 การประเมินความถูกต้องด้วยวิธี 5-fold cross validation



### บทที่ 3

#### การออกแบบระบบและพัฒนาระบบ

งานวิจัยก่อนหน้า [10, 28] ได้ประยุกต์ใช้รูปคลื่นแรงดันและกระแสที่ได้มาจากเครื่องบันทึกความผิดพลาดแบบดิจิทัล (DFR) มาจำแนกสาเหตุของความผิดพลาดบนสายส่ง ได้แก่ ต้นไม้ พายุ และสัตว์ จากข้อมูลรายงานไฟฟ้าขัดข้องพื้นที่ของการไฟฟ้าส่วนภูมิภาค เขต 2 ภาคใต้ [3] ปี 2560 ถึง ปี 2562 พบว่าสาเหตุหลัก 3 อันดับแรก เป็นสาเหตุจาก อุปกรณ์ชำรุด 42.3%, ต้นไม้ 32.4%, สัตว์ 13.1% ซึ่งงานวิจัยก่อนหน้าไม่มีการจำแนกสาเหตุจากอุปกรณ์ชำรุด และเป็น การจำแนกความผิดพลาดบนสายส่งซึ่งมีลักษณะเป็นสายเปลือยและระดับแรงดันที่สูงกว่า ระดับความสูงของเสาส่งและสภาพพื้นที่ที่สายส่งพาดผ่านจะมีลักษณะแตกต่างจากสายจำหน่ายแบบเหนื่อดินของ กฟภ. ซึ่งสายจำหน่ายของ กฟภ. มีลักษณะเป็นสายหุ้มฉนวนชนิด Space Aerial Cable (SAC) เป็นส่วนใหญ่ จึงอาจมีรายละเอียดสาเหตุและคุณลักษณะขณะเกิดความผิดพลาดที่แตกต่างกัน และยังระบุเพียงสาเหตุความผิดพลาดเป็นกลุ่ม ๆ เท่านั้น ไม่ได้ระบุความผิดพลาดเฉพาะให้ละเอียดขึ้น เช่น ภูเขาไฟ หรือ ทางมะพร้าว จากงานวิจัย [4] ข้อมูลนำเข้า (Input) ที่ใช้วิเคราะห์เป็นเชิงกายภาพ ซึ่งข้อมูลเชิงกายภาพที่ กฟภ. มีนั้นยังไม่สมบูรณ์ และวิธีที่ใช้ในระบบจำหน่ายของ กฟภ. สามารถบันทึกค่าสัญญาณรูปคลื่นกระแสและแรงดันเมื่อเกิดความผิดพลาดได้

การเรียนรู้ของเครื่อง (Machine Learning) [10, 28] คือศาสตร์ในการให้คอมพิวเตอร์เรียนรู้จากชุดข้อมูลที่เก็บไว้มาสร้างแบบจำลอง (Model) ที่จะนำไปตอบอนาคตว่าหากข้อมูลนำเข้า (Input) แบบนี้จะได้อะไร ซึ่งในการจำแนกสาเหตุของความผิดพลาดในระบบไฟฟ้าจำเป็นต้องอาศัยข้อมูลที่ค่อนข้างมาก ในสถานการณ์จริงเมื่อเกิดความผิดพลาดจำเป็นต้องทราบถึงสาเหตุให้รวดเร็วที่สุด ซึ่งมนุษย์เองก็มีข้อจำกัด รวมทั้งความเหนื่อยล้า ประสบการณ์ มุมมองที่มีต่อข้อมูล แต่ในขณะที่คอมพิวเตอร์สามารถทำงานได้รวดเร็วกว่า โดยที่ยังคงรักษาประสิทธิภาพการทำงานได้อย่างคงที่ นอกจากนี้ งานวิจัย [10, 28] ชี้ให้เห็นว่า Decision tree ยังสามารถจำแนกได้ค่อนข้างน้อย ต้องสร้างเงื่อนไข และเกณฑ์ตัดสินใจเอง ซึ่งหากข้อมูลมีความซับซ้อนมากอาจทำให้ยากต่อการจำแนกด้วยมนุษย์ โดยที่ Machine Learning สามารถวิเคราะห์ความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลจำนวนมากและเปลี่ยนแปลงตลอดเวลาได้ และไม่อาจทราบล่วงหน้าได้ถึงความสัมพันธ์นั้นว่าคือความสัมพันธ์ของสิ่งใดกับสิ่งใด

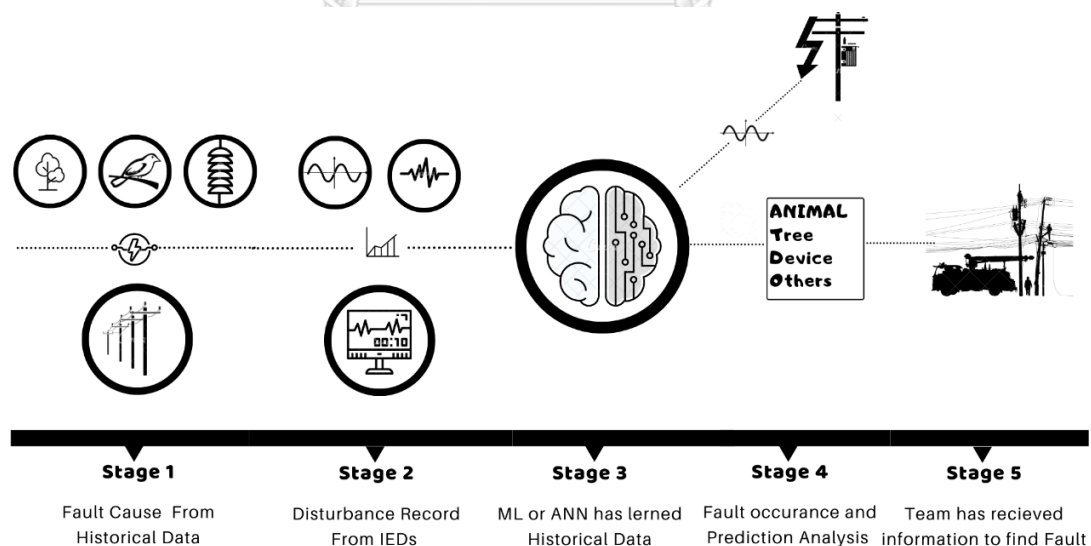
งานวิจัย [4] ใช้ Artificial Neural Network (ANN) มาจำแนกสาเหตุความผิดพลาดในระบบจำหน่ายที่เกิดจาก สัตว์ และ อื่น ๆ โดยใช้ข้อมูลเชิงกายภาพขณะเกิดความผิดพลาดเช่น สภาพอากาศ

ภูมิประเทศ ฤดูกาล วัน เวลาที่เกิด ฯลฯ ซึ่งบันทึกโดยผู้ปฏิบัติงานมาวิเคราะห์และเป็นข้อมูลนำเข้า (Input) ให้ Artificial Neural Network จำแนก โดยสามารถจำแนกได้ถูกต้องถึง 98%

ดังนั้นวิทยานิพนธ์จึงนำเสนอแนวทางการจำแนกสาเหตุของความผิดปกติในสายจำหน่ายเหนือดิน โดยจำแนกเป็นสาเหตุที่เกิดจากกลุ่มของ อุปกรณ์ชำรุด ต้นไม้ และสัตว์ และจำแนกเพิ่มเติมภายในแต่ละกลุ่ม โดยในกลุ่ม 1. อุปกรณ์จะจำแนกเป็นสาเหตุจาก สาย Overhead Ground Wire (OHGW), Spacer และอุปกรณ์อื่น ๆ 2. สัตว์จำแนกเป็นสาเหตุจาก นก งู และสัตว์อื่น ๆ 3. ต้นไม้จำแนกเป็นสาเหตุ ทางมะพร้าว และกิ่งไม้อื่น ๆ แต่ละสาย ดังแสดงในรูปที่ 3-1 และ 3-2 ด้วยการใช้ Machine Learning และ Artificial Neural Network มาจำแนก ซึ่งอธิบายถึงภาพรวมและรายละเอียดต่าง ๆ ดังต่อไปนี้

### 3.1 ภาพรวมของระบบ

งานวิจัยนี้สนใจการนำข้อมูลรูปคลื่นกระแสและแรงดันขณะเกิดความผิดปกติ มาวิเคราะห์และทำนายสาเหตุของความผิดปกติว่าเกิดจากอะไร โดยการรวบรวมข้อมูลรูปคลื่นของเหตุการณ์ที่เคยเกิดขึ้นในอดีตมาจำแนกคุณลักษณะเด่นของแต่ละสาเหตุจากนั้นให้ เครื่องมือที่ใช้ในการจำแนกซึ่งในวิทยานิพนธ์นี้ใช้ Machine Learning หรือ Artificial Neural Network เรียนรู้คุณลักษณะเด่นดังกล่าว และเมื่อมีความผิดปกติเกิดขึ้นอัลกอริทึมจะจำแนกคุณลักษณะและทำนายสาเหตุของความผิดปกติว่าเกิดจากสาเหตุอะไร ดังรูปที่ 3-1

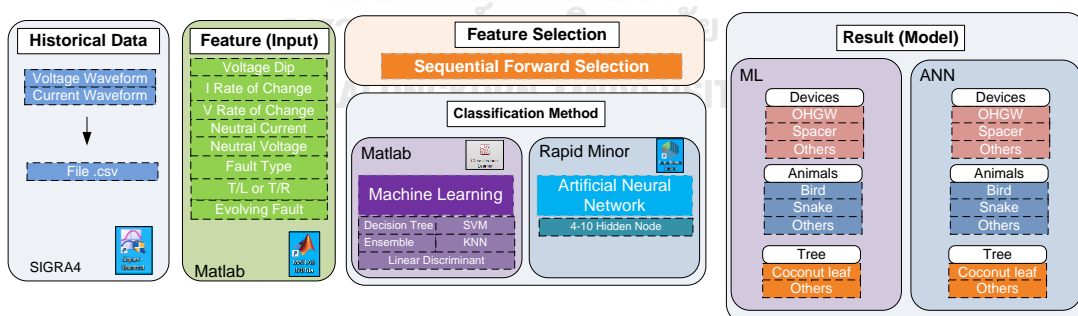


รูปที่ 3-1 ภาพรวมแนวคิดของการจำแนกสาเหตุความผิดปกติในระบบจำหน่าย

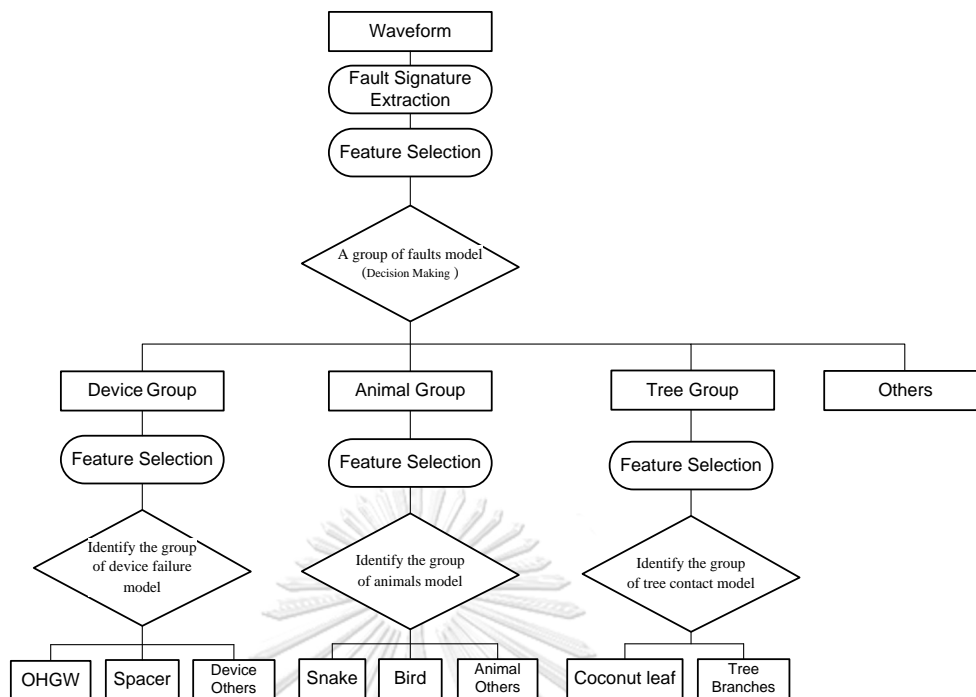
### 3.2 รายละเอียดของระบบการทำงาน

ในวิทยานิพนธ์นี้ประยุกต์ใช้ Machine Learning และ Artificial Neural Network ในการจำแนกสาเหตุความผิดปกติพ่วงจาก 3 กลุ่มสาเหตุ และระบุถึงสาเหตุภายในแต่ละกลุ่ม โดยมีรายละเอียดขั้นตอนระบบการทำงานดังนี้ และไดอะแกรมระบบการทำงานแสดงดังภาพที่ 3-2

- 1) ใช้ข้อมูลรูปคลื่นขณะเกิดความผิดปกติในอดีตที่บันทึกได้จากรีเลย์ที่สถานีไฟฟ้า แปลงไฟล์เป็นนามสกุล .csv ผ่านโปรแกรม SIGRA4
- 2) สกัดคุณลักษณะเด่นออกเป็น 8 คุณลักษณะ ผ่านอัลกอริทึมที่ถูกเขียนไว้เพื่อสกัดคุณลักษณะเด่นโดยโปรแกรม Matlab
- 3) คัดเลือกคุณลักษณะเด่นและจำแนกสาเหตุความผิดปกติโดยแบ่งเป็น 2 การทดสอบ โดยวิธีการจำแนกด้วย Machine Learning ใช้โปรแกรม Matlab และ การจำแนกด้วย Artificial Neural Network ใช้โปรแกรม RapidMiner
- 4) ขั้นตอนของการจำแนกแบ่งเป็น 2 ชั้น ชั้นแรก จะจำแนกสาเหตุออกเป็นกลุ่มประกอบด้วย กลุ่มของ อุปกรณ์ สัตว์ ต้นไม้ และอื่นๆ และ ชั้นที่สอง จำแนกสาเหตุย่อยภายในแต่ละกลุ่ม โดยกลุ่มของอุปกรณ์ประกอบด้วยสาเหตุจาก สาย Over Head Ground Wire (OHGW), Spacer และอุปกรณ์อื่นๆ กลุ่มของสัตว์ประกอบด้วยสาเหตุจาก งู นก และสัตว์อื่นๆ และกลุ่มของต้นไม้ประกอบด้วย ทางมะพร้าวแตะแนวสาย และกิ่งไม้อื่นๆ ดังรูปที่ 3-3
- 5) ทดสอบและเปรียบเทียบผลลัพธ์ของโมเดล



รูปที่ 3-2 ไดอะแกรมแสดงรายละเอียดของระบบการทำงาน



รูปที่ 3-3 ไตอะแกรมขั้นตอนการจำแนกสาเหตุความผิดพลาดในระบบจำหน่ายของ กฟภ.

### 3.3 ข้อมูลที่นำมาใช้ในการวิเคราะห์

จากหลายๆ งานวิจัยที่เกี่ยวข้องในอดีต พบว่ามีการใช้ข้อมูลที่มีการบันทึกเมื่อเกิดความผิดพลาดเป็นข้อมูลนำเข้า (Input) เพื่อใช้ในการวิเคราะห์ โดยข้อมูลที่นำมาวิเคราะห์นั้นมีหลากหลายรูปแบบต่างกันไป ซึ่งอาจแบ่งตามรูปแบบในการบันทึกได้เป็น 1. ข้อมูลที่ถูกบันทึกโดยคน หรือผู้ปฏิบัติงาน 2. ข้อมูลที่ถูกบันทึกโดยเครื่องมือวัดที่ติดตั้งในระบบไฟฟ้า

ข้อมูลที่ถูกจดบันทึกโดยคนนั้นบางครั้งอาจมีการคลาดเคลื่อน และไม่ไปในเกณฑ์เดียวกันได้ หรือมุมมองที่มีต่อข้อมูลต่างกัน เช่น สภาพอากาศฝนตก บางคนอาจพิจารณาและบันทึกข้อมูลว่าฝนตกฟ้าคะนอง หรือพนักงานบางคนไม่เห็นความสำคัญของข้อมูลจึงอาจละเลยการบันทึกถึงรายละเอียดของความผิดพลาดบางอย่างไปทำให้ข้อมูลไม่ครบถ้วน ส่วนข้อมูลที่ได้จากเครื่องมือวัดนั้นจะมีความเที่ยงตรง มีความสม่ำเสมอในการบันทึกข้อมูลมากกว่า ดังนั้นในวิทยานิพนธ์นี้จึงใช้ข้อมูลที่ได้จากการบันทึกโดยรีเลย์เพียงอย่างเดียว โดยจะใช้ข้อมูลของรูปคลื่นแรงดันและกระแสขณะเกิดความผิดพลาดดังรูปที่ 3-4 แปลงเป็นไฟล์นามสกุล .csv ดังรูปที่ 3-4 ผ่านโปรแกรม SIGRA4 มาใช้เป็นข้อมูลนำเข้า (Input) ในการวิเคราะห์เพื่อจำแนกสาเหตุความผิดพลาดในระบบจำหน่ายของ กฟภ.



รูปที่ 3-4 ตัวอย่างข้อมูลรูปคลื่นกระแส และแรงดันเมื่อแปลงเป็นไฟล์นามสกุล .csv

### 3.4 การสกัดคุณลักษณะเด่น

ในวิทยานิพนธ์นี้จะนำข้อมูลรูปคลื่นกระแสและแรงดันขณะเกิดความผิดปกติมาสกัดคุณลักษณะเด่น โดยคัดเลือกประเภทของคุณลักษณะเด่นที่สามารถสกัดได้จากข้อมูลรูปคลื่นกระแสและแรงดันที่ได้เสนอไว้ในหัวข้อที่ 3.3 ประกอบไปด้วย

- 1) ค่าแรงดันตกชั่วขณะ [10]

ค่าแรงดันตกชั่วขณะคำนวณจากค่าความแตกต่างของแรงดันบัสในเวลาปกติกับค่าแรงดันบัสขณะเกิดความผิดปกติสามารถคำนวณดังสมการ (3-1)

$$\% \text{Voltage dip} = \frac{V_p - V_{f,\min}}{V_p} \times 100 \tag{3-1}$$

โดยที่  $V_p$  คือ ค่าแรงดันก่อนเกิดความผิดปกติ

$V_{f,\min}$  คือ ค่าแรงดันต่ำสุดระหว่างที่เกิดความผิดปกติ

- 2) อัตราการเปลี่ยนแปลงค่าของกระแส [18]

ค่าอัตราการเปลี่ยนแปลงของกระแส (di/dt) พิจารณาจากค่าสูงสุดของผลต่างของค่า RMS ของค่ากระแส ณ จุดที่ n+1 กับจุดที่ n ทหารด้วยผลต่างทางเวลา ณ จุดที่ n+1 กับจุดที่ n ค่าสูงสุดของอัตราการเปลี่ยนแปลงกระแสจะบ่งบอกถึงตำแหน่งของเวลาที่เป็นจุดเริ่มต้นของความผิดปกติคือสถานะชั่วครู่ของความผิดปกติ (transient) ดังสมการที่ (3-2)

$$I_{\text{Rate Of Change}} = \max_{n=1}^N \left[ \frac{I_{rms}(t_{n+1}) - I_{rms}(t_n)}{t_{n+1} - t_n} \right] \quad (3-2)$$

โดยที่  $I_{rms}$  คือ รากที่สองของกำลังสองเฉลี่ยของค่ากระแสเฟสที่เกิดความผิดพลาด  
 $N$  คือ จำนวนของจุดของการสุ่มค่ากระแสที่เกิดความผิดพลาด

### 3) กระแสนิวทรัล

ค่ากระแสนิวทรัล คือ ค่ารวมของกระแสลัดวงจรทั้งสามเฟสที่ผ่านลงดิน ในสภาวะปกติของระบบจำหน่ายที่กระแสสมดุลทั้งสามเฟสและไม่มีฮาร์มอนิก ค่ากระแสนิวทรัลจะมีค่าเป็นศูนย์หรือใกล้เคียงศูนย์ แต่ในสภาวะที่เกิดความผิดพลาดขึ้นในระบบจำหน่ายที่กระแสสามารถไหลลงสู่พื้นดินได้จะทำให้เกิดความไม่สมดุลขึ้นในระบบและมีค่ากระแสนิวทรัลเพิ่มขึ้นแบบทันทีทันใดซึ่งเป็นอีกสาเหตุหลักสาเหตุหนึ่งที่ทำให้เกิดแรงดันตกค้างที่กล่าวมาข้างต้น

### 4) แรงดันนิวทรัล

ค่าแรงดันนิวทรัล คือ ค่ารวมของแรงดันทั้งสามเฟส ในสภาวะปกติของระบบจำหน่ายอยู่ในสภาวะสมดุลทั้งสามเฟส ค่าแรงดันรวมทั้งสามเฟสจะเป็นศูนย์หรือใกล้เคียงศูนย์ แต่ในสภาวะที่เกิดความผิดพลาดขึ้นในระบบ ค่าแรงดันนิวทรัลจะเพิ่มขึ้นแบบทันทีทันใดซึ่งอาจสามารถบ่งชี้ถึงสาเหตุความผิดพลาดที่สนใจ

### 5) อัตราการเปลี่ยนแปลงของแรงดัน [18]

การหาค่าอัตราการเปลี่ยนแปลงของแรงดัน ( $dv/dt$ ) โดยพิจารณาจากผลต่างระหว่างค่า RMS ณ จุดที่  $n+1$  และจุดที่  $n$  ต่อค่าระยะเวลา ระหว่างจุดที่  $n+1$  กับจุดที่  $n$  และหาค่าสูงที่สุดของอัตราค่าเปลี่ยนแปลงแรงดัน ณ ตำแหน่งที่มีการเปลี่ยนแปลงของแรงดัน ณ ตำแหน่งที่มีการเปลี่ยนแปลงของแรงดันสูงที่สุดจะเป็นตำแหน่งที่บ่งบอกถึงจุดเริ่มต้นหรือเวลาของการเริ่มเกิดความผิดพลาดเป็นช่วงที่อยู่ในสภาวะชั่วคราว (transient) ดังสมการที่ (3-3)

$$V_{\text{Rate Of Change}} = \max_{n=1}^N \left( \frac{V_{rms}(t_{n+1}) - V_{rms}(t_n)}{t_{n+1} - t_n} \right) \quad (3-3)$$

โดยที่  $V_{rms}$  คือ รากที่สองของกำลังสองเฉลี่ยของค่าแรงดันเฟสที่เกิดความผิดพลาด  
 $N$  คือ จำนวนของจุดของการสุ่มค่าแรงดันเฟสที่เกิดความผิดพลาด

6) ความผิดปกติพร้อมแบบชั่วคราวและความผิดปกติพร้อมแบบถาวร [29]

การทำงานของ Recloser แบ่งออกเป็น 2 กลุ่มคือ

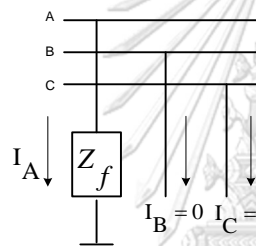
4.1 Recloser สั่งการให้เซอร์กิตเบรกเกอร์ Trip Reclose สำเร็จ เรียกว่า ความผิดปกติพร้อมแบบชั่วคราว (ระบบสามารถกำจัดสาเหตุของความผิดปกติได้)

4.2 Recloser สั่งการให้เซอร์กิตเบรกเกอร์ Trip Reclose ไม่สำเร็จ เรียกว่าความผิดปกติพร้อมแบบถาวร (ระบบไม่สามารถกำจัดสาเหตุความผิดปกติได้)

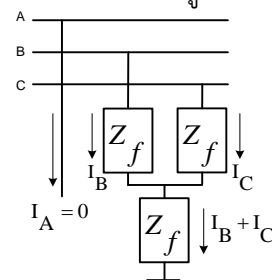
7) การแยกแยะชนิดของความผิดปกติ [30]

การแยกแยะชนิดความผิดปกติ โดยแยกแยะชนิดความผิดปกติออกเป็น 4 ชนิด คือ 1. แบบเฟสเดียวลงดิน 2. แบบคู่เฟสลงดิน 3. แบบเฟสต่อเฟส 4. แบบสามเฟสลงดิน ดังแสดงในรูปที่ 3-5

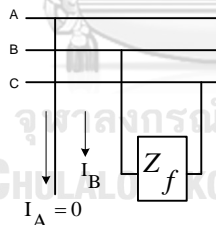
1. ความผิดปกติพร้อมแบบเฟสเดียวลงดิน



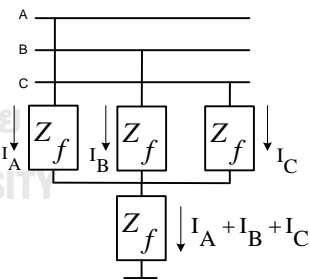
2. ความผิดปกติพร้อมแบบคู่เฟสลงดิน



3. ความผิดปกติพร้อมแบบเฟสต่อเฟส



4. ความผิดปกติพร้อมแบบสามเฟสลงดิน



รูปที่ 3-5 ตัวอย่างความผิดปกติชนิดต่าง ๆ

8) ความผิดปกติที่มีการแปรเปลี่ยนประเภท

ความผิดปกติที่มีการแปรเปลี่ยนประเภท (Evolving Fault) คือ การที่ความผิดปกติสามารถเปลี่ยนประเภทของความผิดปกติได้เมื่อระยะเวลาของการเกิดไม่แตกต่างกันมากนัก เช่น ความผิดปกติที่เกิดจากสเปเซอร์อาจเกิดความผิดปกติชนิดเฟสต่อเฟส และเปลี่ยนแปลงเป็นคู่เฟสลงดิน

### 3.5 การเลือกคุณลักษณะเด่น

จากงานวิจัยในอดีต [18] พบว่าการเลือกคุณลักษณะ (Feature Selection) มีวิธีการแบ่งเป็น 2 กลุ่ม หลักๆคือ 1. วิธีแรปเปอร์ (Wrapper Method) 2. วิธีฟิลเตอร์ (Filter Method)

เนื่องจากวิธี Wrapper Method นั้นใช้ค่าความแม่นยำในการจำแนกในการเลือกชุดคุณลักษณะที่ให้ค่าความแม่นยำในการจำแนกสูงสุด ต่างจากวิธี Filter Method ที่มักจะลำดับความสำคัญของคุณลักษณะต่างๆ มากกว่าที่จะให้ชุดคุณลักษณะเด่นที่เหมาะสมกับโมเดลการจำแนกมากที่สุด ดังนั้นงานวิจัยนี้จึงเลือกใช้วิธี Wrapper Method โดยประยุกต์ใช้การคัดเลือกคุณลักษณะเด่นแบบ Sequential Forward Selection โดยจะแสดงตัวอย่างการเลือกคุณลักษณะเด่นอย่างง่ายดังต่อไปนี้

$$J(x) = -2x_1x_2 + 3x_1 + 5x_2 - 2x_1x_2x_3 + 7x_3 + 4x_4 - 2x_1x_2x_3x_4$$

โดยที่  $J(x)$  คือ objective function

$x_1, x_2, x_3, x_4$  คือ คุณลักษณะเด่น

1. เลือกคุณลักษณะเด่นมา 1 ค่า ได้ทั้งหมด  $({}^4C_1) = \frac{4!}{1!(4-1)!} = 4$  จะเลือกได้ 4 วิธี

ดังนี้

$$J(x_1) = 3 \quad J(x_2) = 5 \quad J(x_3) = 7 \quad J(x_4) = 4$$

จะเห็นว่าคุณลักษณะเด่น  $x_3$  ทำให้เกิดค่า objective function มากที่สุด ดังนั้น  $x_3$  จึงถูกเลือกให้เป็นคุณลักษณะเด่นถัดไป

2. เลือกคุณลักษณะเด่น 2 ค่า โดยค่าแรกคือ  $x_3$  และค่าหลังต้องไม่ซ้ำกับค่าแรก จะได้ 3 วิธี คือ

$$J(x_3x_1) = 10 \quad J(x_3x_2) = 12 \quad J(x_3x_4) = 11$$

จะเห็นว่าคุณลักษณะเด่น  $x_3x_2$  ทำให้เกิดค่า objective function มากที่สุด ดังนั้น  $x_3x_2$  จึงถูกเลือกให้เป็นคุณลักษณะเด่นถัดไป

3. เลือกคุณลักษณะเด่น 3 ค่า โดย 2 ค่าแรกคือ  $x_3x_2$  และค่าหลังต้องไม่ซ้ำกับสองค่าแรก จะได้ 2 วิธี คือ

$$J(x_3x_2x_1) = 11 \quad J(x_3x_2x_4) = 16$$

จะเห็นได้ว่าคุณลักษณะเด่น  $x_3, x_2, x_4$  ทำให้เกิดค่า objective function มากที่สุด  
 ดังนั้น  $x_3, x_2, x_4$  จึงถูกเลือกให้เป็นคุณลักษณะเด่นถัดไป

4. เลือกคุณลักษณะเด่นทั้ง 4 ค่า ดังนี้

$$J(x_3, x_2, x_4, x_1) = 13$$

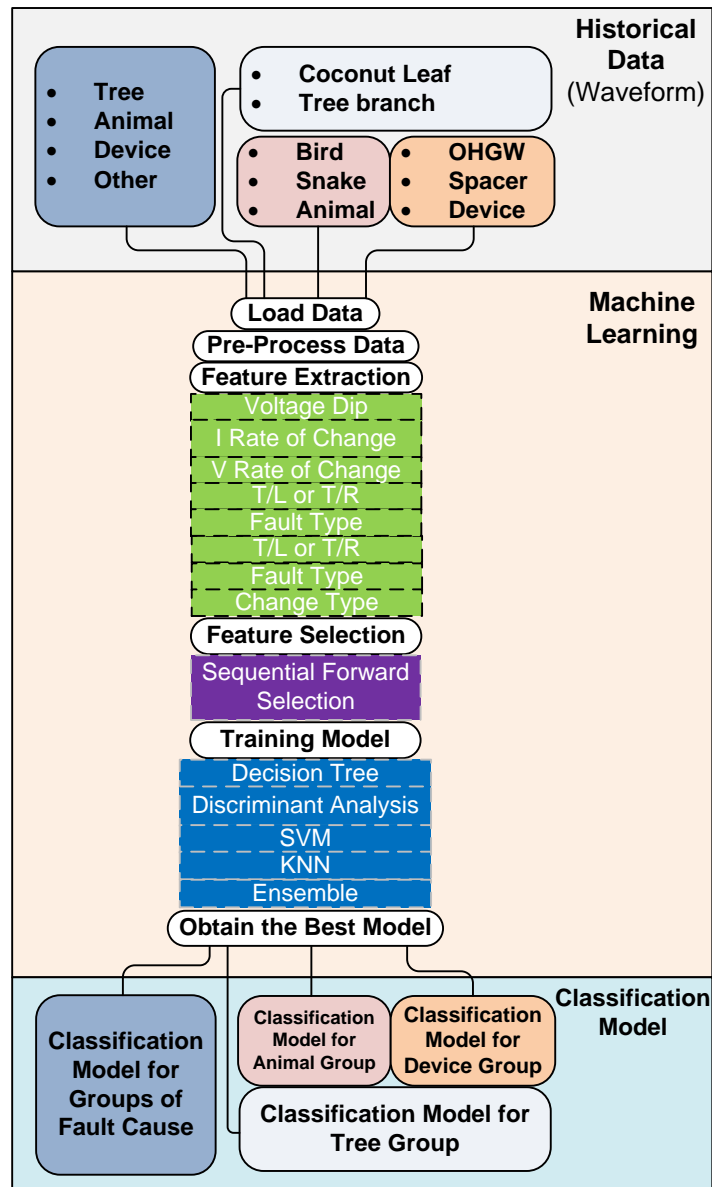
จากทั้ง 4 ขั้นตอน จะเห็นว่า ขั้นตอนที่ 4 ซึ่งเลือกคุณลักษณะเด่นทั้งหมด 4 ค่า มีค่า objective function น้อยกว่าขั้นตอนที่ 3 ที่เลือกคุณลักษณะเด่นเพียง 3 ค่า  $x_3, x_2, x_4$  ดังนั้นการใช้จำนวนคุณลักษณะเด่นที่มากกว่าอาจไม่ได้ส่งผลให้การจำแนกมีความแม่นยำสูงสุด จึงทำให้ต้องมีการเลือกชุดคุณลักษณะเด่นที่เหมาะสมกับโมเดลการจำแนกมากที่สุด

### 3.6 Training Model

Classification Model เป็น Supervised Model มีลักษณะเป็นการแบ่งกลุ่มของข้อมูลโดยมีต้นแบบของการแบ่งเป็นตัว Target ที่มีกลุ่มหรือจำนวนกลุ่มไม่มาก เช่น Yes/No

การทำ Classification มีอัลกอริทึมให้เลือกมากมาย เช่น Linear regression, SVM, Decision Tree, Neural Network หรือ Deep Learning ซึ่งในวิทยานิพนธ์นี้เลือกการประยุกต์ Machine Learning โดยใช้ Classification Learner App ในโปรแกรม Matlab เป็นเครื่องมือในการทำ Classification Model โดยสามารถทำการ Training Model ได้ 5 อัลกอริทึม คือ Decision Tree, Discriminant Analysis, SVM, KNN, Ensemble ได้พร้อมกันและนำผล Accuracy ของแต่ละอัลกอริทึมมาเปรียบเทียบกันเพื่อคัดเลือกอัลกอริทึมที่ดีที่สุด กระบวนการของการเรียนรู้ของ Machine Learning แสดงดังรูปที่ 3-6 และ Artificial Neural Network โดยใช้โปรแกรม Rapid Miner เป็นเครื่องมือในการทำ Classification Model แสดงในรูปที่ 3-7

### 3.6.1 กระบวนการของ Machine Learning

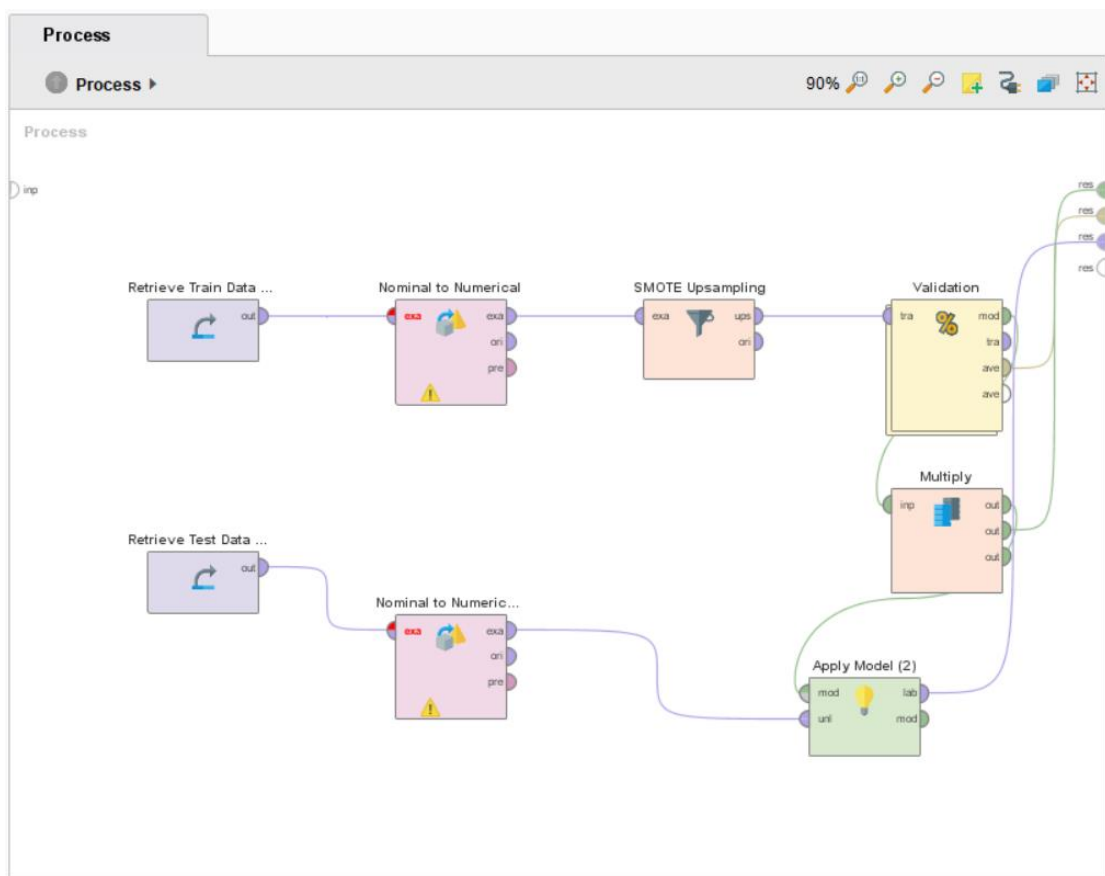


รูปที่ 3-6 กระบวนการของ Machine Learning (The Process of Machine Learning)

กระบวนการของ Machine Learning แสดงดังรูปที่ 3-6 ประกอบด้วย 1. การนำข้อมูลรูปคลื่นขณะเกิดความผิดปกติมารวบรวม 2. จัดการข้อมูลให้อยู่ในรูปแบบที่ต้องการใช้งาน 3. สกัดคุณลักษณะเด่น 4. เลือกคุณลักษณะเด่นที่มีประสิทธิภาพที่สุด 5. สร้าง Training Model และเลือกโมเดลที่ดีที่สุด 6. นำโมเดลไปทดสอบประสิทธิภาพต่อไป

### 3.6.2 กระบวนการของ Artificial Neural Network

กระบวนการ Training และ จำแนกสาเหตุด้วย Artificial Neural Network ในวิทยานิพนธ์นี้ใช้โปรแกรม RapidMiner ดังรูปที่ 3-7 ทดสอบโดยใช้ Artificial Neural Network แบบ 1 ชั้นซ่อน และเปลี่ยนค่าจำนวนโนดซ่อนตั้งแต่ 4 -10 โนด ใช้ชุดข้อมูลฝึกฝนที่ 80% ของจำนวนเหตุการณ์ทั้งหมด โดยวิธี 5-fold cross validation และคัดเลือกโมเดลที่ให้ผลลัพธ์ที่แม่นยำที่สุด



รูปที่ 3-7 ตัวอย่างการจำแนกสาเหตุด้วย Artificial Neural Network โดยโปรแกรม Rapid Miner

### 3.7 การประเมินความถูกต้องของโมเดล

ในงานวิจัยนี้มีการทดสอบโมเดล อยู่ 2 ส่วน ส่วนแรกเป็นการประเมินความถูกต้องภายใน Machine Learning และ Artificial Neural Network โดยใช้วิธี 5-fold cross validation ดังรูปที่ 2-26 เพื่อคัดเลือกโมเดลที่ดีที่สุด และส่วนที่สองเป็นการนำโมเดลที่ดีที่สุดมาทดสอบกับสถานการณ์จริงโดยการนำเหตุการณ์ความผิดปกติพร้อมใหม่มาให้โมเดลที่ได้ทำนายสาเหตุและประเมินความถูกต้อง

### 3.8 การนำไปใช้งาน

หลังจากได้โมเดลที่ผ่านการประเมินความถูกต้องตามที่ต้องการแล้ว เราสามารถนำวิธีการที่เสนอไปใช้งานจริงได้ตามขั้นตอนต่าง ๆ ดังต่อไปนี้

- 1) คอมพิวเตอร์ประจำศูนย์ควบคุมการจ่ายไฟ (Dispatching Center) จะมีซอฟต์แวร์เฉพาะ ชื่อว่า SIRGA4 เพื่อใช้ดูและวิเคราะห์รูปคลื่นเมื่อมีความผิดปกติ และ export ข้อมูลรูปคลื่นขณะเกิดความผิดปกติเป็น ไฟล์นามสกุล .csv เพื่อให้โปรแกรม Matlab วิเคราะห์ต่อไป
- 2) เมื่อเกิดความผิดปกติ รีเลย์ที่ถูกติดตั้งที่สถานีไฟฟ้า จะส่งข้อมูลรูปคลื่นขณะเกิดความผิดปกติผ่านระบบสื่อสารของ กฟภ. (ปัจจุบันยังอยู่ระหว่างการพิจารณาดำเนินการ) มายัง ศูนย์ควบคุมการจ่ายไฟ
- 3) พนักงานประจำศูนย์ควบคุมการจ่ายไฟ ได้รับข้อมูลรูปคลื่น จากนั้นเปิดไฟล์ด้วยโปรแกรม SIRGA4 และ export ออกมาเป็นไฟล์ นามสกุล .csv
- 4) ใช้อัลกอริทึมที่เขียนด้วยโปรแกรม Matlab สกัดคุณลักษณะเด่นเพื่อให้โมเดลทำนายสาเหตุ
- 5) เมื่อโมเดลทำนายสาเหตุแล้ว แจ้งกับผู้ปฏิบัติงานเพื่อตรวจสอบเคลียร์ไลน์ต่อไป



## บทที่ 4

### การพัฒนาและผลการทดสอบ

วิทยานิพนธ์ฉบับนี้ได้ออกแบบขั้นตอนการพัฒนาวิธีให้ได้มาซึ่งโมเดลที่ให้ผลการจำแนกแม่นยำมากที่สุด จากข้อมูลรูปคลื่นกระแสและแรงดันขณะเกิดความผิดปกติที่บันทึกได้จากรีเลย์ที่สถานีไฟฟ้า เกาะสมุย 1 (KMA), นครศรีฯ 2 (NTB), ปากพนัง (PPN), กาญจนดิษฐ์ (KCD), อ่าวลึก (ALA), ท้ายเหมือง (TMG), และ ป่าตอง (PAT) รวม 7 สถานีไฟฟ้า ในพื้นที่รับผิดชอบของการไฟฟ้าส่วนภูมิภาค เขต 2 ภาคใต้ ตั้งแต่ เดือนกุมภาพันธ์ พ.ศ. 2560 ถึง เดือนธันวาคม พ.ศ. 2562 จำนวน 200 เหตุการณ์ แบ่งเป็น ความผิดปกติที่เกิดจากสัตว์ 60 เหตุการณ์ ความผิดปกติที่เกิดจากต้นไม้ 60 เหตุการณ์ ความผิดปกติที่เกิดจากอุปกรณ์ 60 เหตุการณ์ และอื่นๆ 20 เหตุการณ์ เพื่อใช้จำแนกสาเหตุความผิดปกติที่เกิดกับระบบจำหน่าย 33 kV ของการไฟฟ้าส่วนภูมิภาค เขต 2 ภาคใต้ โดยใช้ Machine Learning และเปรียบเทียบความแม่นยำกับโมเดลที่จำแนกด้วย Artificial Neural Network

#### 4.1 การพัฒนาขั้นตอนวิธีการจำแนกด้วย Machine Learning

การพัฒนาขั้นตอนวิธีการจำแนกด้วย Machine Learning แบ่งออกเป็น 2 ขั้นตอน คือ 1. การคัดเลือกคุณลักษณะเด่นที่เหมาะสมตามจำนวนเหตุการณ์ที่นำมาวิเคราะห์ 2. ผลของจำนวนข้อมูลฝึกฝนกับความแม่นยำ

##### 4.1.1 การคัดเลือกคุณลักษณะเด่นที่เหมาะสมตามจำนวนเหตุการณ์ที่นำมาวิเคราะห์

ในการทดสอบนี้จะทำการสร้าง 4 โมเดลที่ให้ความแม่นยำสูงที่สุด คือ 1. โมเดลจำแนกสาเหตุความผิดปกติออกเป็นกลุ่มๆ 2. โมเดลจำแนกสาเหตุความผิดปกติภายในกลุ่มของอุปกรณ์ 3. โมเดลจำแนกสาเหตุความผิดปกติภายในกลุ่มของสัตว์ 4. โมเดลจำแนกสาเหตุความผิดปกติภายในกลุ่มของต้นไม้ โดยในแต่ละโมเดลจะคัดเลือกคุณลักษณะเด่นจากทั้งหมด 8 ค่า ที่ทำให้เกิดความแม่นยำในการจำแนกสูงสุด ทดสอบกับทุกอัลกอริทึมที่ Classification Learner Application ในโปรแกรม Matlab มี จากทั้งหมด 5 อัลกอริทึม คัดเลือกอัลกอริทึมและคุณลักษณะเด่นของแต่ละโมเดลที่ให้ความแม่นยำในการจำแนกสูงสุด

##### 1) การคัดเลือกคุณลักษณะเด่นที่เหมาะสมในการจำแนกสาเหตุความผิดปกติออกเป็นกลุ่มๆ

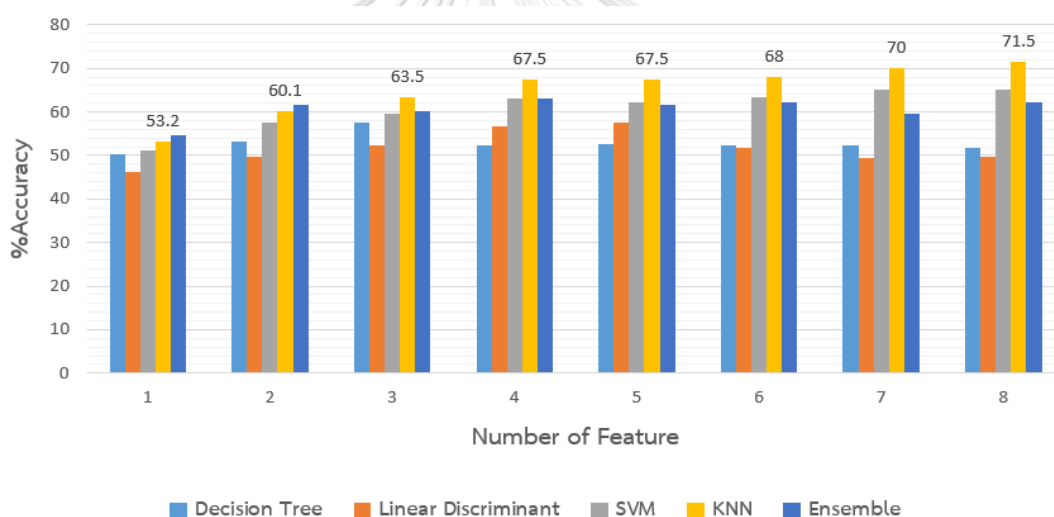
การทดสอบนี้จะคัดเลือกคุณลักษณะเด่นที่ใช้สำหรับจำแนกสาเหตุความผิดปกติออกเป็นกลุ่มๆที่มีความแม่นยำสูงที่สุด จากจำนวนข้อมูลฝึกฝนทั้งหมด 200 เหตุการณ์ ทดสอบความแม่นยำด้วยวิธี 5-fold cross validation ผลการทดสอบดังแสดงในตารางที่ 4-1 ถึง ตารางที่ 4-5





ตารางที่ 4-5 การคัดเลือกคุณลักษณะเด่นของการจำแนกเป็นกลุ่มด้วยวิธี Ensemble

N	ประเภทของคุณลักษณะเด่นที่ถูกเลือก								เครื่องมือที่ใช้จำแนก
	di/dt	In	Vn	dv/dt	reclose	type	change	Vdip	Ensemble
1	✓								54.7
2	✓			✓					61.6
3	✓			✓	✓				60.6
4	✓			✓	✓		✓		63.1
5	✓			✓	✓		✓	✓	61.6
6	✓			✓	✓	✓	✓	✓	62.1
7	✓	✓		✓	✓	✓	✓	✓	59.6
8	✓	✓	✓	✓	✓	✓	✓	✓	62.1



รูปที่ 4-1 ค่าความแม่นยำในการจำแนกเป็นกลุ่มของแต่ละอัลกอริทึมที่คุณลักษณะเด่นต่างกัน

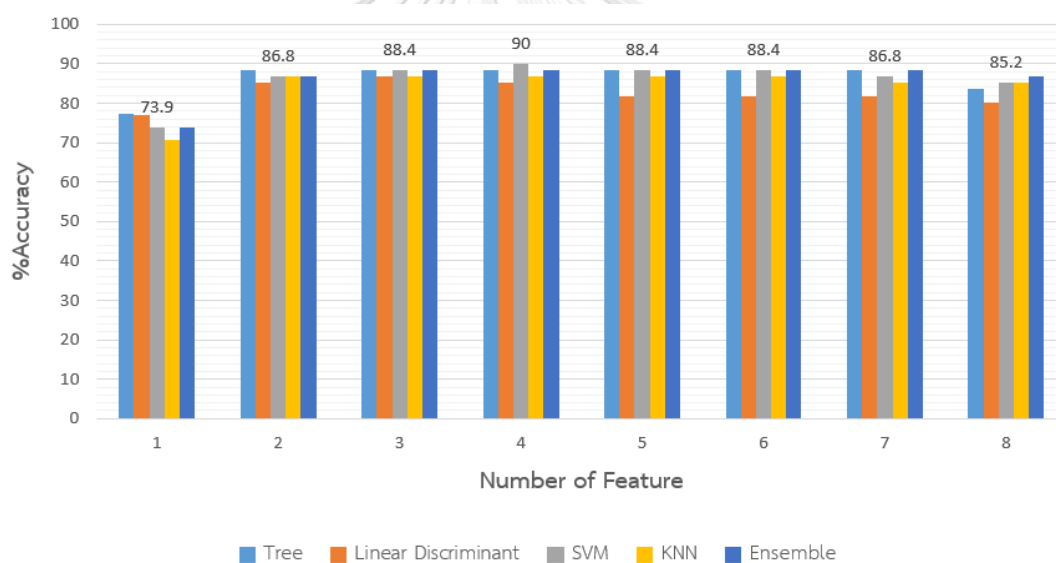
จากตารางที่ 4-4 และ รูปที่ 4-1 ภายใต้เงื่อนไขของคุณลักษณะเด่นและอัลกอริทึมที่ใช้ในการฝึกฝนต่างกัน พบว่าวิธี KNN ให้ความแม่นยำในการจำแนกสาเหตุความผิดปกติออกเป็นกลุ่มๆ สูงสุดที่ 71.50% ด้วย 8 คุณลักษณะเด่นคือ 1. อัตราการเปลี่ยนแปลงของกระแส (di/dt) 2. ค่าสูงสุดของกระแสนิวทรัล (In) 3. ค่าสูงสุดของแรงดันนิวทรัล (Vn) 4. อัตราการเปลี่ยนแปลงของแรงดัน (dv/dt) 5. ความผิดปกติแบบชั่วคราวหรือถาวร (reclose) 6. ชนิดของความผิดปกติ (type) 7. การแปรเปลี่ยนชนิดของความผิดปกติ (change) 8. ค่าแรงดันตกชั่วขณะ (Vdip) โดยมีค่า K = 10





ตารางที่ 4-10 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี Ensemble

N	ประเภทของคุณลักษณะเด่นที่ถูกเลือก								เครื่องมือที่ใช้จำแนก
	di/dt	In	Vn	dv/dt	reclose	type	change	Vdip	Ensemble
1								✓	73.9
2						✓		✓	86.8
3	✓					✓		✓	88.4
4	✓	✓				✓		✓	88.4
5	✓	✓				✓	✓	✓	88.4
6	✓	✓			✓	✓	✓	✓	88.4
7	✓	✓		✓	✓	✓	✓	✓	88.4
8	✓	✓	✓	✓	✓	✓	✓	✓	86.8



รูปที่ 4-2 ค่าความแม่นยำในการจำแนกภายในกลุ่มอุปกรณ์ของแต่ละอัลกอริทึม

จากตารางที่ 4-8 และ รูปที่ 4-2 ภายใต้งี้อื่นๆของคุณลักษณะเด่นและอัลกอริทึมที่ใช้ในการฝึกฝนต่างกัน พบว่าวิธี SVM ให้ความแม่นยำในการจำแนกสาเหตุความผิดปกติภายในกลุ่มของอุปกรณ์สูงสุดที่ 90% ด้วย 4 คุณลักษณะเด่นคือ 1. ค่าสูงสุดของกระแสไฟฟ้า (In) 2. อัตราการเปลี่ยนแปลงของแรงดัน (dv/dt) 3. การแปรเปลี่ยนชนิดของความผิดปกติ (change) 4. ค่าแรงดันตกชั่วขณะ (Vdip)

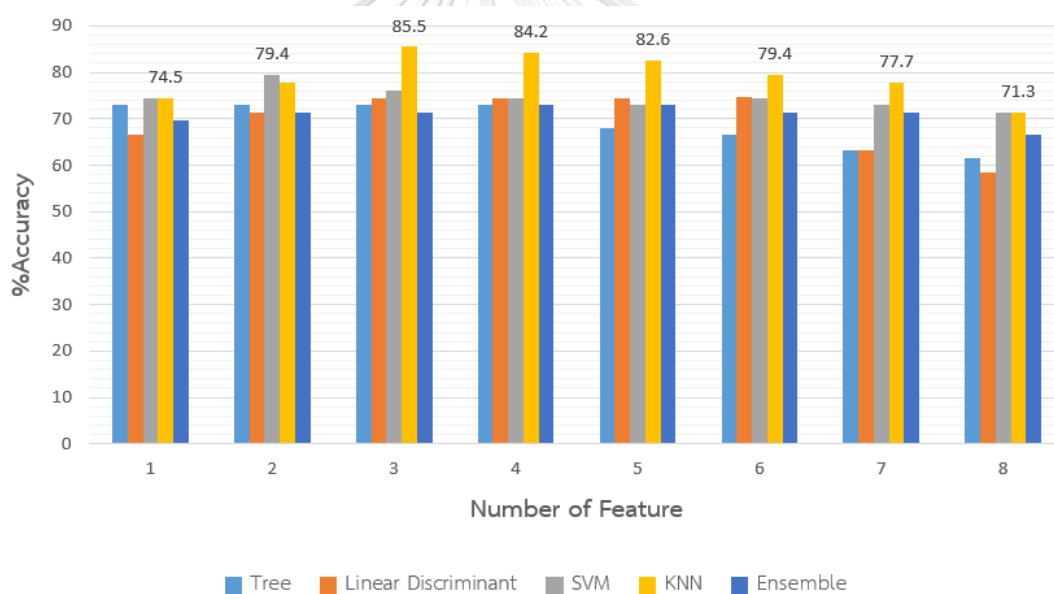






ตารางที่ 4-15 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มสัตว์ด้วยวิธี Ensemble

N	ประเภทของคุณลักษณะเด่นที่ถูกเลือก								เครื่องมือที่ใช้จำแนก
	di/dt	In	Vn	dv/dt	reclose	type	change	Vdip	Ensemble
1								✓	69.7
2					✓			✓	71.3
3					✓	✓		✓	71.3
4		✓			✓	✓		✓	72.9
5		✓		✓	✓	✓		✓	72.9
6		✓	✓	✓	✓	✓		✓	71.3
7		✓	✓	✓	✓	✓	✓	✓	71.3
8	✓	✓	✓	✓	✓	✓	✓	✓	66.5



รูปที่ 4-3 ค่าความแม่นยำในการจำแนกภายในกลุ่มสัตว์ของแต่ละอัลกอริทึม

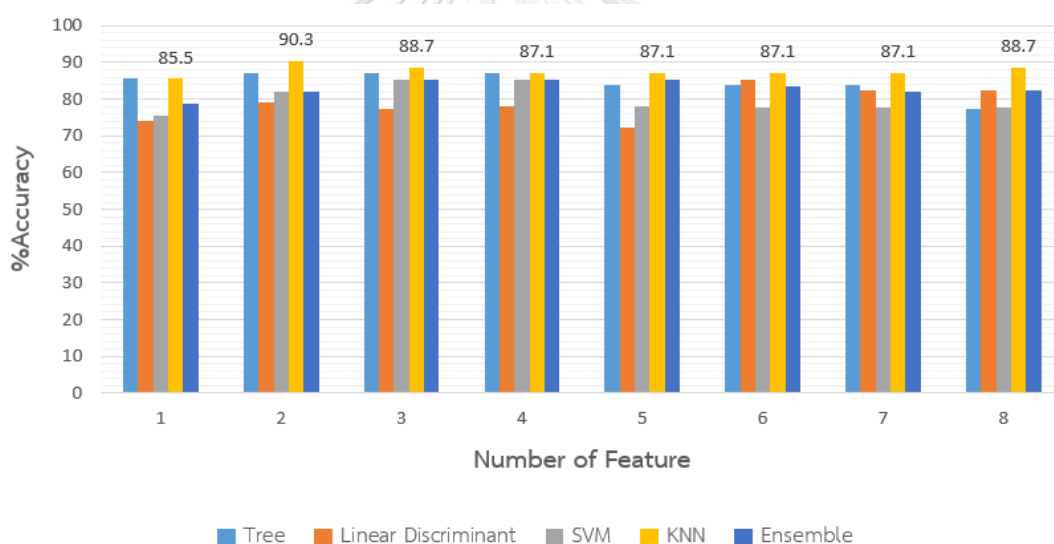
จากตารางที่ 4-14 และ รูปที่ 4-3 ภายใต้เงื่อนไขของคุณลักษณะเด่นและอัลกอริทึมที่ใช้ในการฝึกฝนต่างกัน พบว่าวิธี KNN ให้ความแม่นยำในการจำแนกสาเหตุความผิดปกติภายในกลุ่มของสัตว์สูงสุดที่ 85.5% ด้วย 3 คุณลักษณะเด่นคือ 1. ค่าสูงสุดของกระแสไฟฟ้า (In) 2. อัตราการเปลี่ยนแปลงของแรงดัน (dv/dt) 3. ค่าแรงดันตกชั่วขณะ (Vdip) โดยมีค่า K=10





ตารางที่ 4-20 การคัดเลือกคุณลักษณะเด่นของการจำแนกภายในกลุ่มต้นไม้ด้วยวิธี Ensemble

N	ประเภทของคุณลักษณะเด่นที่ถูกเลือก								เครื่องมือที่ใช้จำแนก
	di/dt	In	Vn	dv/dt	reclose	type	change	Vdip	Ensemble
1	✓								78.7
2	✓			✓					81.9
3	✓			✓		✓			85.2
4	✓	✓		✓		✓			85.2
5	✓	✓		✓	✓	✓			85.2
6	✓	✓		✓	✓	✓	✓		83.5
7	✓	✓	✓	✓	✓	✓	✓		81.9
8	✓	✓	✓	✓	✓	✓	✓	✓	82.3



รูปที่ 4-4 ค่าความแม่นยำในการจำแนกภายในกลุ่มต้นไม้ของแต่ละอัลกอริทึม

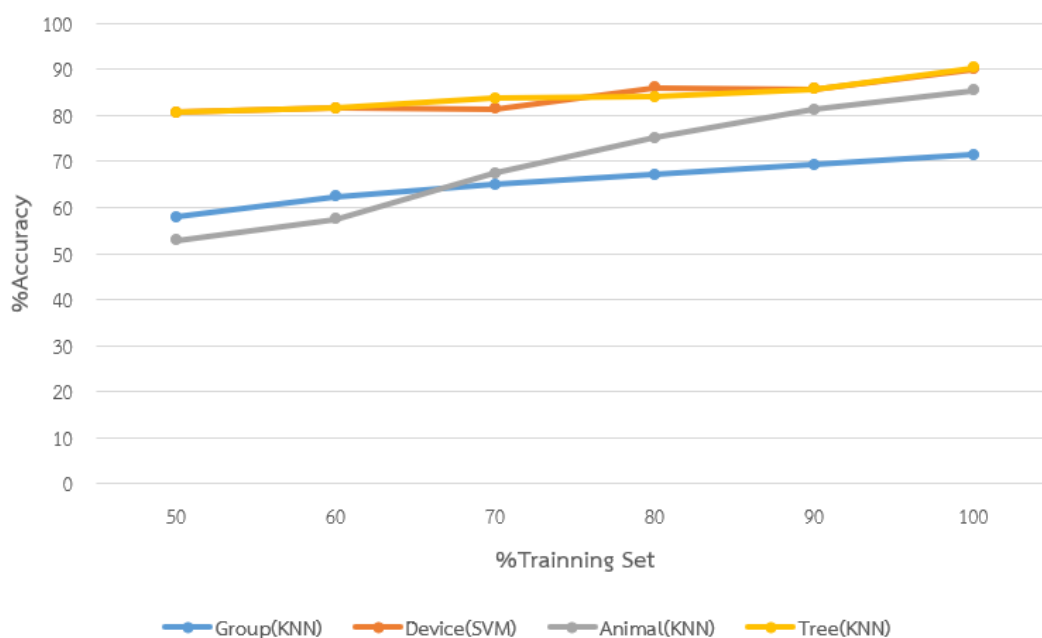
จากตารางที่ 4-19 และ รูปที่ 4-4 ภายใต้เงื่อนไขของคุณลักษณะเด่นและอัลกอริทึมที่ใช้ในการฝึกฝนต่างกัน พบว่าวิธี KNN ให้ความแม่นยำในการจำแนกสาเหตุความผิดปกติภายในกลุ่มของต้นไม้สูงสุดที่ 90.30% ด้วย 2 คุณลักษณะเด่นคือ 1. อัตราการเปลี่ยนแปลงของกระแส (In) 2. ชนิดของความผิดปกติ (type) โดยมีค่า K=10

#### 4.1.2 ผลของจำนวนข้อมูลฝึกฝนกับความแม่นยำ

ในการทดสอบนี้จะทำการทดสอบที่เปลี่ยนค่าของชุดข้อมูลฝึกฝนที่มีค่าตั้งแต่ 50-100% ของจำนวนข้อมูลทั้งหมด ทดสอบการจำแนกด้วยวิธีที่ให้ความแม่นยำสูงสุดของแต่ละโมเดล จะนำมาหาค่าความแม่นยำในการจำแนกของแต่ละการทดสอบ ผลการทดสอบดังแสดงในตารางที่ 4-21

ตารางที่ 4-21 ผลความแม่นยำของแต่ละจำนวนเหตุการณ์ที่ใช้จำแนกในแต่ละโมเดล

จำนวนของข้อมูลฝึกฝน (%)	ผลการจำแนกของแต่ละโมเดล			
	โมเดลจำแนกเป็นกลุ่ม (KNN)	โมเดลจำแนกในกลุ่มอุปกรณ์ (SVM)	โมเดลจำแนกในกลุ่มสัตว์ (KNN)	โมเดลจำแนกในกลุ่มต้นไม้ (KNN)
50	57.9	80.6	52.9	80.6
60	62.4	81.6	57.5	81.6
70	65	81.4	67.4	83.7
80	67.1	86	75.2	84
90	69.3	85.7	81.3	85.7
100	71.5	90	85.5	90.3



รูปที่ 4-5 ค่าความแม่นยำในการจำแนกของแต่ละจำนวนข้อมูลฝึกฝนในแต่ละโมเดล

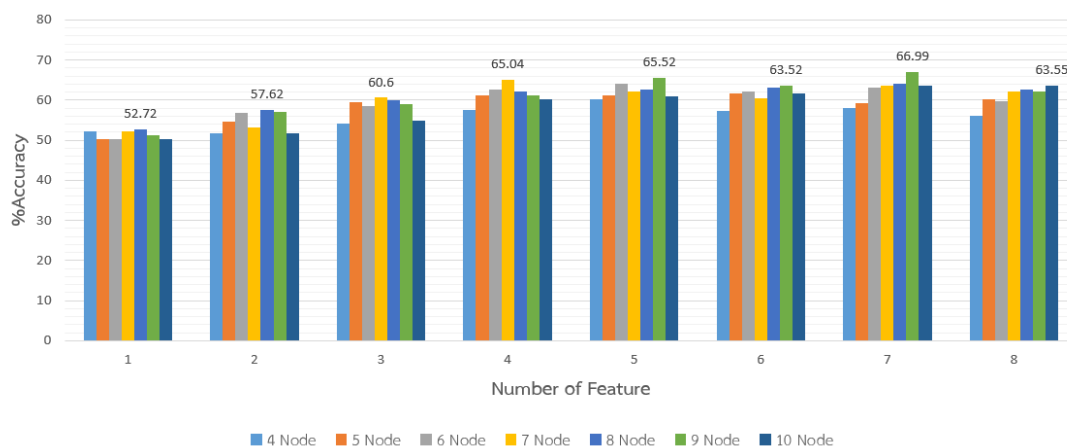
จากตารางที่ 4-21 และ รูปที่ 4-5 ภายใต้เงื่อนไขของจำนวนข้อมูลที่ใช้ในการฝึกฝนที่แตกต่างกัน พบว่าจำนวนของข้อมูลฝึกฝนมีผลต่อความแม่นยำในการจำแนก โดยร้อยละของข้อมูล











รูปที่ 4-6 ค่าความแม่นยำในการจำแนกเป็นกลุ่มของแต่ละจำนวนโนดซ่อน

จากตารางที่ 4-27 และ รูปที่ 4-6 ภายใต้เงื่อนไขของคุณลักษณะเด่นและจำนวนโนดซ่อนที่ใช้ในการฝึกฝนต่างกัน พบว่าที่จำนวน 9 โนดซ่อน ให้ความแม่นยำในการจำแนกสาเหตุความผิดปกติออกเป็นกลุ่มๆ สูงสุดที่ 66.99% ด้วย 7 คุณลักษณะเด่นคือ 1. อัตราการเปลี่ยนแปลงของกระแส ( $di/dt$ ) 2. ค่าสูงสุดของแรงดันนิวทรัล ( $V_n$ ) 3. อัตราการเปลี่ยนแปลงของแรงดัน ( $dv/dt$ ) 4. ความผิดปกติแบบชั่วคราวหรือถาวร (reclose) 5. ชนิดของความผิดปกติ (type) 6. การแปรเปลี่ยนชนิดของความผิดปกติ (change) 7. ค่าแรงดันตกชั่วขณะ ( $V_{dip}$ )

- 2) การคัดเลือกคุณลักษณะเด่นที่จำนวนโนดซ่อนที่ต่างกันของการจำแนกภายในกลุ่มของอุปกรณ์

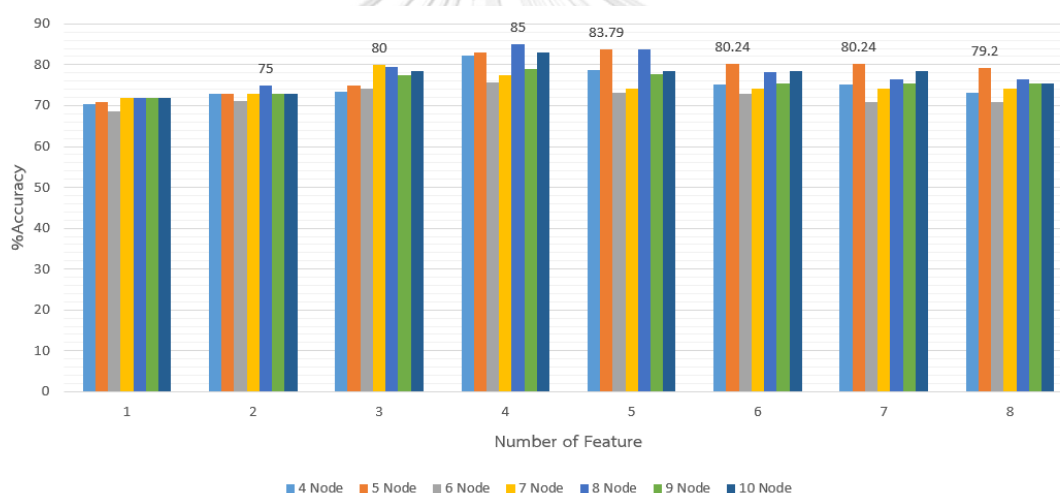
การทดสอบนี้จะคัดเลือกคุณลักษณะเด่นที่ใช้สำหรับจำแนกสาเหตุความผิดปกติภายในกลุ่มของอุปกรณ์ที่มีความแม่นยำสูงที่สุด จากจำนวนข้อมูลฝึกฝนที่เกิดจากอุปกรณ์ชำรุดทั้งหมด 60 เหตุการณ์ ที่จำนวนโนดซ่อนตั้งแต่ 4-10 โนดซ่อน ทดสอบความแม่นยำด้วยวิธี 5-fold cross validation ผลการทดสอบดังแสดงในตารางที่ 4-29 ถึง ตารางที่ 4-35





ตารางที่ 4-35 การคัดเลือกคุณลักษณะเด่นที่ 10 โหนดซ้อน ของการจำแนกภายในกลุ่มอุปกรณ์

N	ประเภทของคุณลักษณะเด่นที่ถูกเลือก								เครื่องมือที่ใช้จำแนก 10 โหนดซ้อน
	di/dt	In	Vn	dv/dt	reclose	type	change	Vdip	
1								✓	71.79
2							✓	✓	72.79
3				✓			✓	✓	78.43
4			✓	✓			✓	✓	83
5			✓	✓	✓		✓	✓	78.43
6		✓	✓	✓	✓		✓	✓	78.43
7	✓	✓	✓	✓	✓		✓	✓	78.43
8	✓	✓	✓	✓	✓	✓	✓	✓	75.34



รูปที่ 4-7 ค่าความแม่นยำในการจำแนกภายในกลุ่มอุปกรณ์ ของแต่ละจำนวนโหนดซ้อน

จากตารางที่ 4-33 และ ภาพที่ 4-7 ภายใต้งานไขของคุณลักษณะเด่นและจำนวนโหนดซ้อนที่ใช้ในการฝึกฝนต่างกัน พบว่าที่จำนวน 8 โหนดซ้อน ให้ความแม่นยำในการจำแนกสาเหตุความผิดปกติภายในกลุ่มอุปกรณ์ สูงสุดที่ 85% ด้วย 4 คุณลักษณะเด่นคือ 1. ค่าสูงสุดของแรงดันนิวทรัล (Vn) 2. อัตราการเปลี่ยนแปลงของแรงดัน (dv/dt) 3. การแปรเปลี่ยนชนิดของความผิดปกติ (change) 4. ค่าแรงดันตกชั่วขณะ (Vdip)





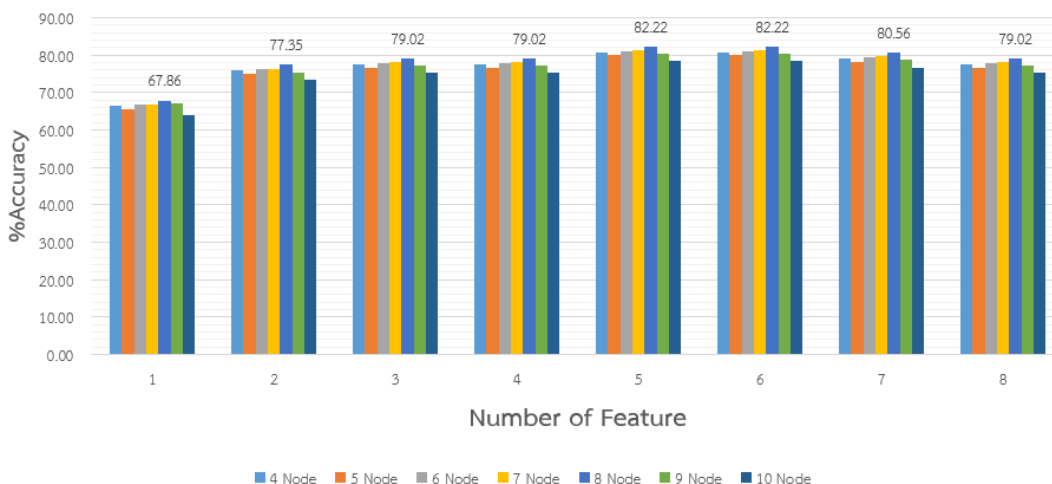












รูปที่ 4-9 ค่าความแม่นยำในการจำแนกภายในกลุ่มของต้นไม้ ของแต่ละจำนวนโนดซ่อน

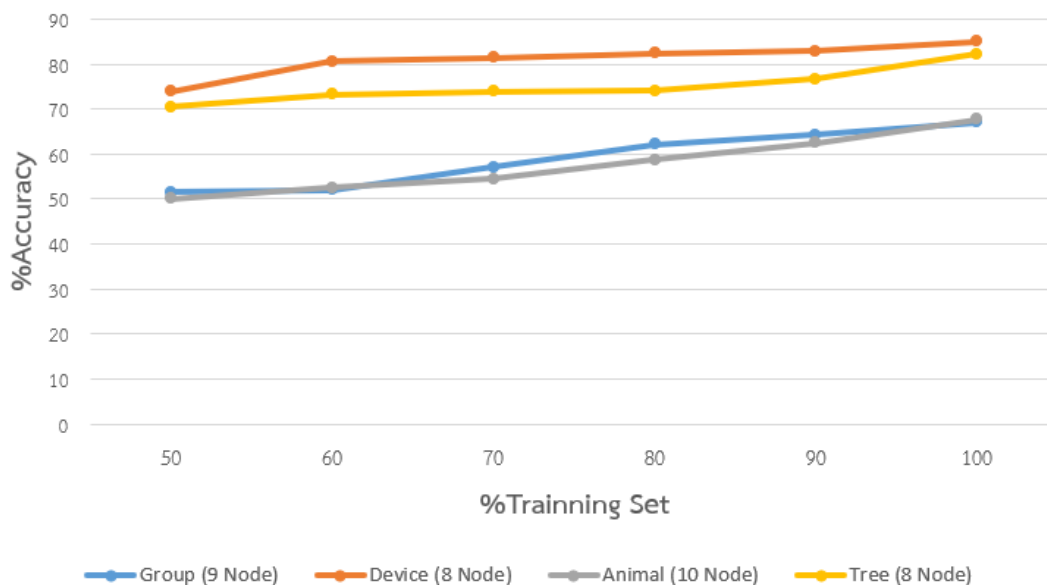
จากตารางที่ 4-47 และ รูปที่ 4-9 ภายใต้เงื่อนไขของคุณลักษณะเด่นและจำนวนโนดซ่อนที่ใช้ในการฝึกฝนต่างกัน พบว่าที่จำนวน 8 โนดซ่อน ให้ความแม่นยำในการจำแนกสาเหตุความผิดปกติภายในกลุ่มอุปกรณ์ สูงสุดที่ 82.22% ด้วย 5 คุณลักษณะเด่นคือ 1. ค่าสูงสุดของกระแสไนวรัล (In) 2. อัตราการเปลี่ยนแปลงของแรงดัน (dv/dt) 3. ความผิดปกติแบบชั่วคราวหรือถาวร (reclose) 4. การแปรเปลี่ยนชนิดของความผิดปกติ (change) 5. ค่าแรงดันตกชั่วขณะ (Vdip)

#### 4.2.2 ผลของจำนวนข้อมูลฝึกฝนกับความแม่นยำ

ในการทดสอบนี้จะทำการทดสอบที่เปลี่ยนค่าของชุดข้อมูลฝึกฝนที่มีค่าตั้งแต่ 50-100% ของจำนวนข้อมูลทั้งหมด ทดสอบการจำแนกด้วย Artificial Neural Network ที่โนดซ่อนที่ให้ความแม่นยำสูงสุดของแต่ละโมเดล จะนำมาหาค่าความแม่นยำในการจำแนกของแต่ละการทดสอบ ผลการทดสอบดังแสดงในตารางที่ 4-50

ตารางที่ 4-50 ผลความแม่นยำของแต่ละจำนวนเหตุการณ์ที่ใช้จำแนกในแต่ละโมเดล

จำนวนของข้อมูลฝึกฝน (%)	ผลการจำแนกของแต่ละโมเดล			
	โมเดลจำแนกเป็นกลุ่ม (9 โนดซ่อน)	โมเดลจำแนกในกลุ่มอุปกรณ์ (8 โนดซ่อน)	โมเดลจำแนกในกลุ่มสัตว์ (10 โนดซ่อน)	โมเดลจำแนกในกลุ่มต้นไม้ (8 โนดซ่อน)
50	51.52	73.93	50	70.48
60	52	80.56	52.50	73.21
70	57.12	81.33	54.44	73.89
80	62.14	82.36	58.73	74
90	64.24	82.88	62.42	76.67
100	66.99	85	67.69	82.22



รูปที่ 4-10 ค่าความแม่นยำในการจำแนกของแต่ละจำนวนข้อมูลฝึกฝนในแต่ละโมเดล

จากตารางที่ 4-50 และ รูปที่ 4.10 ภายใต้เงื่อนไขของจำนวนข้อมูลที่ใช้ในการฝึกฝนที่แตกต่างกัน พบว่าจำนวนของข้อมูลฝึกฝนมีผลต่อความแม่นยำในการจำแนก โดยร้อยละของข้อมูลฝึกฝนทั้งหมดที่ทำให้เกิดความแม่นยำสูงสุดของทุกโมเดลที่ใช้ในการจำแนกคือ ร้อยละ 100 ของจำนวนเหตุการณ์ทั้งหมด

#### 4.3 ผลการคัดเลือกคุณลักษณะเด่นและโมเดลที่เหมาะสม

ผลการเลือกคุณลักษณะเด่นที่เหมาะสมโดยวิธี Sequential Forward Selection ที่ทำให้แต่ละโมเดลมีประสิทธิภาพที่ดีที่สุด และ อัลกอริทึมที่ให้ Machine Learning ให้ผลการจำแนกมีความแม่นยำสูงสุดแสดงดังตารางที่ 4-51

ผลการเลือกคุณลักษณะเด่นที่เหมาะสมโดยวิธี Sequential Forward Selection ที่ทำให้แต่ละโมเดลมีประสิทธิภาพที่ดีที่สุด และจำนวนโนดซ่อนที่ส่งผลให้ Artificial Neural Network ให้ผลการจำแนกมีความแม่นยำสูงสุดแสดงดังตารางที่ 4-52

ตารางที่ 4-51 ผลการเลือกคุณลักษณะเด่นและอัลกอริทึมที่เหมาะสมของ Machine Learning

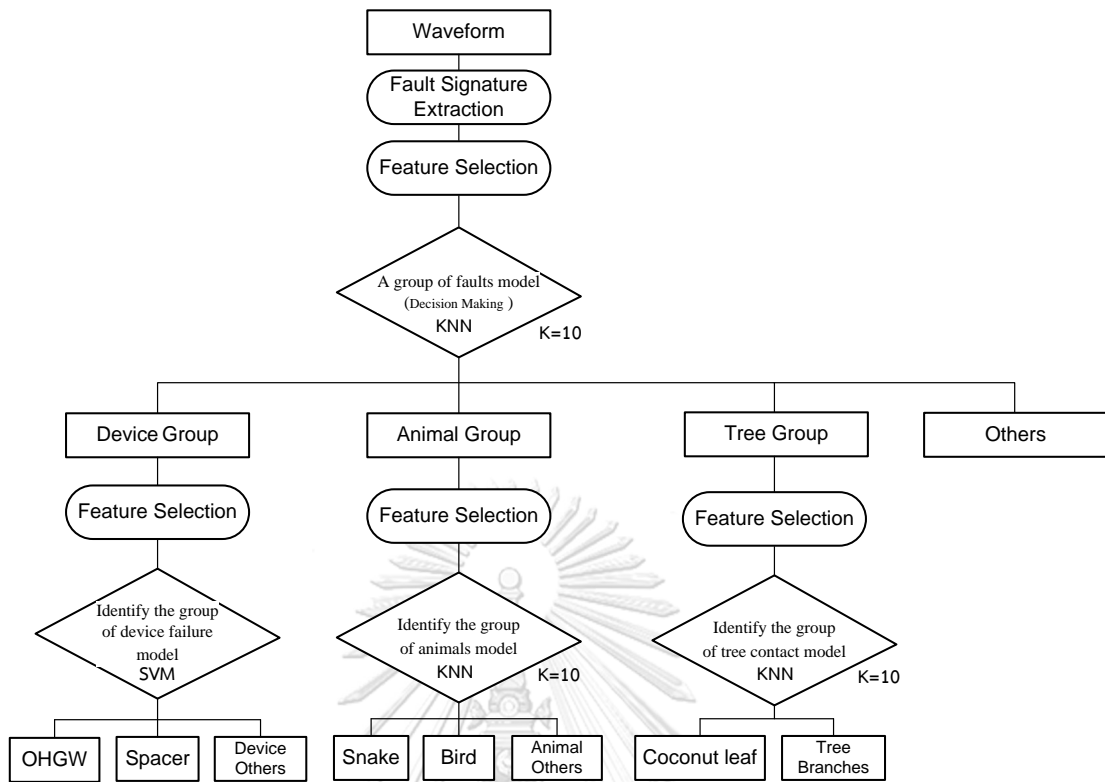
การจำแนกสาเหตุความผิดปกติ	คุณลักษณะเด่น	Training Model	
		Method	Result (%)
จำแนกออกเป็นกลุ่มๆ	di/dt, ln, Vn, dv/dt, reclose, type, change, Vdip	KNN	71.5
จำแนกภายในกลุ่มของอุปกรณ์	ln, dv/dt, change, Vdip	SVM	90
จำแนกภายในกลุ่มของสัตว์	ln, dv/dt, Vdip	KNN	85.5
จำแนกภายในกลุ่มของต้นไม้	di/dt, type	KNN	90.3

ตารางที่ 4-52 ผลการเลือกคุณลักษณะเด่นและจำนวนโนดซ่อนเหมาะสมของ ANN

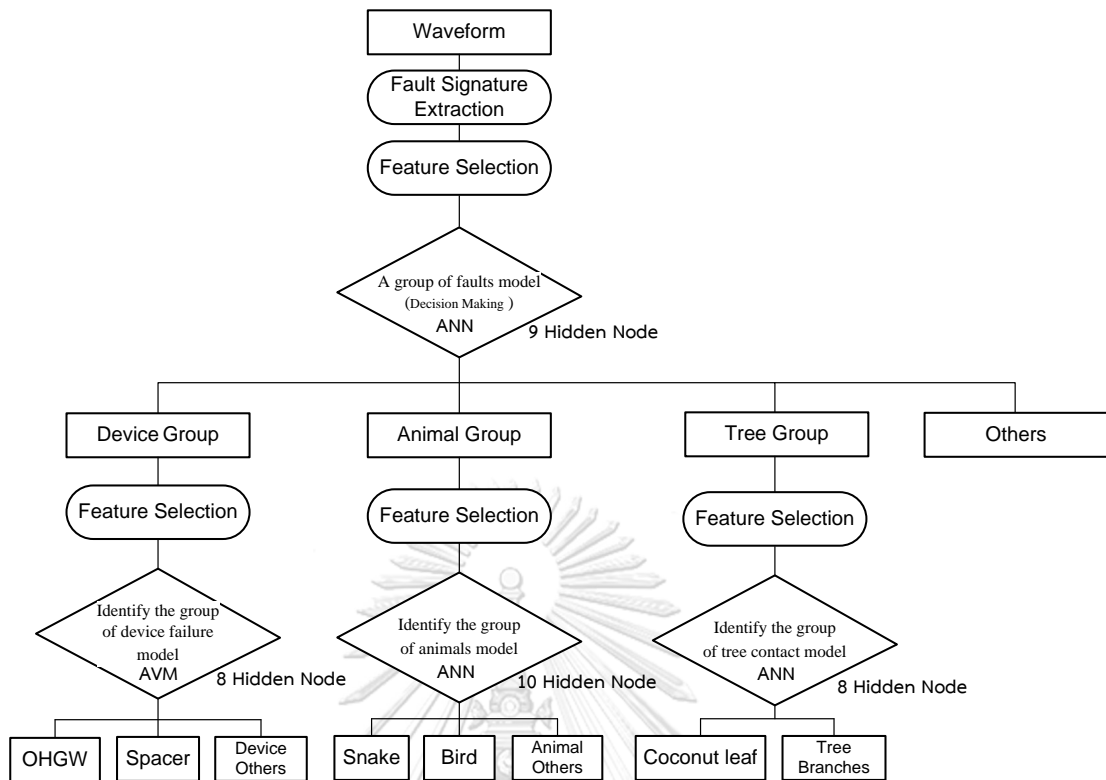
การจำแนกสาเหตุความผิดปกติ	คุณลักษณะเด่น	Training Model	
		จำนวนโนดซ่อน	Result (%)
จำแนกออกเป็นกลุ่มๆ	di/dt, Vn, dv/dt, reclose, type, change, Vdip	9	66.99
จำแนกภายในกลุ่มของอุปกรณ์	Vn, dv/dt, change, Vdip	8	85
จำแนกภายในกลุ่มของสัตว์	ln, dv/dt, reclose, change	10	67.69
จำแนกภายในกลุ่มของต้นไม้	ln, dv/dt, reclose, change, Vdip	8	82.22

#### 4.4 ผลการทดสอบกับเหตุการณ์จริง

ในการทดสอบกับเหตุการณ์จริงนี้เป็นการนำเหตุการณ์ความผิดปกติพร้อมใหม่อีก 77 เหตุการณ์ แบ่งเป็น ความผิดปกติที่เกิดจาก สัตว์ 25 เหตุการณ์ อุปกรณ์ชำรุด 23 เหตุการณ์ ต้นไม้ 18 เหตุการณ์ และอื่นๆ 11 เหตุการณ์ มาทดสอบกับแต่ละโมเดลที่ให้ประสิทธิภาพดีที่สุดที่สร้างจาก Machine Learning และ Artificial Neural Network โดยขั้นตอนการทดสอบดังแสดงในรูปที่ 4-11 และ 4-12



รูปที่ 4-11 ไดอะแกรมขั้นตอนการจำแนกด้วยโมเดลที่สร้างจาก Machine Learning



รูปที่ 4-12 ไดอะแกรมขั้นตอนการจำแนกด้วยโมเดลที่สร้างจาก Artificial Neural Network

#### 4.4.1 ผลการทดสอบกับเหตุการณ์จริงด้วยวิธี Machine Learning

ผลการทดสอบกับเหตุการณ์จริงด้วยวิธี Machine Learning แสดงดังตารางที่ 4-53 ถึง ตารางที่ 4-56

ตารางที่ 4-53 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกสาเหตุเป็นกลุ่มๆ ด้วยวิธี ML  
Accuracy = 72.72%

สาเหตุที่คาดคะเน (กลุ่ม)	สาเหตุจริง				Class Precision
	สัตว์	อุปกรณ์	ต้นไม้	อื่น ๆ	
สัตว์	17	1	0	0	94.44%
อุปกรณ์ชำรุด	6	18	0	5	62.07%
ต้นไม้	2	4	18	3	66.67%
อื่น ๆ	0	0	0	3	100%
Class Recall	68%	78.26%	100%	27.27%	



ตารางที่ 4-54 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี ML

Accuracy = 83.33%

สาเหตุที่คาดคะเน	สาเหตุจริง			Class Precision
	OHW	Spacer	อุปกรณ์อื่น ๆ	
OHW	6	0	0	100%
Spacer	1	3	0	75%
อุปกรณ์อื่น ๆ	0	2	6	75%
Class Recall	85.71%	60%	100%	

ตารางที่ 4-55 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มสัตว์ด้วยวิธี ML

Accuracy = 81.21%

สาเหตุที่คาดคะเน	สาเหตุจริง			Class Precision
	นก	งู	สัตว์อื่น ๆ	
นก	5	1	0	83.33%
งู	1	7	0	87.5%
สัตว์อื่น ๆ	0	1	2	66.67%
Class Recall	83.33%	77.78%	100%	

ตารางที่ 4-56 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มต้นไม้ด้วยวิธี ML

Accuracy = 88.89%

สาเหตุที่คาดคะเน	สาเหตุจริง		Class Precision
	ทางมะพร้าว	ต้นไม้อื่น ๆ	
ทางมะพร้าว	4	0	100%
ต้นไม้อื่น ๆ	2	12	85.71%
Class Recall	66.67%	100%	

ผลการทดสอบในตารางที่ 4-53 แสดงผลการจำแนกสาเหตุความผิดปกติเป็นกลุ่มๆ ด้วย Machine Learning พบว่าสามารถคาดคะเนสาเหตุของความผิดปกติว่ามาจากสัตว์ 17 เหตุการณ์ (17/25) มาจากอุปกรณ์ 18 เหตุการณ์ (18/23) มาจากต้นไม้ 18 เหตุการณ์ (18/18) และมาจากสาเหตุอื่นๆ 3 เหตุการณ์ (3/11) โดยผลการคาดคะเนทั้งหมดมีความแม่นยำ 72.72%

ผลการทดสอบในตารางที่ 4-54 แสดงผลการจำแนกสาเหตุความผิดปกติภายในกลุ่มอุปกรณ์ด้วย Machine Learning พบว่าสามารถคาดคะเนสาเหตุของความผิดปกติว่ามาจาก

OHGW 6 เหตุการณ์ (6/7) มาจาก Spacer 3 เหตุการณ์ (3/5) และมาจากอุปกรณ์อื่นๆ 6 เหตุการณ์ (6/6) โดยผลการคาดคะเนทั้งหมดมีความแม่นยำเท่ากับ 83.33%

ผลการทดสอบในตารางที่ 4-55 แสดงผลการจำแนกสาเหตุความผิดปกติภายในในกลุ่มสัตว์ ด้วย Machine Learning พบว่าสามารถคาดคะเนสาเหตุของความผิดปกติว่ามาจากนก 5 เหตุการณ์ (5/6) มาจากงู 7 เหตุการณ์ (7/9) และมาจากสัตว์อื่นๆ 2 เหตุการณ์ (2/2) โดยผลการคาดคะเนทั้งหมดมีความแม่นยำเท่ากับ 81.21%

ผลการทดสอบในตารางที่ 4-56 แสดงผลการจำแนกสาเหตุความผิดปกติภายในในกลุ่มต้นไม้ ด้วย Machine Learning พบว่าสามารถคาดคะเนสาเหตุของความผิดปกติว่ามาจากทางมะพร้าว 4 เหตุการณ์ (4/6) และมาจากต้นไม้อื่นๆ 12 เหตุการณ์ (12/12) โดยผลการคาดคะเนทั้งหมดมีความแม่นยำเท่ากับ 88.89%

#### 4.4.2 ผลการทดสอบกับเหตุการณ์จริงด้วยวิธี Artificial Neural Network

ผลการทดสอบกับเหตุการณ์ด้วยวิธี Artificial Neural Network แสดงดังตารางที่ 4-57 ถึง ตารางที่ 4-60

ตารางที่ 4-57 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกสาเหตุเป็นกลุ่มๆ ด้วยวิธี ANN  
Accuracy = 58.44%

สาเหตุที่คาดคะเน (กลุ่ม)	สาเหตุจริง				Class Precision
	สัตว์	อุปกรณ์	ต้นไม้	อื่น ๆ	
สัตว์	8	1	0	2	72.72%
อุปกรณ์ชำรุด	13	16	2	1	50%
ต้นไม้	4	6	16	3	55.17%
อื่น ๆ	0	0	0	5	100%
Class Recall	32%	69.56	88.89%	45.45%	

ตารางที่ 4-58 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มอุปกรณ์ด้วยวิธี ANN  
Accuracy = 75%

สาเหตุที่คาดคะเน	สาเหตุจริง			Class Precision
	OHGW	Spacer	อุปกรณ์อื่น ๆ	
OHGW	5	0	0	100%
Spacer	1	4	1	66.67%
อุปกรณ์อื่น ๆ	1	1	3	60%
Class Recall	71.43%	80%	75%	

ตารางที่ 4-59 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มสัตว์ด้วยวิธี ANN

Accuracy = 62.5%

สาเหตุที่คาดคะเน	สาเหตุจริง			Class Precision
	นก	งู	สัตว์อื่น ๆ	
นก	2	0	0	100%
งู	3	1	0	25%
สัตว์อื่น ๆ	0	0	2	100%
Class Recall	40%	100%	100%	

ตารางที่ 4-60 ผลการทดสอบกับสถานการณ์จริงด้วยโมเดลจำแนกภายในกลุ่มต้นไม้ด้วยวิธี ANN

Accuracy = 75%

สาเหตุที่คาดคะเน	สาเหตุจริง		Class Precision
	ทางมะพร้าว	ต้นไม้อื่น ๆ	
ทางมะพร้าว	2	0	100%
ต้นไม้อื่น ๆ	4	10	71.43%
Class Recall	33.33%	100%	

ผลการทดสอบในตารางที่ 4-57 แสดงผลการจำแนกสาเหตุความผิดปกติเป็นกลุ่มๆ ด้วย Artificial Neural Network พบว่าสามารถคาดคะเนสาเหตุของความผิดปกติว่ามาจากสัตว์ 8 เหตุการณ์ (8/25) มาจากอุปกรณ์ 16 เหตุการณ์ (16/23) มาจากต้นไม้ 16 เหตุการณ์ (16/18) และมาจากสาเหตุอื่นๆ 5 เหตุการณ์ (5/11) โดยผลการคาดคะเนทั้งหมดมีความแม่นยำ 58.44%

ผลการทดสอบในตารางที่ 4-58 แสดงผลการจำแนกสาเหตุความผิดปกติภายในกลุ่มอุปกรณ์ด้วย Artificial Neural Network พบว่าสามารถคาดคะเนสาเหตุของความผิดปกติว่ามาจาก OHGW 5 เหตุการณ์ (5/7) มาจาก Spacer 4 เหตุการณ์ (4/5) และมาจากอุปกรณ์อื่นๆ 3 เหตุการณ์ (3/4) โดยผลการคาดคะเนทั้งหมดมีความแม่นยำเท่ากับ 75%

ผลการทดสอบในตารางที่ 4-59 แสดงผลการจำแนกสาเหตุความผิดปกติภายในกลุ่มสัตว์ด้วย Artificial Neural Network พบว่าสามารถคาดคะเนสาเหตุของความผิดปกติว่ามาจากนก 2 เหตุการณ์ (2/5) มาจากงู 1 เหตุการณ์ (1/1) และมาจากสัตว์อื่นๆ 2 เหตุการณ์ (2/2) โดยผลการคาดคะเนทั้งหมดมีความแม่นยำเท่ากับ 62.5%

ผลการทดสอบในตารางที่ 4-60 แสดงผลการจำแนกสาเหตุความผิดปกติภายในกลุ่มต้นไม้ด้วย Artificial Neural Network พบว่าสามารถคาดคะเนสาเหตุของความผิดปกติว่ามาจากทางมะพร้าว 2 เหตุการณ์ (2/6) และมาจากต้นไม้อื่นๆ 10 เหตุการณ์ (10/10) โดยผลการคาดคะเนทั้งหมดมีความแม่นยำเท่ากับ 75%

#### 4.5 การวิเคราะห์ผล

จากผลการทดสอบทั้งหมดสามารถแบ่งการวิเคราะห์ที่ได้เป็น 3 ประเด็นหลักๆ คือ 1. ความแตกต่างของผลการจำแนกของแต่ละวิธี 2. ความแตกต่างของค่าความแม่นยำระหว่างการจำแนกเป็นกลุ่มๆ และการจำแนกภายในแต่ละกลุ่มของสาเหตุที่เกิดจากอุปกรณ์ สัตว์ และต้นไม้ 3. การปรับปรุงค่าความแม่นยำในการจำแนก

1. ในวิทยานิพนธ์เล่มนี้ใช้การเรียนรู้ของเครื่องแบบมีผู้สอน (Supervised Machine Learning) ในการจำแนกสาเหตุความผิดปกติ และนำไปเปรียบเทียบกับ การจำแนกโดยใช้ Artificial Neural Network ซึ่งหากดูกระบวนการของวิธี Artificial Neural Network ก็สามารถเปรียบได้ว่าเป็น Machine Learning ชนิดหนึ่ง เพียงแต่ Artificial Neural Network สามารถทำได้ทั้ง การเรียนรู้แบบมีผู้สอน (Supervised Learning) และการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Machine Learning) โดยจากผลการทดสอบในวิทยานิพนธ์เล่มนี้จากชุดข้อมูลฝึกฝนจำนวน 200 เหตุการณ์ ในทุกโมเดลทั้งการทดสอบโดยวิธี 5-fold cross validation และทดสอบโมเดลด้วยเหตุการณ์จริง พบว่า Supervised Machine Learning มีความแม่นยำมากกว่า Artificial Neural Network ในทุกการทดสอบ ซึ่งจะเห็นได้ว่าถึงแม้ว่า Artificial Neural Network จะมีขีดความสามารถในการใช้งานที่กว้าง แต่ไม่ได้หมายความว่า Artificial Neural Network จะให้ประสิทธิภาพที่ดีกว่าเสมอไป

2. จากผลการทดสอบด้วยวิธี Machine Learning พบว่าเมื่อนำมาทดสอบกับสถานการณ์จริง 77 เหตุการณ์ สามารถจำแนกกลุ่มของสาเหตุความผิดปกติได้ถูกต้องร้อยละ 72.72 สามารถจำแนกสาเหตุความผิดปกติในกลุ่มของอุปกรณ์ได้ถูกต้องร้อยละ 83.33 จำแนกสาเหตุความผิดปกติในกลุ่มของสัตว์ได้ถูกต้องร้อยละ 81.21 และสามารถจำแนกสาเหตุความผิดปกติในกลุ่มของต้นไม้ได้ถูกต้องร้อยละ 88.89 ซึ่งจะเห็นว่าระหว่างการจำแนกเป็นกลุ่มและการจำแนกสาเหตุภายในแต่ละกลุ่มมีค่าความแม่นยำแตกต่างกันมาก และจากตารางที่ 4-51 จะเห็นว่าค่าคุณลักษณะเด่นของโมเดลจำแนกเป็นกลุ่มๆ มีทั้งหมด 8 ค่า ซึ่งแตกต่างจากโมเดลที่ใช้จำแนกภายในกลุ่มของอุปกรณ์ สัตว์ และต้นไม้ ที่มีค่าคุณลักษณะเด่นเพียง 4 ค่า 3 ค่า และ 2 ค่า เมื่อวิเคราะห์ในมิติของเปอร์เซ็นต์ความถูกต้องและจำนวนของค่าคุณลักษณะเด่นจะเห็นว่ามีความแตกต่างกันของการจำแนกเป็นกลุ่ม และการจำแนกภายในแต่ละกลุ่ม ในส่วนของการจำแนกเป็นกลุ่มๆ ที่มีค่าความแม่นยำน้อยกว่าอาจมีผลมาจากปัญหา การเรียนรู้มากเกินไป (Overfitting) ที่เกิดจากค่าคุณลักษณะเด่นที่มากเกินไปหรืออาจเกิดจากค่าคุณลักษณะเด่นบางค่าของงานวิจัยนี้ยังไม่เหมาะสมเพียงพอต่อการจำแนกเป็นกลุ่ม จึงอาจส่งผลให้การจำแนกสาเหตุออกเป็นกลุ่มยังมีค่าน้อยกว่า และในส่วนของ การทดสอบด้วยวิธี Artificial Neural Network ก็มีลักษณะเช่นเดียวกัน

3. จากผลความแม่นยำของแต่ละจำนวนเหตุการณ์ที่แสดงดังรูปที่ 4-5 และ 4-10 ที่จำแนกด้วยวิธี Machine Learning และ Artificial Neural Network พบว่าการทดสอบความแม่นยำของแต่ละจำนวนเหตุการณ์ตั้งแต่ 50% ถึง 100% ของจำนวนเหตุการณ์ทั้งหมด ในทุกๆ การเพิ่มขึ้นของจำนวนเหตุการณ์จะทำให้ผลของความแม่นยำเพิ่มขึ้นตามไปด้วยและมีแนวโน้มว่าจะเพิ่มขึ้นอีกหากมีจำนวนเหตุการณ์เพิ่มขึ้น ซึ่งให้เห็นว่าจำนวนข้อมูลเหตุการณ์ที่นำมาใช้ในการเรียนรู้ของงานวิจัยนี้ยังไม่ถึงจุดที่โมเดลเหมาะสมที่สุด (Optimal) ซึ่งหากสามารถเพิ่มจำนวนเหตุการณ์อาจจะสามารถเพิ่มความแม่นยำในการจำแนกได้



## บทที่ 5 สรุปผล

### 5.1 สรุปผลการวิจัย

วิทยานิพนธ์ฉบับนี้กล่าวถึงการวิเคราะห์คุณลักษณะของรูปคลื่นกระแสและแรงดันขณะเกิดความผิดปกติ 8 คุณลักษณะประกอบด้วย 1. อัตราการเปลี่ยนแปลงของกระแส 2. ค่าสูงสุดของกระแสนิวทรัล 3. ค่าสูงสุดของแรงดันนิวทรัล 4. อัตราการเปลี่ยนแปลงของแรงดัน 5. ความผิดปกติแบบชั่วคราวหรือถาวร 6. ชนิดของความผิดปกติ 7. ความผิดปกติที่มีการแปรเปลี่ยนประเภท 8. ค่าแรงดันตกชั่วขณะ โดยข้อมูลความผิดปกติที่เกิดขึ้นในระบบจำหน่ายของการไฟฟ้าส่วนภูมิภาค เขต 2 ภาคใต้ ประกอบไปด้วยสาเหตุ สัตว์ อุปกรณ์ชำรุด และต้นไม้ อย่างละ 60 เหตุการณ์ และอื่นๆ 20 เหตุการณ์ รวม 200 เหตุการณ์ จาก 7 สถานีไฟฟ้า ประกอบด้วย สถานีไฟฟ้า เกาะสมุย 1 (KMA), นครศรีฯ 2 (NTB), ปากพนัง (PPN), กาญจนดิษฐ์ (KCD), อ่าวลึก (ALA), ท้ายเหมือง (TMG), และ ป่าตอง (PAT) ซึ่งแต่ละสาเหตุจะมีค่าลักษณะเด่นที่แตกต่างกันอย่างมีนัยสำคัญ

จากการวิเคราะห์คุณลักษณะของรูปคลื่นกระแสและแรงดันขณะเกิดความผิดปกติเพื่อจำแนกสาเหตุโดยใช้ Machine Learning พบว่า ถึงแม้ว่าการจำแนกเพื่อจัดว่าอยู่ในกลุ่มของสาเหตุใดยังมีเปอร์เซ็นต์ความแม่นยำที่ค่อนข้างน้อยที่ 72.72% แต่ถ้าสามารถจำแนกได้ถูกต้องว่าอยู่ในกลุ่มใดก็จะสามารถจำแนกเพิ่มเติมให้ละเอียดขึ้น โดยสามารถระบุได้ว่าเกิดจากอะไร ซึ่งมีเปอร์เซ็นต์ความแม่นยำอยู่ที่ 81.21% ถึง 88.89% ซึ่งในงานวิจัยก่อนหน้า [10, 28] ระบุเพียงสาเหตุความผิดปกติเป็นกลุ่ม ๆ เท่านั้น เช่นกลุ่มของสัตว์ ต้นไม้ ไม่ได้ระบุความผิดปกติเฉพาะให้ละเอียดขึ้น เช่น งู สเปเซอร์ หรือทางมะพร้าว และในส่วนของเปรียบเทียบระหว่าง Machine Learning กับ Artificial Neural Network เมื่อทดสอบกับข้อมูลฝึกฝนจำนวน 200 เหตุการณ์และเหตุการณ์จริง 77 เหตุการณ์ ที่เกิดขึ้นในระบบจำหน่ายของการไฟฟ้าส่วนภูมิภาค เขต 2 ภาคใต้ พบว่า Machine Learning ให้ความแม่นยำมากกว่า Artificial Neural Network เมื่อกลับมาดูในงานวิจัยก่อนหน้า [4] ที่สามารถจำแนกสาเหตุโดยใช้ Artificial Neural Network ได้ถูกต้องถึง 98% นั้น ใช้ข้อมูลเชิงกายภาพ จำนวน 1,337 เหตุการณ์ ซึ่งมากกว่าจำนวนชุดฝึกของวิทยานิพนธ์นี้มาก แต่ด้วยข้อจำกัดด้านข้อมูล เหตุการณ์ความผิดปกติที่ทราบสาเหตุชัดเจนที่สามารถนำมาใช้เป็นข้อมูลฝึกฝนในวิทยานิพนธ์นี้ได้ มีประมาณ 200 เหตุการณ์ จึงอาจสรุปในส่วนของ การนำ Artificial Neural Network มาประยุกต์ใช้ของวิทยานิพนธ์นี้ได้ว่า จำนวนชุดฝึก 200 เหตุการณ์อาจจะน้อยเกินไป และอาจเป็นสาเหตุหนึ่งที่ทำให้ Artificial Neural Network ไม่เหมาะที่จะนำมาใช้กับการจำแนกสาเหตุความผิดปกติ (Fault Cause Classification)

สำหรับเหตุการณ์ที่ยังจำแนกไม่ถูกต้องนั้น จากการวิเคราะห์ผลการทดสอบจำนวนข้อมูลฝึกฝนกับความแม่นยำชี้ให้เห็นว่ายังสามารถปรับปรุงความแม่นยำได้โดยเพิ่มจำนวนข้อมูลเหตุการณ์ที่

นำมาใช้วิเคราะห์ และอาจค้นหาหรือเพิ่มเติมคุณลักษณะเด่นที่มีความเหมาะสมมากกว่าที่วิทยานิพนธ์นี้นำมาใช้

## 5.2 ข้อเสนอแนะสำหรับงานวิจัยในอนาคต

1. ปรับปรุงการสกัดคุณลักษณะเด่นให้มีความเหมาะสมกับการจำแนกสาเหตุของความผิดปกติที่จะทำการศึกษาในอนาคตมากขึ้น เช่น ตำแหน่งมุมที่เกิดความผิดปกติ หรือ นำข้อมูลสภาวะภายนอกเช่น สภาพอากาศ ลักษณะพื้นที่ที่ระบบจำหน่ายพาดผ่าน (ป่า, เมือง) รูปแบบของการติดตั้งสายเคเบิลอากาศ (เคเบิลสเปเซอร์, ลูกถ้วยประกอบคอน, ลูกถ้วยประกอบเหล็กคอนเคเบิลทางโค้ง) ฯลฯ มาประกอบการวิเคราะห์ คาดว่าน่าจะมีแนวโน้มทำให้ความแม่นยำในการทำนายเพิ่มมากขึ้น เนื่องจาก บางสภาพอากาศ เช่น ลมแรงอาจโน้มกิ่งไม้หรือทางมะพร้าวแห้งมาสัมผัสแนวสาย, ลักษณะพื้นที่ เช่น ป่าอาจมีแนวโน้มเกิดจากสาเหตุจากต้นไม้ สัตว์ มากกว่า, รูปแบบของการติดตั้งสายเคเบิลอาจจะสามารถจำแนกสาเหตุให้แคบลงเช่น หากเป็นลูกถ้วยประกอบคอนอาจตัดสาเหตุที่เกิดจากสเปเซอร์ได้เป็นต้น
2. ต่อยอดและปรับปรุงโดยการทดสอบกับเครื่องมือจำแนกอื่น เช่น การเรียนรู้เชิงลึก (Deep Learning)

## บรรณานุกรม

1. ขจรศักดิ์ ทองพบ, *One Point Knowledge “Work Process of Outage Management System”*. 2561.
2. สุชาติ ปรีชาธร, วิศวกรรมกรรมการป้องกันระบบไฟฟ้าแรงสูง (เล่ม 1). 2555, กรุงเทพมหานคร: ซีเอ็ดยูเคชั่น.
3. การไฟฟ้าส่วนภูมิภาค, รายงานไฟฟ้าขัดข้อง พื้นที่ กพต.2 ปี 2560-2562 (ข้อมูลอุปกรณ์ป้องกันไลน์เมน Breaker และ Recloser). 2562.
4. Chow, M.-y., S.O. Yee, and L.S. Taylor, *Recognizing animal-caused faults in power distribution systems using artificial neural networks*. IEEE Transactions on Power Delivery, 1993. 8(3): p. 1268 - 1274.
5. Chow, M.-y., S.O. Yee, and L.S. Taylor, *Neural-fuzzy hybrid system for distribution fault causes identification*, in *Proceedings of the Second International Forum on Applications of Neural Networks to Power Systems*. 1993: Yokohama, Japan, Japan.
6. Xu, L. and M.-y. Chow, *A Classification Approach for Power Distribution Systems Fault Cause Identification*. IEEE Transactions on Power Systems, 2006. 21(1): p. 53-60.
7. Xu, L., M.-Y. Chow, and L.S. Taylor, *Power Distribution Fault Cause Identification With Imbalanced Data Using the Data Mining-Based Fuzzy Classification SES-Algorithm*. IEEE Transactions on Power Systems, 2007. 22(1): p. 164-171.
8. Zhang, H. and R. Che, *Fault cause identification based on characteristics of transition resistances for transmission lines*, in *2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*. 2015, IEEE: Changsha, China.
9. Kulkarni, S., et al., *Waveform characterization of animal contact, tree contact, and lightning induced faults*, in *IEEE PES General Meeting*. 2010: Providence, RI, USA.
10. รัฐพล หงษ์ดีลกุล, การจำแนกความผิดปกติพร้อมบนสายส่งด้วยลักษณะสมบัติของลูกคลื่น, in *วิศวกรรมไฟฟ้า*. 2559, จุฬาลงกรณ์มหาวิทยาลัย.



11. Ingeteam Power Technology S.A., *MULTIFUNCTION PROTECTION INGEPAC EF-MD User Manual*.
12. Xu, L., et al., *On the Investigation of Artificial Immune Systems on Imbalanced Data Classification for Power Distribution System Fault Cause Identification*, in *2006 IEEE International Conference on Evolutionary Computation*. 2006, IEEE: Vancouver, BC, Canada.
13. Zin, A.A.M. and S.P.A. Karim, *The Application of Fault Signature Analysis in Tenaga Nasional Berhad Malaysia*. IEEE Transactions on Power Delivery, 2007. 22(4): p. 2047-2056.
14. Núñez, V.B., et al., *Feature analysis and classification methodology for overhead distribution fault events.pdf*, in *IEEE PES General Meeting*. 2010, IEEE: Providence, RI, USA.
15. Qin, X., et al., *Research on Distribution Network Fault Recognition Method Based on Time-Frequency Characteristics of Fault Waveforms*. IEEE Access, 2018. 6: p. 7291-7300.
16. ปริญญา สงวนสัตย์, *Artificial Intelligence with Machine Learning*. 1 ed. 2562, นนทบุรี: บริษัท ไอดีซี พรีเมียม จำกัด.
17. โกเมศ อัมพวัน, *Data mining-การทำเหมืองข้อมูล*. 2562.
18. พลสัมพันธ์ พงษ์ประยูร, *การวิเคราะห์ซิกเนเจอร์ของความผิดปกติพ่วงในระบบส่งไฟฟ้าโดยใช้การรู้จำรูปแบบ*, in *วิศวกรรมไฟฟ้า*. 2554, จุฬาลงกรณ์มหาวิทยาลัย.
19. นพดล จตุไพบูลย์, *การพัฒนาเทคโนโลยีเพื่อการปฏิสัมพันธ์ระหว่างมนุษย์และคอมพิวเตอร์โดยใช้สัญญาณสมองร่วมกับสัญญาณอื่นๆหลายรูปแบบ*, in *วิศวกรรมคอมพิวเตอร์*. 2557, จุฬาลงกรณ์มหาวิทยาลัย.
20. Xu, L. and M.-y. Chow, *Power distribution systems fault cause identification using logistic regression and artificial neural network.pdf*, in *Proceedings of the 13th International Conference on, Intelligent Systems Application to Power Systems*. 2005, IEEE: Arlington, VA, USA.
21. Xu, L., M.-y. Chow, and L.S. Taylor, *Data Mining and Analysis of Tree-Caused Faults in Power Distribution Systems.pdf*, in *2006 IEEE PES Power Systems Conference and Exposition*. 2006, IEEE: Atlanta, GA, USA.
22. จิรายุทธ์ กิตติจันทรรัตน์, *การวิเคราะห์อุปกรณ์ผิดปกติพ่วงบนเครือข่ายระบบส่งไฟฟ้าโดยใช้ระบบผู้เชี่ยวชาญ*,

- in วิศวกรรมไฟฟ้า. 2553, จุฬาลงกรณ์ มหาวิทยาลัย.
23. Li, L., R. Che, and H. Zang, *A fault cause identification methodology for transmission lines based on support vector machines*, in 2016 *IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*. 2016, IEEE: Xi'an, China.
  24. Ma, C.Z.a.Y., *Ensemble Machine Learning Methods and Applications*. 2012: Springer.
  25. Kohavi, E.B.a.R., *An empirical comparison of voting classification algorithms: Bagging, boosting, and variants*. *Machine Learning* 36, 1999.
  26. Breiman, L., *Bagging Predictors*. 1994, University of California.
  27. อสมา กุลวานิชไชยน์นธ์, *Big Data Series II: Think Like a Data Scientist* คิดแบบ นักวิทยาศาสตร์ข้อมูล, ed. 2. 2562, สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมิกราช: บริษัท คอราไลน์ จำกัด (สำนักงานใหญ่).
  28. Hongdilokkul, R. and C. Banmongkol, *Classification of transmission line faults with waveform characterization*, in 2016 *13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. 2016, IEEE: Chiang Mai, Thailand.
  29. Brown, R.E., *Electric Power Distribution Reliability*, ed. n. Edition. 2008, Boca Raton.
  30. Bollen, M.H.J., et al., *Classification of Underlying Causes of Power Quality Disturbances: Deterministic versus Statistical Methods*. *EURASIP Journal on Advances in Signal Processing*, 2007. 2007(1).



จุฬาลงกรณ์มหาวิทยาลัย  
**CHULALONGKORN UNIVERSITY**

## ประวัติผู้เขียน

ชื่อ-สกุล	นายวิษณุ พรหมรัตน์
วัน เดือน ปี เกิด	6 ธันวาคม 2531
สถานที่เกิด	จังหวัดนครศรีธรรมราช
วุฒิการศึกษา	วิศวกรรมศาสตรบัณฑิต คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย วิทยาเขต สงขลา
ที่อยู่ปัจจุบัน	140/57 ม.4 ต.จันดี อ.ฉวาง จ.นครศรีธรรมราช 80250
ผลงานตีพิมพ์	วิษณุ พรหมรัตน์, วาทีต เบญจพลกุล “การจำแนกสาเหตุความผิดปกติในระบบจำหน่าย 33 kV ของการไฟฟ้าส่วนภูมิภาค เขต 2 ภาคใต้ โดยใ้การใช้การเรียนรู้ของเครื่อง” การประชุมเครือข่ายวิศวกรรมไฟฟ้า ครั้งที่ 12, EENET, 2563