

หน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางกับกลไกจุดสนใจสำหรับการทำนายความล้มเหลว
ของรีโมทคอมเพรสเซอร์โดยใช้การสกัดลักษณะสำคัญร่วมกับเทคนิคการลดข้อมูล



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2562
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Attention-Based Bidirectional Long Short-Term Memory for Remote Compressor
Failure Prediction Using Feature Extraction Together with Data Reduction Techniques



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science
Department of Computer Engineering
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2019
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

หน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางกับกลไกจุดสนใจ
สำหรับการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์โดยใช้การ
สกัดลักษณะสำคัญร่วมกับเทคนิคการลดข้อมูล

โดย

นายวิระศักดิ์ ชมภู

สาขาวิชา

วิทยาศาสตร์คอมพิวเตอร์

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์

..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุกรี สิ้นธุภิญโญ)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.พีรพล เวทีกุล)

..... กรรมการภายนอกมหาวิทยาลัย
(รองศาสตราจารย์ ดร.ชลวิช นัทธี)

วิระศักดิ์ ชมภู : หน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางกับกลไกจุดสนใจสำหรับการทำนาย
ความล้มเหลวของรีโมทคอมเพรสเซอร์โดยใช้การสกัดลักษณะสำคัญร่วมกับเทคนิคการลดข้อมูล. (
Attention-Based Bidirectional Long Short-Term Memory for Remote Compressor
Failure Prediction Using Feature Extraction Together with Data Reduction Techniques)
อ.ที่ปรึกษาหลัก : ศ. ดร.บุญเสริม กิจศิริกุล

ในอุตสาหกรรมน้ำมันและก๊าซธรรมชาติ รีโมทคอมเพรสเซอร์ถือเป็นอุปกรณ์สำคัญ ซึ่งความล้มเหลว
ของอุปกรณ์ที่มีความรุนแรงระดับสูงสามารถทำให้เกิดความเสียหายอย่างใหญ่หลวงต่อมนุษย์และสิ่งแวดล้อมได้
วิศวกรต้องใช้ข้อมูลจากเครื่องมือวัดหลายจุดเพื่อวิเคราะห์ด้วยวิธีการทางสถิติในการวางแผนการบำรุงรักษาและ
การดำเนินงาน แต่เมื่อข้อมูลมีปริมาณมหาศาลจึงเป็นความท้าทายอย่างยิ่งในการวิเคราะห์โดยมนุษย์ ดังนั้นการ
ทำนายความล้มเหลวล่วงหน้าของรีโมทคอมเพรสเซอร์จึงมีความสำคัญอย่างยิ่ง วิทยานิพนธ์นี้
นำเสนอ หน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางกับกลไกจุดสนใจ ซึ่งมีความสามารถในการเรียนรู้ที่มี
การอ้างอิงระยะยาวของข้อมูลอนุกรมเวลาและกลไกจุดสนใจที่ช่วยเพิ่มช่วยให้แบบจำลองสามารถเลือกลำดับ
ของเอาต์พุตที่เหมาะสม และเสริมประสิทธิภาพโดยใช้โครงข่ายคอนโวลูชันในการสกัดคุณลักษณะสำคัญอย่าง
อัตโนมัติจากคุณลักษณะท้องถิ่นที่อิสระต่อเวลา เพื่อเสริมประสิทธิภาพความครอบคลุมแก่แบบจำลอง
วิทยานิพนธ์นี้ยังได้นำเสนอเทคนิคการลดข้อมูล เพื่อปรับปรุงประสิทธิภาพโดยแสดงประสิทธิภาพของกระบวนการฝึก
ด้วยชุดข้อมูลขนาดใหญ่ ผู้วิจัยประเมินประสิทธิภาพของวิธีการที่นำเสนอโดยใช้ชุดข้อมูลจริงของรีโมท
คอมเพรสเซอร์ เปรียบเทียบคะแนน F1 กับแบบจำลองการเรียนรู้ของเครื่องและแบบจำลองโครงข่ายประสาท
เทียมหลากหลายสถาปัตยกรรม ผลการทดลองชี้ให้เห็นว่าวิธีการที่นำเสนอนี้มีประสิทธิภาพการทำนายความ
ล้มเหลวของรีโมทคอมเพรสเซอร์ที่เหนือกว่าวิธีอื่น

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์
ปีการศึกษา 2562

ลายมือชื่อนิสิต
ลายมือชื่อ อ.ที่ปรึกษาหลัก

6170964221 : MAJOR COMPUTER SCIENCE

KEYWORD: Machine Learning, Deep Learning, Long Short-Term Memory, Attention Mechanism, Failure Prediction, Remote Compressor, Convolutional Neural Networks

Wirasak Chomphu : Attention-Based Bidirectional Long Short-Term Memory for Remote Compressor Failure Prediction Using Feature Extraction Together with Data Reduction Techniques. Advisor: Prof. Dr. BOONSERM KIJSIRIKUL, D.Eng.

In the oil and gas industry, Remote Compressors (RCs) are considered critical equipment. Their high severity of accidental failures can cause catastrophic damage to human beings and the environment. Engineers must capture multiple sensor measurements for analysis using statistical methods to schedule maintenance activities or operating plans. When dealing with a large amount of data, it becomes exceptionally challenging for interpretation by human efforts. Thus, the RC failure prediction is approaching criticality. This thesis proposes Attention-based Bidirectional Long Short-Term Memory (ABD-LSTM), capable of effectively learning long-term dependencies of time series data and attention mechanism dynamically selecting relevant sequence outputs for capable of predicting the accidental failures. The model enhanced by Convolutional Neural Networks (CNNs), which can automatically extract essential information from local features independent of time to enhance the model performance. This thesis also proposes feature extraction and data reduction techniques as complementary methods to improve the effectiveness of the training process in a large-scale dataset. We evaluate the proposed method performance using F1 score on actual RC datasets through comparisons with other classical machine learning models and several neural network architectures. The results indicate that the proposed method achieves superior prediction performance and outperforms all its counterparts.

Field of Study: Computer Science

Student's Signature

Academic Year: 2019

Advisor's Signature

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์เป็นอย่างยิ่งจาก ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล ซึ่งเป็นอาจารย์ที่ปรึกษา ซึ่งท่านได้ให้ความรู้ คำปรึกษา คำแนะนำ แนวทางในการจัดทำวิทยานิพนธ์และข้อคิดเห็นอันเป็นประโยชน์อย่างยิ่งในการจัดทำวิทยานิพนธ์ ตลอดจนการตรวจสอบและแนะนำแก้ไขข้อบกพร่องจนทำให้วิทยานิพนธ์นี้สำเร็จลุล่วงด้วยดี

ขอขอบพระคุณอาจารย์กรรมการสอบวิทยานิพนธ์ ซึ่งประกอบด้วย รองศาสตราจารย์ ดร.ชลวิช นัทธี ผู้ช่วยศาสตราจารย์ ดร.สุกรี สิ้นธุภิญโญ และผู้ช่วยศาสตราจารย์ ดร.พีรพล เวทีกุล ที่สละเวลามาเป็นกรรมการสอบวิทยานิพนธ์ อีกทั้งยังให้คำแนะนำในการทำวิทยานิพนธ์ให้สำเร็จลุล่วงไปด้วยดี

ขอขอบคุณ บริษัท เซฟรอนประเทศไทยสำรวจและผลิต จำกัด ที่สนับสนุนทุนการศึกษา ข้อมูล เครื่องมือคอมพิวเตอร์ประมวลผลงานวิจัย ผู้เชี่ยวชาญในการตรวจสอบข้อมูลและคำแนะนำเกี่ยวกับข้อมูล การสนับสนุนค่าใช้จ่ายในการเดินทางไปเผยแพร่รายงานวิจัยต่างประเทศ การสนับสนุนจากผู้บริหาร หัวหน้างานและกำลังใจจากเพื่อนร่วมงาน

ขอบคุณนายธนัทวิทย์ อังศรวณีย์ นางสาวรมณ รอบโลกและนางสาวธัญญา พิรพัฒนาการ สมาชิกห้องปฏิบัติการภายใต้ที่ปรึกษาของ ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล และสมาชิกห้องปฏิบัติการอื่น ๆ ที่ได้ให้การช่วยเหลือและให้กำลังใจมาเป็นอย่างดี

สุดท้ายนี้ขอกราบขอบคุณพ่อชัยสวัสดิ์ ชมภู ขอขอบคุณพี่น้องและญาติสนิททุกท่านที่ให้การสนับสนุนและเป็นที่กำลังใจที่ทำให้ประสบความสำเร็จลุล่วงไปด้วยดี

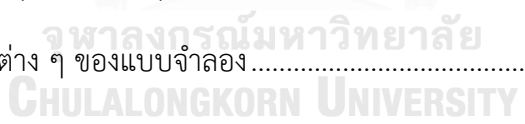
วิระศักดิ์ ชมภู

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญ

	หน้า
.....	ค
บทคัดย่อภาษาไทย.....	ค
.....	ง
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ฌ
สารบัญรูปภาพ.....	ฎ
บทที่ 1 ที่มาและความสำคัญ.....	1
1.1 ที่มาและความสำคัญ.....	1
1.2 วัตถุประสงค์.....	4
1.3 ขอบเขตการดำเนินงาน.....	4
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	4
1.5 ขั้นตอนการดำเนินงาน.....	5
1.6 สรุปผลภาพรวมของการทำวิจัย.....	5
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 ทฤษฎีที่เกี่ยวข้อง.....	7
2.2 งานวิจัยที่เกี่ยวข้อง.....	23
บทที่ 3 แบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าและเทคนิคการเพิ่มประสิทธิภาพโดยการสกัดลักษณะสำคัญร่วมกับเทคนิคการลดข้อมูล.....	27
3.1 แบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้า.....	27

3.2 การเพิ่มประสิทธิภาพให้กับแบบจำลอง	28
3.4 การแปลงเมทริกซ์ลำดับเวลาสามมิติ (Time Sequence Matrix 3D Transformation - TSMT).	34
3.5 เทคนิคการลดข้อมูล	35
การแปลงเมทริกซ์ลำดับเวลาสามมิติ	36
3.6 การวิเคราะห์องค์ประกอบสำคัญ.....	39
3.6 การเลือกคุณลักษณะ.....	40
บทที่ 4 แนวทางและวิธีการดำเนินงาน	42
4.1 แหล่งข้อมูล	42
4.2 การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูลเชิงตรวจสอบ	43
4.3 การประมวลผลข้อมูลก่อนหน้า (Data Pre-processing).....	43
4.4 การทดลอง	47
4.5 การแบ่งข้อมูลในการทดลอง.....	49
.....	49
4.6 การฝึก (Training).....	50
4.7 การประเมินผล (Evaluation)	54
4.8 การกำหนดค่าต่าง ๆ ของแบบจำลอง.....	57
บทที่ 5 ผลการดำเนินงาน	59
5.1 ผลการใช้เทคนิคลดข้อมูล.....	59
5.2 การทดลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองและวิธีการเลือกคุณลักษณะ.....	60
5.3 การเปรียบเทียบประสิทธิภาพของแบบจำลองและการเลือกแบบจำลอง.....	64
บทที่ 6 สรุปการดำเนินงาน	67
6.1 สรุปการวิจัย	67
6.2 ข้อจำกัดของงานวิจัย	68
6.3 แนวทางในการวิจัยในอนาคต	68



ภาคผนวก.....	70
ภาคผนวก ก คุณลักษณะที่เลือกนำมาใช้ในกระบวนการทดลอง.....	71
ภาคผนวก ข ตารางค่าไอเกินที่ได้จากการวิเคราะห์ห้องค์ประกอบสำคัญ.....	80
ภาคผนวก ค ตารางแสดงรายละเอียดของคุณลักษณะที่ใช้ในการเลือกคุณลักษณะแต่ละการ ทดลอง	85
บรรณานุกรม.....	93
ประวัติผู้เขียน.....	96



สารบัญตาราง

	หน้า
ตารางที่ 1 แผนการดำเนินงาน	6
ตารางที่ 2 ประเภทของฟังก์ชันกระตุ้น ค่าผลลัพธ์ และสมการ	17
ตารางที่ 3 คำอธิบายขั้นตอนการประมวลผลข้อมูลก่อนในรูปที่ 22	36
ตารางที่ 4 โครงสร้างตารางข้อมูล	45
ตารางที่ 5 ฟังก์ชันทางสถิติที่ใช้ในการสกัดคุณลักษณะสำคัญ	46
ตารางที่ 6 ฟังก์ชันทางสถิติที่ใช้ในการสกัดลักษณะสำคัญ (ต่อ)	47
ตารางที่ 7 ชื่อย่อและชื่อเต็มแบบจำลองในการทดลอง	48
ตารางที่ 7 ชื่อย่อและชื่อเต็มแบบจำลองในการทดลอง (ต่อ)	49
ตารางที่ 8 จำนวนข้อมูลที่ใช้ในการทดลองสำหรับชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบี	49
ตารางที่ 9 จำนวนข้อมูลที่ใช้ในการทดลองสำหรับชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดดี	49
ตารางที่ 10 สถิติข้อมูลเมื่อแบ่งตามคลาสของชุดข้อมูลฝึก ชุดข้อมูลประเมินและชุดข้อมูลทดสอบ .	50
ตารางที่ 11 รายละเอียดสปรังก์ครัสเตอร์ที่ใช้ในการประมวลผลข้อมูลก่อน การพัฒนาขั้นตอนวิธีและการทดลอง	54
ตารางที่ 12 เมทริกซ์ความสับสน	55
ตารางที่ 13 โครงสร้าง ค่าพารามิเตอร์ต่าง ๆ ของแบบจำลองในการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าในแต่ละแบบสำหรับโครงข่ายประสาทเทียมการเรียนรู้เชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้น	56
ตารางที่ 14 ชื่อเต็มของอักขรย่อแต่ละฟังก์ชันหรือองค์ประกอบของ โครงสร้าง ค่าพารามิเตอร์ต่าง ๆ ของแบบจำลองโครงข่ายประสาทเทียมการเรียนรู้เชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้นในตารางที่ ตารางที่ 13	57
ตารางที่ 15 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ โดยใช้แบบจำลองต้นไม้ตัดสินใจเอ็กซ์ทรีมเกรเดียนต์บูตดิ้งและไลต์เอ็กซ์จีบีเอ็ม	58

ตารางที่ 16 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ โดยใช้แบบจำลองถดถอยโลจิสติก.....	58
ตารางที่ 17 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ โดยใช้แบบจำลองป่าแบบสุ่ม	58
ตารางที่ 18 ผลการใช้เทคนิคการลดข้อมูลและการเลือกคุณลักษณะด้วยไลท์จีบีเอ็ม การวิเคราะห์องค์ประกอบสำคัญ และวิธีพื้นฐานร่วมกับการวิเคราะห์องค์ประกอบสำคัญ เปรียบเทียบกับวิธีพื้นฐานบนชุดข้อมูลฝึกรีโมทคอมเพรสเซอร์ชนิดบี โดย ‘*’ หมายถึงวิธีการเลือกคุณลักษณะที่นำเสนอ.....	59
ตารางที่ 19 ผลการใช้เทคนิคการลดข้อมูลและการเลือกคุณลักษณะด้วยไลท์จีบีเอ็ม การวิเคราะห์องค์ประกอบสำคัญ และวิธีพื้นฐานร่วมกับการวิเคราะห์องค์ประกอบสำคัญ เปรียบเทียบกับวิธีพื้นฐานบนชุดข้อมูลฝึกรีโมทคอมเพรสเซอร์ชนิดดี.....	60
ตารางที่ 20 F1 เฉลี่ยของแบบจำลองวิธีการเลือกคุณลักษณะบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบีและดี โดย ‘*’ หมายถึงวิธีการเลือกคุณลักษณะที่นำเสนอหรือค่า F1 สูงสุด และ ‘**’ หมายถึงแบบจำลองที่นำเสนอ.....	61
ตารางที่ 21 ค่า F1 เฉลี่ยของแต่ละแบบจำลองที่นำเสนอโดยใช้การเลือกคุณลักษณะบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดดีและดี และค่า F1 เฉลี่ยที่เพิ่มขึ้นเมื่อเทียบกับแบบจำลองที่นำเสนอ.....	64
ตารางที่ 22 ค่า F1 เฉลี่ยของแต่ละแบบจำลองที่นำเสนอโดยใช้การเลือกคุณลักษณะบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบีและดี และค่า F1 เฉลี่ยที่เพิ่มขึ้นเมื่อเทียบกับแบบจำลองที่นำเสนอ (ต่อ)	65
ตารางที่ 23 ค่า F1 ของแบบจำลอง ABDLSTM&ACNN ที่ดีที่สุดบนชุดข้อมูลทดสอบของชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบี.....	65
ตารางที่ 24 ค่า F1 ของแบบจำลอง ABDLSTM&ACNN ที่ดีที่สุด บนชุดข้อมูลทดสอบของชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดดี	65
ตารางที่ 25 การเปรียบเทียบค่า F1 เฉลี่ยของของคลาส 0 และ คลาส 1 ของชุดข้อมูลบี.....	66
ตารางที่ 26 การเปรียบเทียบค่า F1 เฉลี่ยของของคลาส 0 และ คลาส 1 ของชุดข้อมูลดี	66

สารบัญรูปภาพ

หน้า

รูปที่ 1 ภาพจากมุมมองด้านข้างและด้านบนของรีโมทคอมเพรสเซอร์ แสดงความยาว ความสูงและความกว้างของรีโมทคอมเพรสเซอร์ รุ่น C-SERIES 3516 ของ Exterran Holdings, Inc.....	1
รูปที่ 2 โครงสร้างพื้นฐานของแหล่งผลิตปิโตรเลียมของประเทศไทย ประกอบไปด้วยแท่นผลิตกลาง (CPP) แท่นที่พักอาศัย (Living Quarter Platform - LQP) แท่นหลุมผลิต (WP) เรือกักเก็บ (Floating Storage and Off loading terminal - FSO) และท่อ (pipeline).....	2
รูปที่ 3 (ก) โครงข่ายประสาทเทียมแบบวงกลับอย่างง่าย และ (ข) โครงข่ายประสาทเทียมแบบวงกลับอย่างง่ายตามเวลา [6, 9]	8
รูปที่ 4 ปัญหาวนิชชิงเกรเดียนต์ของโครงข่ายประสาทเทียมแบบวงกลับโดยความเข้มของสีในแต่ละบัพแสดงถึงการสูญหายไปของสารสนเทศนำเข้าลำดับที่ 1 ตามเวลา โดยที่ยังมีความเข้มมากแสดงให้เห็นถึงการสูญหายมาก โดยการสูญเสียดลงเมื่อเวลาผ่านไปเขียนทับจากการกระตุ้นของชั้นซ่อน [6]	8
รูปที่ 5 บล็อกหน่วยความจำของหน่วยความจำสั้นระยะยาว หนึ่งเซลล์หน่วยความจำ (memory cell) ประกอบไปด้วยสามประตู ได้แก่ ประตูนำเข้า (input gate) ประตูนำออก (output gate) และประตูลืม (forget gate) เป็นหน่วยการประมวลผลอย่างไม่เป็นเชิงเส้น (nonlinear summation units) ที่เก็บสารสนเทศจากฟังก์ชันกระตุ้น ของทั้งภายในและภายนอกบล็อก และควบคุมการกระตุ้นในเซลล์ [6]	9
รูปที่ 6 โครงข่ายหน่วยความจำระยะสั้นแบบยาว ประกอบด้วยชั้นนำเข้าจำนวน 4 บัพ ชั้นนำออกจำนวน 5 บัพและชั้นซ่อนจำนวน 2 บัพหรือ 2 บล็อกหน่วยความจำ [6].....	10
รูปที่ 7 การรักษาแตรนเดียร์นสารสนเทศโดยหน่วยความจำระยะสั้นแบบยาว [6]	11
รูปที่ 8 เซลล์หน่วยความจำระยะสั้นแบบยาว [6]	12
รูปที่ 9 โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง	13
รูปที่ 10 ชั้นคอนโวลูชันที่ประกอบไปด้วยองค์ประกอบกลางของเคอร์เนลจะถูกวางไว้เหนือเวกเตอร์อินพุตซึ่งจะถูกคำนวณและแทนที่ด้วยผลรวมถ่วงน้ำหนักของตัวเองและพิกเซลใกล้เคียง	15
รูปที่ 11 กระบวนการของโครงข่ายประสาทเทียมคอนโวลูชัน 1 มิติสำหรับอนุกรมเวลา	16

รูปที่ 12 โครงสร้างแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าโดยใช้ โครงข่ายประสาทเทียมแบบสองทิศทางร่วมกับกลไกจุดสนใจ	27
รูปที่ 13 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองที่มีจำนวนชั้นหน่วยความจำระยะ สั้นแบบยาวแบบสองทิศทางตั้งแต่หนึ่งถึงสี่ชั้น ทดสอบบนชุดข้อมูลประเมินผล รีโมทคอมเพรสเซอร์ ชนิดบี ของการเรียนรู้ 100 รอบแรก	28
รูปที่ 14 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองที่มีจำนวนชั้นคอนโวลูชันสามถึง หกชั้น ทดสอบบนชุดข้อมูลประเมินผล รีโมทคอมเพรสเซอร์ชนิดบี ของการเรียนรู้ 100 รอบแรก.....	29
รูปที่ 15 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองด้วยการปรับค่าการลดจำนวน (Drop-out) ทดสอบบนชุดข้อมูลประเมินผล รีโมทคอมเพรสเซอร์ชนิดบี ของการเรียนรู้ 100 รอบ แรก	30
รูปที่ 16 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองด้วยการปรับขนาดเบตของ ข้อมูลฝึก ทดสอบบนชุดข้อมูลประเมินผล รีโมทคอมเพรสเซอร์ชนิดบี ของการเรียนรู้ 100 รอบแรก	31
รูปที่ 17 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองด้วยการปรับค่าเรกูลาร์ไรเซชัน แบบแอลสอง ข้อมูลบนชุดข้อมูลประเมินผลชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบี ของการเรียนรู้ 100 รอบแรก	32
รูปที่ 18 โครงสร้างแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าโดยใช้ โครงข่ายประสาทเทียมหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางกลไกจุดสนใจผสานกับ โครงข่ายประสาทเทียมคอนโวลูชันกลไกจุดสนใจ.....	33
รูปที่ 19 อัลกอริทึมการแปลงเมทริกซ์ลำดับเวลาสามมิติ.....	34
รูปที่ 20 วิธีการแปลงข้อมูลอนุกรมเวลาแบบหลายตัวแปรโดยวิธีการเมทริกซ์ลำดับเวลาสามมิติหรือ การเลื่อนซ้อนทุก ๆ 48 ชั้นเวลา จากเวลาที่ 1 ถึงเวลาที่ 48 และวิธีการเลื่อนซ้อนจากเวลาที่ 2 ถึง เวลาที่ 49.....	35
รูปที่ 21 เมทริกซ์ลำดับเวลาจากหนึ่งอนุกรมเวลาสามมิติ	35
รูปที่ 22 เทคนิคการเพิ่มประสิทธิภาพโดยการสกัดลักษณะสำคัญร่วมกับเทคนิคการลดข้อมูล	36
รูปที่ 23 แนวคิดที่นำเสนอเพื่อลดปริมาณข้อมูลด้วยการสกัดคุณลักษณะสำคัญ โดยใช้ฟังก์ชันทาง สถิติ.....	37

รูปที่ 24 อัลกอริทึมการแปลงเมทริกซ์ลำดับเวลาสามมิติโดยการลดปริมาณข้อมูล	38
รูปที่ 25 การเลือกองค์ประกอบสำคัญโดยใช้เกณฑ์เปอร์เซ็นต์ความแปรปรวน ใช้ค่าขีดแบ่ง $L = 0.975$ ซึ่งทำให้สามารถเลือก 8 อันดับแรกขององค์ประกอบสำคัญไปใช้ในกระบวนการขั้นต่อไป.....	40
รูปที่ 26 การวิเคราะห์คะแนนความสำคัญสะสมของคุณลักษณะ โดยใช้แบบจำลองโลทจีบีเอ็ม	41
รูปที่ 27 การทดลอง 5 รอบ ในแต่ละรอบ กล้องเทอา่อนแสดงสัดส่วนชุดข้อมูลฝึก กล้องสีเทาเข้ม แสดงสัดส่วนชุดข้อมูลประเมินผล และกล้องสีขาวแสดงสัดส่วนชุดข้อมูลทดสอบ.....	49
รูปที่ 28 แสดงการเปรียบเทียบค่าฟังก์ชันต้นทุนระหว่างแบบจำลองที่นำเสนอกับแบบจำลองโครงข่ายประสาทเทียมเชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้นบนชุดข้อมูลประเมินผลของรีโมทคอมเพรสเซอร์ชนิดบี	52
รูปที่ 29 แสดงการเปรียบเทียบค่า F1 ระหว่างแบบจำลองที่นำเสนอกับแบบจำลองโครงข่ายประสาทเทียมเชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้นบนชุดข้อมูลประเมินผลของรีโมทคอมเพรสเซอร์ชนิดบี.....	52
รูปที่ 30 กราฟแสดงการเปรียบเทียบค่าฟังก์ชันต้นทุนต่อการเรียนรู้อันระหว่างแบบจำลองที่นำเสนอในแต่ละรอบการทดลองบนชุดข้อมูลประเมินผลของรีโมทคอมเพรสเซอร์ชนิดบี	53
รูปที่ 31 กราฟแสดงค่าฟังก์ชันต้นทุนต่อการเรียนรู้อันของรอบการทดลองที่สี่โดยจะใช้น้ำหนักในการเรียนรู้อันที่ได้ค่าฟังก์ชันต้นทุนที่ต่ำที่สุดในการประเมินประสิทธิภาพของแบบจำลองด้วยชุดข้อมูลทดสอบ	53
รูปที่ 32 ค่า F1 เกลี่ยบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบี ในรูปแบบกราฟแท่งจากการทดลองแบบจำลองที่นำเสนอกับวิธีการเลือกคุณลักษณะกับแบบจำลองการเรียนรู้เชิงลึกและแบบจำลองโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้น โดย ‘**’ หมายถึงแบบจำลองที่นำเสนอ	62
รูปที่ 33 ค่า F1 เกลี่ยบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดดี ในรูปแบบกราฟแท่งจากการทดลองแบบจำลองที่นำเสนอกับวิธีการเลือกคุณลักษณะกับแบบจำลองการเรียนรู้เชิงลึกและแบบจำลองโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้น โดย ‘**’ หมายถึงแบบจำลองที่นำเสนอ.....	62
รูปที่ 34 ค่า F1 เกลี่ยบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบี ในรูปแบบกราฟแท่งจากการทดลองแบบจำลองที่นำเสนอกับวิธีการเลือกคุณลักษณะกับแบบจำลองการเรียนรู้เชิงลึกที่นำเสนอกับแบบจำลองการเรียนรู้ของเครื่อง โดย ‘**’ หมายถึงแบบจำลองที่นำเสนอ.....	63

รูปที่ 35 ค่า F1 เกลี่ยบนชุดข้อมูลทดสอบของริโมทคอมเพรสเซอร์ชนิดดี ในรูปแบบกราฟแท่งจาก
การทดลองแบบจำลองที่นำเสนอเกี่ยวกับวิธีการเลือกคุณลักษณะกับแบบจำลองการเรียนรู้เชิงลึกที่
นำเสนอเกี่ยวกับแบบจำลองการเรียนรู้ของเครื่อง โดย ‘**’ หมายถึงแบบจำลองที่นำเสนอ..... 63

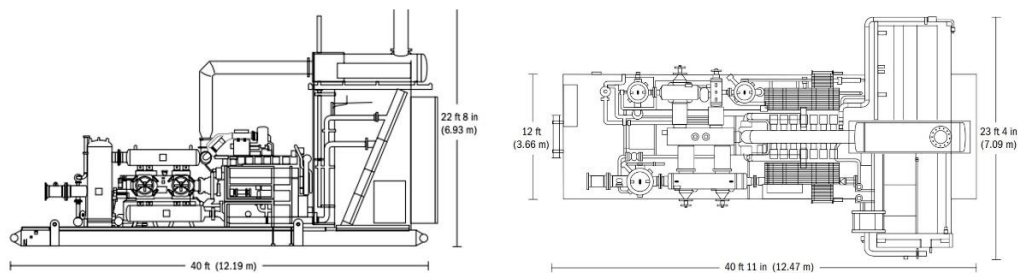


บทที่ 1

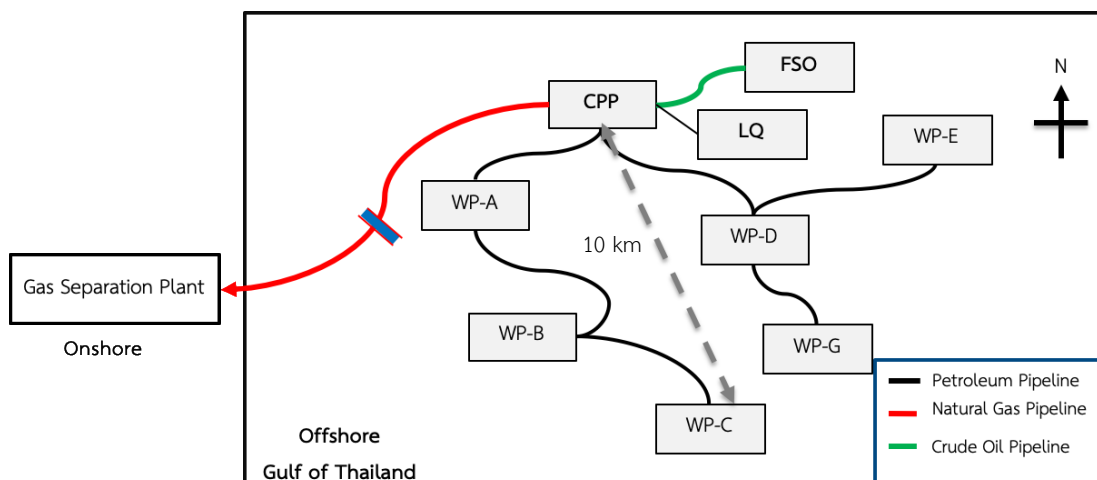
ที่มาและความสำคัญ

1.1 ที่มาและความสำคัญ

ในอุตสาหกรรมการผลิตและสำรวจปิโตรเลียม ก๊าซคอมเพรสเซอร์ (gas compressor) เป็นอุปกรณ์เครื่องกลที่ใช้ในการเพิ่มความดันและลดปริมาตรของก๊าซธรรมชาติ ทั้งยังถูกนำมาใช้ในกระบวนการทำให้เป็นของเหลว (liquefaction process) ทั้งในโรงงานบนชายฝั่ง (onshore plant) และโรงงานนอกชายฝั่ง (offshore plant) ในรูปที่ 1 [1] แสดงแบบของก๊าซคอมเพรสเซอร์ มีความยาว 12.9 เมตร สูง 6.93 เมตร และกว้างด้านหน้า 3.66 เมตร เมื่อนำไปติดตั้งบนแท่นหลุมผลิต (Wellhead Platform - WP) จะเรียกว่ารีโมทคอมเพรสเซอร์ (Remote Compressor - RC) มีหน้าที่หลักในการเพิ่มความดันให้กับก๊าซธรรมชาติในปิโตรเลียม เนื่องจากปิโตรเลียมไหลขึ้นมาจากแหล่งกักเก็บ (reservoir) ได้พื้นดินผ่านท่อส่ง (pipeline) เรียกว่าหลุมผลิต (well) ด้วยความดันธรรมชาติและเมื่อปิโตรเลียมไหลตามท่อส่งที่มีความยาวมาก ๆ จะสูญเสียความดัน จนมีไม่พอที่จะทำให้ปิโตรเลียมสามารถไหลไปยังแท่นผลิตกลาง (Central Processing Platform - CPP) ได้ วิธีการหนึ่งในการแก้ไขปัญหาคือการใช้รีโมทคอมเพรสเซอร์ แต่อย่างไรก็ตามเมื่อรีโมทคอมเพรสเซอร์หยุดทำงานโดยไม่ได้วางแผนไว้ล่วงหน้า เนื่องจากเซ็นเซอร์ความปลอดภัยตรวจพบความผิดปกติของสัญญาณ เกินกว่าค่ามาตรฐานที่กำหนดไว้ ระบบรักษาความปลอดภัย จะสั่งหยุดการทำงานของเครื่องโดยทันที จึงต้องมีการบำรุงรักษาเชิงแก้ไขอย่างไม่ได้วางแผนไว้ล่วงหน้า (Unplanned Corrective Maintenance - UCM) ดังนั้นรีโมท-คอมเพรสเซอร์ จึงถือว่าเป็นอุปกรณ์ที่มีความสำคัญต่อกระบวนการผลิตเป็นอย่างมาก เพราะผลกระทบของการหยุดทำงานของเครื่อง ทำให้ปิโตรเลียมไม่สามารถไหลไปยังแท่นผลิตกลางได้



รูปที่ 1 ภาพจากมุมมองด้านข้างและด้านบนของรีโมทคอมเพรสเซอร์ แสดงความยาว ความสูงและความกว้างของรีโมทคอมเพรสเซอร์ รุ่น C-SERIES 3516 ของ Exterran Holdings, Inc. [ที่มา <https://www.exterran.com/Content/Docs/Products/CSeries-3516-Gas-Compression-Package-English-A4.pdf> Accessed: January 2019]



รูปที่ 2 โครงสร้างพื้นฐานของแหล่งผลิตปิโตรเลียมของประเทศไทย ประกอบไปด้วยแท่นผลิตกลาง (CPP) แท่นที่พักอาศัย (Living Quarter Platform - LQP) แท่นหลุมผลิต (WP) เรือกักเก็บ (Floating Storage and Off loading terminal - FSO) และท่อ (pipeline)

รูปที่ 2 แสดงให้เห็นโครงสร้างพื้นฐานของแหล่งผลิตปิโตรเลียมของประเทศไทย ประกอบไปด้วยแท่นผลิตกลาง แท่นหลุมผลิตที่ห่างไกลกัน การบำรุงรักษาอุปกรณ์และเครื่องจักรที่แท่นหลุมผลิตต้องใช้การเดินทางทางเรือ หากเรือทคอมเพรสเซอร์หยุดการทำงานอย่างไม่ได้วางแผนล่วงหน้า หรือเกิดความล้มเหลว การจะเดินทางไปยังแท่นหลุมผลิตเพื่อทำการตรวจสอบและซ่อมบำรุงจะต้องมีการวางแผนเตรียมเจ้าหน้าที่ อุปกรณ์ ให้เพียงพอ และกำหนดเส้นทางเดินเรือให้คุ้มค่าที่สุด ซึ่งเป็นกระบวนการที่ใช้เวลาไม่ต่ำกว่าหนึ่งวัน ในบางกรณีการซ่อมบำรุงอาจใช้เวลานานหลายวัน ทำให้สูญเสียโอกาสในการผลิต (Loss Production Opportunity - LPO) และเกิดค่าบำรุงรักษาที่ไม่คาดคิด[1] ดังนั้นการวางแผนงานและระบบการบำรุงรักษา จึงมีความสำคัญเป็นอย่างมาก นโยบายการบำรุงรักษามี 3 ประเภท ได้แก่ 1) การบำรุงรักษาเชิงแก้ไข (corrective maintenance) คือจะดำเนินการแก้ไขเมื่อพบว่าอุปกรณ์เสียหาย 2) การบำรุงรักษาเชิงป้องกัน (preventive maintenance) ดำเนินการบำรุงรักษาตามรอบที่กำหนดไว้ และ 3) การบำรุงรักษาเชิงพยากรณ์ (predictive maintenance - PdM)[2] ที่จะพยากรณ์เหตุเสียหายหรืออาการเสียล่วงหน้า เพื่อให้ผู้มีหน้าที่ดำเนินการบำรุงรักษาวางแผนได้อย่างมีประสิทธิภาพ การบำรุงรักษาเชิงแก้ไข การบำรุงรักษาเชิงป้องกัน มีข้อด้อยที่ไม่สามารถช่วยให้ผู้ปฏิบัติงานเตรียมการไว้ล่วงหน้าว่าจะเกิดเหตุเสียหายกับอุปกรณ์ใดจึงทำให้มีปัญหาด้านค่าใช้จ่ายที่เพิ่มมากขึ้น[3] กระบวนการทำงานในปัจจุบันต้องอาศัยความเชี่ยวชาญของผู้ปฏิบัติงานในการตรวจจบบรรยากาศที่มีความผิดปกติโดยการวิเคราะห์รูปแบบสัญญาณของเซ็นเซอร์แต่ละตัวแต่เนื่องจากการเพิ่มขึ้นของอุปกรณ์และเครื่องมือวัด การติดตามและตรวจจบบางการทำงานของตัวอุปกรณ์เพื่อจะคาดการณ์ว่าจะเกิดเหตุเสียหายจึงเป็นความท้าทายต่อการดำเนินงาน การวางแผน รวมถึงการใช้ทรัพยากรบุคคลและการลงทุนเพิ่มขึ้นเป็นอย่างมาก

ด้วยการพัฒนาอย่างรวดเร็วของเทคโนโลยีสารสนเทศและเทคนิคด้านปัญญาประดิษฐ์ ประกอบกับการส่งเสริมยุคอุตสาหกรรม 4.0 (Industry 4.0 Era)[4] การค้นหาและรู้ถึงรูปแบบความผิดปกติก่อนที่จะเกิดเหตุเสียของอุปกรณ์ โดยประยุกต์ใช้การวิเคราะห์ข้อมูลเวลาจริง (real-time data analytics)[5] ที่สามารถประมวลเหตุการณ์ทั้งหลายขณะที่เหตุการณ์เหล่านั้นกำลังเกิดหรือใน ช่วงเวลาที่เฉพาะเจาะจง เป็นสิ่งที่สามารถช่วยลดงานให้กับผู้ดูแลอุปกรณ์นั้น ๆ เป็นอย่างมาก มีการวิจัยการประยุกต์ใช้การเรียนรู้ของเครื่อง (machine learning) เพื่อใช้ในการทำนายเหตุขัดข้องครั้งต่อไปของ โปรดักชันก๊าซคอมเพรสเซอร์ (production gas compressor) ณ ที่เรือผลิตและกักเก็บ (Floating Production Storage and Offloading - FPSO) โดยใช้ข้อมูลการสั่นสะเทือนและข้อมูลประวัติเหตุการณ์ความล้มเหลวของเครื่อง และได้นำเสนอ แบบจำลอง 3 แบบ ได้แก่ 1) แบบจำลอง มาร์คอฟ (Markov model) 2) แบบจำลองการถดถอยเชิงเส้น (linear regression model) และ 3) แบบจำลองแบบผสม (hybrid model) โดยได้ผลลัพธ์จาก ค่าปรับปรุงร้อยละความคลาดเคลื่อน สมบูรณ์-เฉลี่ย (Modified Absolute Percentage Error - MAPE) ที่ 37.09% ผู้วิจัยรายงานว่า มีข้อจำกัดในการปรับปรุงค่าพารามิเตอร์ในเวลาจริงและการเสื่อมของอุปกรณ์อื่นเนื่องจากความไม่แน่นอนต่าง ๆ [1] ถึงอย่างไรก็ตามองค์ประกอบและหน้าที่หลักของโปรดักชันแก๊สคอมเพรสเซอร์และ รีโมทคอมเพรสเซอร์มีความแตกต่างกันอย่างมาก ในกรณีที่ข้อมูลสัญญาณของเครื่องมือวัดที่เก็บไว้ในระบบฐานข้อมูลอนุกรมเวลา (time series database) มีปริมาณมาก การทำนายโดยใช้แบบจำลอง โครงข่ายประสาทเทียมที่นิยมนำมาใช้ในการสร้างแบบจำลองของความสัมพันธ์ของข้อมูลที่ไม่เป็นเชิงเส้น (non-linear) ซึ่งโดยทั่วไปนั้นถ้าสร้างแบบจำลองที่มีชั้นซ่อน (hidden layer) ชั้นเดียว มีความสามารถที่ต่ำกว่าในการนำมาสร้างแบบจำลองที่มีความไม่เป็นเชิงเส้นมากกว่า เมื่อเทียบกับการใช้โครงข่ายประสาทเทียมเชิงลึก (deep neural network)[5] Jolliffe และคณะ [6] นำเสนอ การใช้หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory – LSTM) เนื่องจากมีความเหมาะสมในค้นหารูปแบบของข้อมูลอนุกรมเวลา ตลอดจนทำนายความล้มเหลวของอุปกรณ์ มีงานวิจัยที่เกี่ยวข้องโดยการใช้ โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network - CNN) เพื่อใช้ในการพยากรณ์อายุการใช้งานที่คงเหลือของเครื่องจักรโดยใช้ร่วมกันกับหน่วยความจำระยะสั้นแบบยาว[7]และการพยากรณ์อนุกรมเวลาแบบหลายตัวแปรโดยใช้การผสมผสานโครงข่ายประสาทเทียมคอนโวลูชันและหน่วยความจำระยะสั้นแบบยาวเพื่อเสริมประสิทธิภาพในการพยากรณ์[8]

งานวิจัยนี้มีจุดมุ่งหมายเพื่อนำเสนอโครงสร้างแบบจำลอง การทำนายความล้มเหลวของ รีโมทคอมเพรสเซอร์ล่วงหน้าโดยใช้เทคนิคการเรียนรู้เชิงลึก เทคนิคการประมวลผลข้อมูลก่อน เพื่อให้สามารถประมวลผลข้อมูลอนุกรมเวลาขนาดใหญ่ได้อย่างมีประสิทธิภาพ อีกทั้งเพื่อเพิ่มความแม่นยำที่สูงเมื่อเปรียบเทียบกับอัลกอริทึมและวิธีการประมวลผลข้อมูลก่อน ที่ได้รับความนิยมในงานวิจัยที่เกี่ยวข้อง อีกทั้งเพื่อเป็นแนวทางในการประยุกต์ใช้แบบจำลองและเทคนิคการประมวลผลข้อมูล

ก่อนที่นำเสนอ ในการทำนายความล้มเหลวล่วงหน้าของอุปกรณ์ชนิดอื่นในอุตสาหกรรมการผลิตและสำรวจปิโตรเลียมได้

1.2 วัตถุประสงค์

งานวิจัยนี้มีวัตถุประสงค์ 3 ข้อดังต่อไปนี้

1) เพื่อสร้างแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าโดยใช้โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางที่ใช้กลไกจุดสนใจร่วมกับโครงข่ายประสาทเทียมคอนโวลูชัน

2) เพื่อแสดงการเปรียบเทียบประสิทธิภาพระหว่างการใช้เทคนิคการเลือกคุณลักษณะ โดยการลดมิติของข้อมูลด้วยการวิเคราะห์องค์ประกอบสำคัญร่วมกับการลดปริมาณข้อมูลโดยการสกัดคุณลักษณะโดยใช้ฟังก์ชันทางสถิติกับวิธีพื้นฐาน

3) เพื่อแสดงการเปรียบเทียบประสิทธิภาพแบบจำลองที่นำเสนอกับแบบจำลองโครงข่ายประสาทเทียมคอนโวลูชันร่วมกับกลไกจุดสนใจ โครงข่ายประสาทเทียมคอนโวลูชัน โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางร่วมกับกลไกจุดสนใจ โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง โครงข่ายหน่วยความจำระยะสั้นแบบยาว โครงข่ายประสาทเทียม ป่าแบบสุ่ม ถดถอยโลจิสติก ต้นไม้ตัดสินใจเอ็กซ์ทรีมเกรเดียนต์บูตติง และไลท์เกรเดียนต์บูตติงแมทซิง

1.3 ขอบเขตการดำเนินงาน

วิทยานิพนธ์นี้มีขอบเขตในการทำวิจัย 3 ข้อดังต่อไปนี้

1) ข้อมูลของวิทยานิพนธ์ฉบับนี้ได้แก่ ข้อมูลประวัติการดำเนินการของรีโมทคอมเพรสเซอร์จากระบบฐานข้อมูลอนุกรมเวลา เช่น ข้อมูลอุณหภูมิ ความเร็วของกังหัน แรงสั่นสะเทือนของเครื่อง ความดัน และข้อมูลประวัติเหตุเสียของรีโมทคอมเพรสเซอร์ ตั้งแต่ เดือน มกราคม พ.ศ. 2559 ถึง เดือน กุมภาพันธ์ พ.ศ. 2561

2) ใช้ภาษา ภาษา Python TensorFlow Keras pySpark Spark และ MS Azure Databricks ในกระบวนการสร้างแบบจำลอง

3) เปรียบเทียบประสิทธิภาพของแบบจำลองโดยใช้ F1 และพิจารณา ความแม่นยำ (precision) ค่าความระลึก (recall) ประกอบในการเลือกแบบจำลองที่จะนำไปใช้จริง

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1) สามารถแสดงให้เห็นถึงการเพิ่มประสิทธิภาพใน การทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าได้

2) สามารถให้ข้อมูลสนับสนุนการตัดสินใจของวิศวกรผู้ดูแลอุปกรณ์ ในการเตรียมความพร้อม จัดสรรครุ เครื่องมือ ผู้เชี่ยวชาญ และวางแผนเส้นทางเดินเรือ เพื่อไปยังแท่นหลุมผลิต เพื่อตรวจสอบและซ่อมบำรุงได้อย่างคุ้มค่าและมีประสิทธิภาพเพิ่มขึ้น

3) นำขั้นตอนวิธีการพัฒนาแบบจำลองนี้ไปประยุกต์ใช้การสร้างแบบจำลองสำหรับการทำนายความล้มเหลวล่วงหน้ากับอุปกรณ์ชนิดอื่นได้

1.5 ขั้นตอนการดำเนินงาน

1) ศึกษางานวิจัยที่เกี่ยวข้องกับการบำรุงรักษาเชิงพยากรณ์ และ ศึกษางานวิจัยที่เกี่ยวข้องกับการทำนายความล้มเหลวของอุปกรณ์ในธุรกิจน้ำมันและก๊าซธรรมชาตินอกชายฝั่ง

2) ศึกษาหลักการทํางาน ประเภท องค์ประกอบของรีโมทคอมเพรสเซอร์ของแท่นหลุมผลิตนอกชายฝั่ง

3) ศึกษาอัลกอริทึมและเทคนิคการเรียนรู้ของเครื่อง ที่คาดว่าจะใช้ในงานวิจัยในการจำแนกประเภทความล้มเหลวและการการทำนายความล้มเหลวล่วงหน้า

4) ศึกษาเครื่องมือที่ใช้ในงานวิจัย เช่น เครื่องมือในการจัดเก็บและประมวลผลข้อมูลจากระบบ PI Historian และเครื่องมือในการใช้สร้างแบบจำลอง

5) ออกแบบการทดลอง ออกแบบและสร้างระบบฐานข้อมูลจัดเก็บข้อมูลอนุกรมเวลาและข้อมูลเชิงความสัมพันธ์ที่ใช้ในการทำวิจัย

6) สร้างแบบจำลองในการทำนายความล้มเหลวล่วงหน้าของรีโมทคอมเพรสเซอร์

7) ทดสอบและวัดผลความถูกต้องในการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์

8) วิเคราะห์ผลการทดลอง

9) สรุปผลและเรียบเรียงวิทยานิพนธ์

1.6 สรุปผลภาพรวมของการทำวิจัย

จากวัตถุประสงค์และขอบเขตการดำเนินงานทั้งหมด ผู้วิจัยได้ทำการทดลองและดำเนินการสำเร็จเรียบร้อยแล้ว โดยได้สรุปไว้ในบทที่ 6

ตารางที่ 1 แผนการดำเนินงาน

ลำดับ	ขั้นตอนงานวิจัย	2562										2563								
		ก.พ.	มี.ค.	เม.ย.	พ.ค.	มี.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.	
1	ศึกษางานวิจัยที่เกี่ยวข้อง																			
2	ศึกษาทฤษฎีที่เกี่ยวข้องกับการบำรุงรักษาเชิงพยากรณ์																			
3	ศึกษาขั้นตอนการทำงานวิจัยและออกแบบการทดลองเบื้องต้น																			
4	ทำการทดลองเบื้องต้นและสรุปผลการทดลอง																			
5	สอบหัวข้อวิทยานิพนธ์																			
6	ดำเนินการทำการทดลองตามที่นำเสนอ																			
7	วิเคราะห์ผลการทดลอง																			
8	สรุปผลและเรียบเรียงวิทยานิพนธ์																			
9	สอบวิทยานิพนธ์																			

บทที่ 2

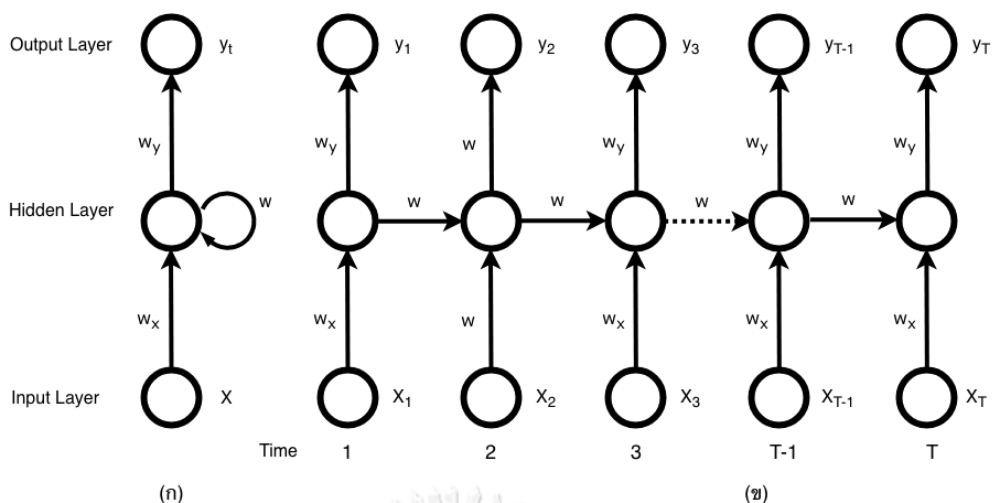
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

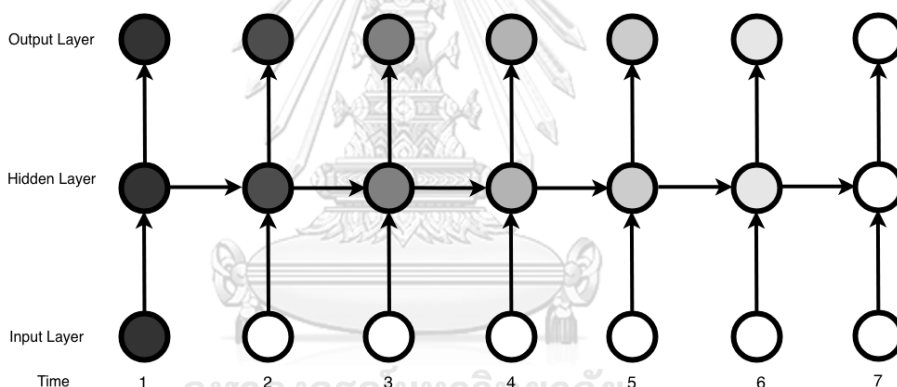
วิทยานิพนธ์นี้เกี่ยวข้องกับโครงข่ายประสาทเทียมแบบวนกลับ หน่วยความจำระยะสั้นแบบยาว โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง โครงข่ายประสาทเทียมคอนโวลูชัน กลไกจุดสนใจ การวิเคราะห์ส่วนองค์ประกอบสำคัญ และการวัดประสิทธิภาพด้วยค่า F1

2.1.1 โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network - RNN)

โครงสร้างของโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า [6] เป็นโครงข่ายแบบไม่มีวงวน ซึ่งลักษณะที่สำคัญของโครงข่ายประสาทเทียมแบบวนกลับคือมีวงวน ข้อมูลนำเข้าและข้อมูลนำออกที่เป็นอนุกรมเวลา ซึ่งแตกต่างจากลักษณะของข้อมูลนำเข้าและข้อมูลนำออกแบบโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้าที่เป็นข้อมูลที่เพิ่มขึ้นอยู่กับเวลา ในการจำแนกหรือรู้จำข้อมูล จะจำแนกข้อมูลข้อมูลแต่ละตัว ในขณะที่โครงข่ายประสาทเทียมแบบวนกลับทำการรู้จำหรือจำแนกข้อมูลนำเข้าเป็นอนุกรมเวลา และให้ข้อมูลนำออกเป็นอนุกรมเวลา ทำให้มีประโยชน์ที่สำคัญ คือนอกจากสามารถเรียนรู้สารสนเทศจากสถานะตนเองแล้วยังสามารถเรียนรู้สารสนเทศจากสถานะก่อนหน้าได้คล้ายกับการเรียนรู้บริบทของสิ่งที่เรียนรู้ทำให้เหมาะสมกับการรู้จำข้อมูลรูปแบบสัญญาณจากเครื่องมือวัดที่เป็นไปตามลำดับ โครงข่ายประสาทเทียมแบบวนกลับอย่างง่าย (Simple Recurrent Neural Network) แสดงในรูปที่ 3 (ก) โดยมีชั้นนำเข้า (input layer) ชั้นนำออก (output layer) และชั้นซ่อนชั้นละ 1 บัพ (node) และมีค่าถ่วงน้ำหนัก (weight) 3 ค่า ประกอบด้วยค่าถ่วงน้ำหนักจากชั้นนำเข้าไปยังชั้นซ่อนแทนด้วย w_x ค่าถ่วงน้ำหนักจากชั้นซ่อนไปยังชั้นซ่อนแทนด้วย w และค่าถ่วงน้ำหนักจากชั้นซ่อนไปยังชั้นนำออกแทนด้วย w_y เมื่อมีข้อมูลนำเข้าและข้อมูลนำออกจำนวน T ตัว สามารถขยายรูปที่ 3 (ก) เป็นดังรูปที่ 3 (ข)[6]



รูปที่ 3 (ก) โครงข่ายประสาทเทียมแบบวนกลับอย่างง่าย และ (ข) โครงข่ายประสาทเทียมแบบวนกลับอย่างง่ายตามเวลา [6, 9]

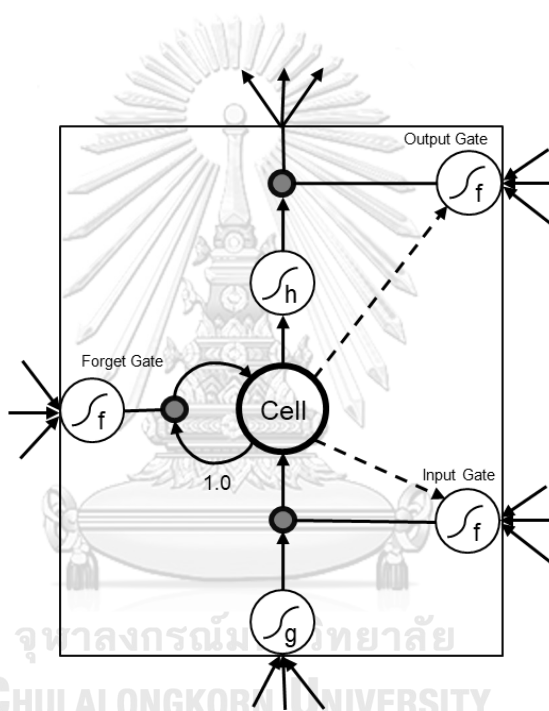


รูปที่ 4 ปัญหา Vanishing Gradient Problem ของโครงข่ายประสาทเทียมแบบวนกลับโดยความเข้มของสีในแต่ละบัพแสดงถึงการสูญหายไปของสารสนเทศนำเข้าลำดับที่ 1 ตามเวลา โดยที่ยังมีความเข้มมากแสดงให้เห็นถึงการสูญหายมาก โดยการสูญเสียลดลงเมื่อเวลาผ่านไปเขียนทับจากการกระตุ้นของชั้นซ่อน [6]

ปัญหาที่เรียกว่า ปัญหา Vanishing Gradient Problem (Vanishing Gradient Problem) เกิดจากเมื่อทำการเรียนรู้ข้อมูลนำเข้าทั้งหมดจนเสร็จสิ้น จะได้ข้อมูลนำออกตั้งแต่ 1 จนถึง T ซึ่งถ้าค่าผลคูณของข้อมูลนำเข้าลำดับที่ 1 (x_1) กับ w ทุกรอบจะทำให้ข้อมูลลำดับที่ 1 จะสูญหายไปตามเวลา แสดงในรูปที่ 4 ส่งผลให้โครงข่ายนั้นไม่สามารถรักษาสารสนเทศไว้ได้เป็นระยะเวลาอันยาวนาน และส่งผลให้ไม่เหมาะสำหรับการนำเข้าข้อมูลขนาดใหญ่ที่มีรูปแบบการขึ้นตรงต่อลำดับระยะยาวเพราะสารสนเทศของข้อมูลนำเข้าจะสูญหายไปตามเวลาอย่างรวดเร็วทำให้ไม่สามารถนำสารสนเทศของสถานะก่อนหน้ามาใช้ให้เกิดประโยชน์เท่าที่ควร

2.1.2 หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory - LSTM) [6]

สถาปัตยกรรมของหน่วยความจำระยะยาวถูกออกแบบมาเพื่อให้มีที่แก้ไขปัญหานวนิชเชิงเรเดียนต์ การพยากรณ์สารสนเทศนำออกไม่เพียงขึ้นอยู่กับสารสนเทศนำเข้าในเวลาเดียวกัน แต่ยังนำสารสนเทศของสถานะก่อนหน้ามาประมวลผลร่วมในการพยากรณ์ เมื่อเวลาผ่านไปสารสนเทศที่ควรอยู่ไม่สูญหายไปประหว้างทางตามเวลาหรือคงอยู่เพียงช่วงระยะเวลาสั้น ๆ โดยหน่วยความจำระยะยาว สามารถรักษาสารสนเทศที่มีความสำคัญให้คงอยู่ได้นานขึ้น มาช่วยพยากรณ์สารสนเทศนำออกในสถานะปัจจุบัน ซึ่งมีความเหมือนกับการดูบริบทของสารสนเทศข้างเคียงเพื่อช่วย หน่วยความจำระยะยาว สามารถเก็บค่าคงที่หรือสารสนเทศ



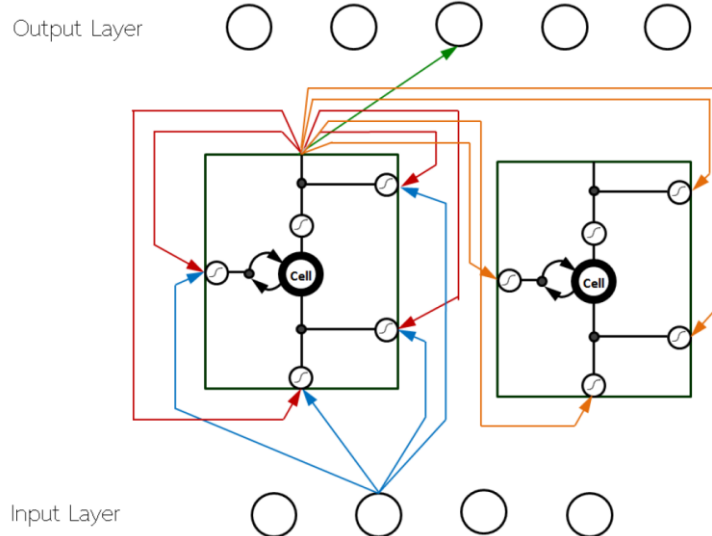
รูปที่ 5 บล็อกหน่วยความจำของหน่วยความจำระยะยาว หนึ่งในเซลล์หน่วยความจำ (memory

cell) ประกอบไปด้วยสามประตู ได้แก่ ประตูนำเข้า (input gate)

ประตูนำออก (output gate) และประตูลืม (forget gate) เป็นหน่วยการประมวลผลอย่างไม่เป็นเชิงเส้น (nonlinear summation units) ที่เก็บสารสนเทศจากฟังก์ชันกระตุ้น ของทั้งภายในและภายนอกบล็อก และควบคุมการกระตุ้นในเซลล์ [6]

จากรูปที่ 5 ประตูทั้งสามทำหน้าที่ควบคุมการรับและส่งสารสนเทศจากภายในและภายนอกบล็อกหน่วยความจำ ผลรวมสารสนเทศที่ได้จะถูกนำมาผ่านฟังก์ชันกระตุ้น (activation function) และผ่านตัวคูณ (วงกลมสีดำ) ประตูนำเข้าทำหน้าที่ควบคุมการรับข้อมูลนำเข้า

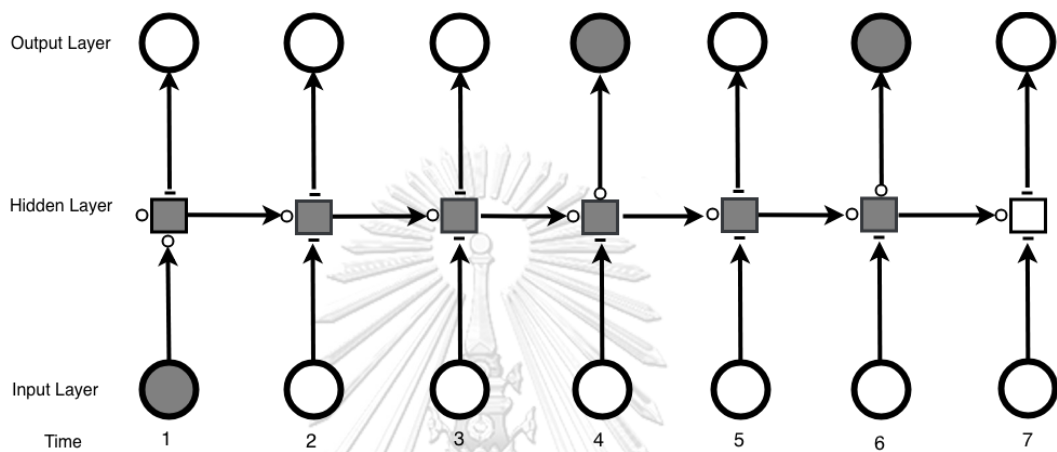
สถานะปัจจุบัน โดยนำค่าที่ได้จากผลรวมและข้อมูลนำเข้ามาผ่านตัวคูณก่อนเข้าสู่เซลล์หน่วยความจำ ประตุนำออกทำหน้าที่ควบคุมการส่งค่าคงที่ถูกเก็บในเซลล์หน่วยความจำในสถานะปัจจุบัน โดยนำค่าที่ได้จากผลรวมและค่าคงที่ในเซลล์หน่วยความจำผ่านฟังก์ชันกระตุ้นและตัวคูณก่อนส่งสารสนเทศไปเป็นข้อมูลนำออกและส่งไปยังสถานะถัดไป ประตูลิมนำหน้าที่จดจำหรือเปลี่ยนแปลงค่าคงที่ในเซลล์หน่วยความจำ โดยค่าที่ได้จากผลรวมจะเป็นตัวกำหนดการจดจำหรือเปลี่ยนแปลง ค่าคงที่ซึ่งถ้าจะนำค่าที่อยู่ภายในเซลล์หน่วยความจำจะเชื่อมต่อกับตัวเอง ฟังก์ชันกระตุ้น 'f' มักใช้โลจิสติกซิกมอยด์ (logistic sigmoid) ค่าอยู่ระหว่าง 0 (ประตูปิด) ถึง 1 (ประตูเปิด) ฟังก์ชันกระตุ้นที่เข้าและออกจากเซลล์ 'g' และ 'h' มักใช้ ฟังก์ชัน ไฮเพอร์โบลิกแทนเจนต์ (hyperbolic tangent - tanh) หรือ ใช้โลจิสติกซิกมอยด์ ซึ่งบางครั้ง 'h' อาจเป็นฟังก์ชันเอกลักษณ์ (identity function) ค่าถ่วงน้ำหนักเท่ากับ 1.0 การส่งสารสนเทศไปยังบัพทั้งหมดภายในโครงข่ายหน่วยความจำระยะสั้นแบบยาวสามารถแสดงได้ดังรูปที่ 6[6]



รูปที่ 6 โครงข่ายหน่วยความจำระยะสั้นแบบยาว ประกอบด้วยชั้นนำเข้าจำนวน 4 บัพ ชั้นนำออกจำนวน 5 บัพและชั้นซ่อนจำนวน 2 บัพหรือ 2 บล็อกหน่วยความจำ [6]

จากรูปที่ 6 เส้นเชื่อมที่มีสีส้มเป็นค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำและค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำไปยังบัพในชั้นนำออกแสดงด้วยเส้นเชื่อมสีเขียว บล็อกหน่วยความจำสามารถช่วยนำสารสนเทศออกจากบล็อกหน่วยความจำไปยังชั้นสารสนเทศนำออกผ่านประตุนำออกและสารสนเทศของข้อมูลนำเข้าในสถานะก่อนหน้าสามารถถูกจำหรือลืมไปจากสถานะเดิม โดยมีสารสนเทศนำเข้าได้ 4 ทางในแต่ละบล็อกหน่วยความจำและมีเพียง 1 สารสนเทศนำออก โดยมีเส้นเชื่อมที่แสดงค่าถ่วงน้ำหนักเป็นเพียง 1 เท่านั้น และเส้นเชื่อมสีฟ้าแสดงค่าถ่วง

น้ำหนักระหว่างบัพในชั้นสารสนเทศนำเข้าต่อบล็อกหน่วยความจำ ค่าถ่วงน้ำหนักระหว่างบล็อกหน่วยความจำเดียวกันแสดงด้วยเส้นเชื่อมสี ดังนั้นจึงทำให้ในแต่ละสถานะมีการนำสารสนเทศของข้อมูลเข้าในสถานะก่อนหน้าและสถานะเดียวกันมาใช้ผ่านประตูนำเข้าและประตูลิ้ม ทำให้สารสนเทศในโครงข่ายหน่วยความจำระยะสั้นแบบยาวเดินทางข้ามช่วงเวลาไปยังสถานะต่อไปได้ยาวนานยิ่งขึ้น ดังรูปที่ 7 [8]



รูปที่ 7 การรักษาแกระนเดียร์นสารสนเทศโดยหน่วยความจำระยะสั้นแบบยาว [6]

ในรูปที่ 7[6] ความเข้มของสีในแต่ละบัพบ่งชี้ถึงการสูญเสียแกระนเดียร์นสารสนเทศไปยังหน่วยนำเข้าในหนึ่งหน่วยเวลา ในกรณีนี้ บัพสี่เทาคือรูปทรงที่ได้รับสารสนเทศจากเวลาที่ 1 ที่มีความสูญเสียมากที่สุดและสีขาวคือไม่เกี่ยวข้องหรือไม่มีการสูญเสียของสารสนเทศจากเวลาที่ 1 เพื่อให้เป็นการอธิบายอย่างง่าย ประตูเปิดแทนด้วยสัญลักษณ์วงกลมขนาดเล็ก (o) และ ประตูปิดแทนด้วยเส้นตรง (-) โดยจะถูกวางทางขวาล่าง ขวาบนและซ้ายของบล็อกหน่วยความจำ คือ ประตูนำเข้า และประตูลิ้ม โดยเซลล์หน่วยความจำจะจดจำ หน่วยนำเข้าอันดับแรกหรือ ณ เวลาที่หนึ่งตราบจนประตูลิ้มจะเปิดออกและประตูนำเข้าปิด การสูญเสียของชั้นนำออกสามารถสลับเปิดและปิดด้วยประตูนำเข้าโดยไม่สูญเสียสารสนเทศไประหว่างทาง

การนำโครงข่ายประสาทเทียมแบบวนกลับมาขยายไปตามเวลาจะได้โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า ข้อมูลนำเข้าที่เรียนรู้เป็นค่าช่วงเวลาแบบไม่ต่อเนื่อง และให้ข้อมูลนำออกเป็นความน่าจะเป็นในช่วงเวลาแบบไม่ต่อเนื่องขนาดน้อยกว่าหรือเท่ากับข้อมูลนำเข้า หลังจากนั้นสามารถใช้การจำแนกประเภทการเชื่อมต่อตามเวลาแปลงข้อมูลนำออกเพื่อแปลเป็นข้อมูลที่สนใจต่อไป โดยแนวคิดของการจำแนกประเภทการเชื่อมต่อตามเวลาคือการแปลงข้อมูลนำออกจากโครงข่ายประสาทเทียมแบบวนกลับเป็นค่าความน่าจะเป็นของป้ายกำกับทั้งหมดและนำชุดข้อมูลที่ได้ไปเปลี่ยนเป็นข้อมูลนำออกที่แท้จริงต่อไป

2.1.3 หน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง (Bidirectional Long Short-term Memory - BLSTM) [7]

จากรูปที่ 8 เซลล์หน่วยความจำระยะสั้นแบบยาว [7] เมื่อกำหนดให้ ลำดับนำเข้า (input sequence) $x = x_1, \dots, x_T$, อาร์เอ็นเอ็น ค่ารวมเวกเตอร์ภายในคือ $h = (h_1, \dots, h_T)$ และ ลำดับเวกเตอร์นำออก $y = (y_1, \dots, y_T)$ โดยการวนซ้ำสมการดังต่อไปนี้ จาก $t = 1$ ถึง T

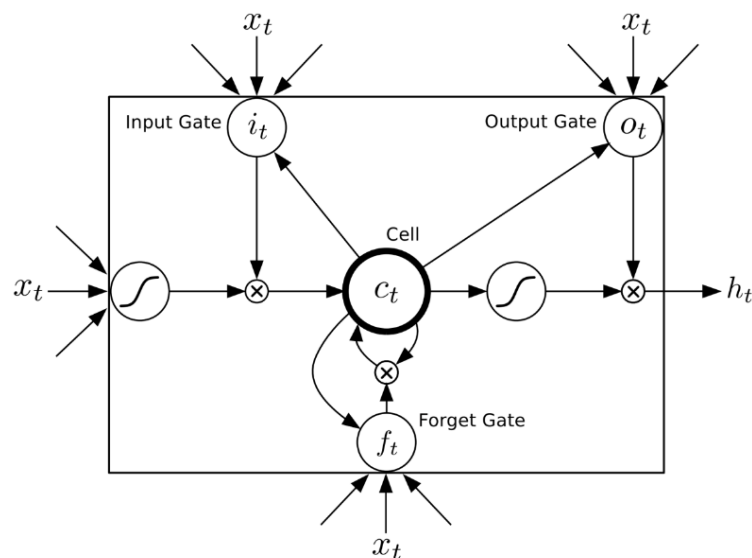
$$h_t = H(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (2.1)$$

$$y_t = W_{yh}h_t + b_y \quad (2.2)$$

เมื่อ W แสดงให้เห็นถึงเมทริกซ์น้ำหนัก เช่น W_{xh} คือ เมทริกซ์น้ำหนักนำเข้าที่ชั้นซ่อน (input-hidden weight matrix)

b แสดงให้เห็นถึงค่าเวกเตอร์อคติ เช่น b_h คือ เวกเตอร์ค่าอคติชั้นซ่อน (hidden bias vector)

H คือ ฟังก์ชันของชั้นซ่อน (hidden layer function)



รูปที่ 8 เซลล์หน่วยความจำระยะสั้นแบบยาว [6]

เมื่อ

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (2.3)$$

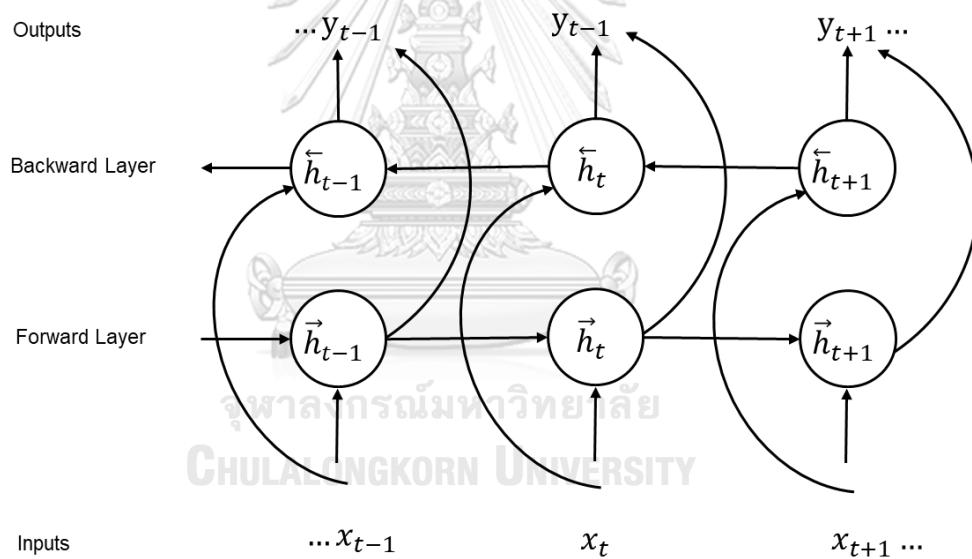
$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (2.4)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}x_t + W_{hc}h_{t-1} + b_c) \quad (2.5)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (2.6)$$

$$h_t = o_t \tanh(c_t) \quad (2.7)$$

และ σ คือ ฟังก์ชัน โลจิสติกซิกมอยด์ และ i, f, o และ c คือ ประตุนำเข้า ประตูลืม ประตุนำออก และ เวกเตอร์เซลล์กระตุ้น ตามลำดับ ทั้งหมดมีขนาดเท่ากันในเวกเตอร์ซ่อน h



รูปที่ 9 โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง

โครงข่ายประสาทเทียมแบบวกกลับแบบสองทิศทาง (Bidirectional Recurrent Neural Network - BRNN) คำนวณ ลำดับซ่อนอยู่ไปข้างหน้า \vec{h} ลำดับซ่อนอยู่ไปข้างหลัง \overleftarrow{h} และ ลำดับนำออก y โดยการเขียนไปตามลำดับจากข้างหลังจาก $t = T$ ถึง 1 และชั้นด้านหน้าจาก $t = 1$ ถึง T ต่อจากนั้นทำการปรับปรุงชั้นนำออก

$$\vec{h}_t = H(W_{x\vec{h}}x_t + W_{\vec{h}\vec{h}}\vec{h}_{t-1} + b_{\vec{h}}) \quad (2.8)$$

$$\vec{h}_t = H(W_{x\vec{h}}x_t + W_{\vec{h}\vec{h}}\vec{h}_{t+1} + b_{\vec{h}}) \quad (2.9)$$

$$y_t = W_{x\vec{h}}\vec{h}_t + W_{\vec{h}y}\vec{h}_t + b_y \quad (2.10)$$

สมมติว่าฟังก์ชันเลเยอร์เดียวกันนี้ใช้สำหรับ N เลเยอร์ในสแต็ก h^n คือลำดับเวกเตอร์ที่ซ่อนอยู่ คำนวณ จาก $n = 1$ ถึง N และ $t = 1$ ถึง T

$$h_t^n = H(W_{h^{n-1}h^n}h_t^{n-1} + W_{h^n h^n}h_t^{n-1} + b_h^n) \quad (2.11)$$

เมื่อกำหนดให้ $h^0 = x$ โครจข่ายนำออก y_t คือ

$$y_t = W_{h^N y}h_t^N + b_y \quad (2.12)$$

การรวมปีอาร์เอ็นเอ็นเข้ากับแอลเอสทีเอ็ม เกิดเป็นแอลเอสทีเอ็มแบบสองทิศทาง ซึ่งสามารถเข้าถึงบริบทระยะยาวในทั้งสองทิศทางนำเข้าดังแสดงในรูป 9

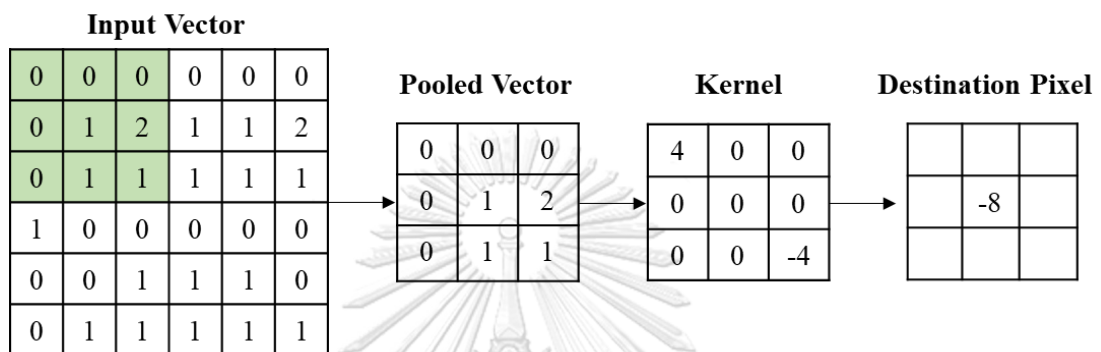
2.1.4 โครจข่ายประสาทเทียมคอนโวลูชัน (Convolution Neural Network - CNN) [10]

โครจข่ายประสาทเทียมคอนโวลูชัน เป็นโครจข่ายประสาทเทียมเชิงลึก เริ่มแรกเกิดจากงานวิจัยด้านการรู้จำตัวอักษรและรูปภาพ ด้วยสถาปัตยกรรมที่เรียบง่าย แต่มีความแม่นยำ มีความคล้ายคลึงกับโครจข่ายประสาทเทียมแบบธรรมดา ซึ่งประกอบด้วยนิรอรอลที่ปรับตัวเองให้เหมาะสมผ่านการเรียนรู้ แต่ละนิรอรอลจะยังคงได้รับอินพุตและดำเนินการพื้นฐานของโครจข่ายประสาทเทียมแบบธรรมดา นิรอรอลชั้นภายในซีเอ็นเอ็นนั้นประกอบไปด้วยนิรอรอลที่แบ่งออกเป็นสามมิติคือมิติเชิงพื้นที่ของอินพุต ความสูง ความกว้าง และความลึก นิรอรอลภายในเลเยอร์ที่กำหนดจะเชื่อมต่อกับพื้นที่เล็ก ๆ ของเลเยอร์ก่อนหน้านั้นเท่านั้น ตัวอย่าง เช่น อินพุตจะมีค่ามิติของ $64 \times 64 \times 3$ (ความสูง ความกว้างและความลึก) ซึ่งนำไปสู่ชั้นเอาต์พุตสุดท้ายประกอบด้วยขนาด $1 \times 1 \times n$ โดยที่ n แทนจำนวนคลาสที่เป็นไปได้ ดังนั้นจะย่อมิติข้อมูลเข้าแบบเต็มรูปแบบให้เป็นระดับค่าที่น้อยลงในมิติความลึก ฟังก์ชันการทำงานพื้นฐานของโครจข่ายประสาทเทียมคอนโวลูชัน สามารถแบ่งออกเป็น 3 ส่วนหลัก ๆ ดังต่อไปนี้

1. ชั้นคอนโวลูชัน (Convolutional layer)

ในโครจข่ายประสาทเทียมคอนโวลูชัน ประกอบด้วย ชั้นคอนโวลูชันทำหน้าที่สำคัญ โดยมีชุดของตัวกรอง เคอร์เนลฟิลเตอร์ (kernel filter) ซึ่งทำการเรียนรู้พารามิเตอร์ ที่จะถูกกำหนด โดยการสุ่มในขั้นตอนการเรียนรู้เริ่มต้น ความสูง (height) และน้ำหนัก (weight) ของตัวกรองจะเล็ก

กว่าปริมาตรของสารสนเทศนำเข้าแต่แต่ละตัวมีความสอดคล้องเพื่อคำนวณ ฟีดเจอร์แมพ (feature map) ที่ใช้ในเซลล์ประสาท กลางอีกนัยหนึ่งตัวกรองจะเลื่อนไปตามความกว้างและความสูงของข้อมูลนำเข้าและผลคูณจุด (dot products) ระหว่างข้อมูลนำเข้าและตัวกรองในทุก ๆ หน่วยพื้นที่ หลังจากนั้นจะปรับค่าโดยการเรียนรู้แบบแพร่กระจายย้อนกลับ จำนวนผลลัพธ์ที่ได้จากชั้นคอนโวลูชันจะเท่ากับจำนวนฟีดเจอร์แมพในชั้นของคอนโวลูชันมักจะตามด้วยฟังก์ชันกระตุ้น ชั้นตอนดังกล่าวนี้แสดงในรูปที่ 10



รูปที่ 10 ชั้นคอนโวลูชันที่ประกอบไปด้วยองค์ประกอบกลางของคอร์เนลจะถูกวางไว้เหนือเวกเตอร์อินพุตซึ่งจะถูกคำนวณและแทนที่ด้วยผลรวมถ่วงน้ำหนักของตัวเองและฟังก์ชันกระตุ้น

[ที่มา : https://www.researchgate.net/publication/285164623_An_Introduction_to_Convolutional_Neural_Networks]

2. ชั้นพูลลิ่ง (Pooling Layer)

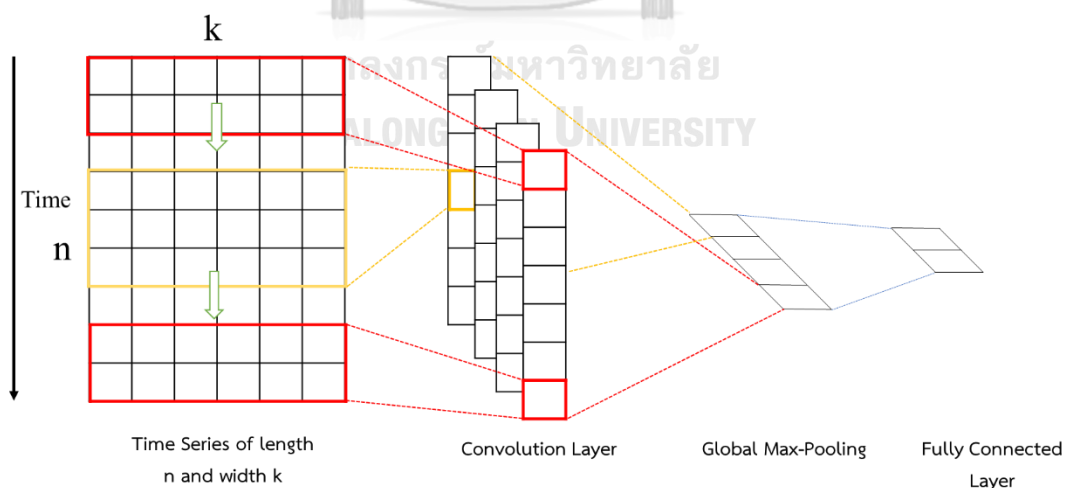
ชั้นพูลลิ่งมีหน้าที่ในการลดมิติของข้อมูลหรืออินพุต และค่อย ๆ ลดจำนวนพารามิเตอร์และความซับซ้อนในการคำนวณของแบบจำลอง โดยชั้นพูลลิ่งดำเนินการโดยฟังก์ชันกระตุ้นแมพ (activation map) กับอินพุตเวกเตอร์และทำการปรับขนาดมิติของข้อมูลโดยใช้ฟังก์ชันมากที่สุด (max) โดยทั่วไปนิยมใช้การทำพูลลิ่งแบบหาค่ามากที่สุด (max pooling) หรือการทำพูลลิ่งแบบหาค่าเฉลี่ย (average pooling)

3. ชั้นการเชื่อมโยงเต็มรูปแบบ (Fully-Connected Layer)

ชั้นที่เชื่อมต่ออย่างสมบูรณ์หรือเชื่อมโยงเต็มรูปแบบ ในเครือข่ายประสาทเทียมคือโดยชั้นหน่วยนำเข้าทั้งหมดจากชั้นหนึ่งเชื่อมต่อกับทุกหน่วยการเปิดใช้งานของชั้นถัดไป ในรูปแบบการเรียนรู้ของเครื่องส่วนใหญ่ขั้นสุดท้ายเป็นชั้นที่เชื่อมต่อเต็มซึ่งรวบรวมข้อมูลที่แยกโดยชั้นก่อนหน้าเพื่อสร้างเอาต์พุตสุดท้าย เป็นชั้นที่ใช้เวลามากที่สุดเป็นอันดับสองรองจากชั้นคอนโวลูชัน การเชื่อมโยงเต็มรูปแบบเป็นขั้นสุดท้ายของนิรอรอลโครงข่ายประสาทเทียมคอนโวลูชัน

2.1.5 โครงข่ายประสาทเทียมคอนโวลูชัน 1 มิติสำหรับอนุกรมเวลา (1-D Convolution Neural Network for Time Series)

กำหนดให้ข้อมูลอนุกรมเวลา ความยาว n และมีความกว้าง k เมื่อความยาวเรียกว่าจำนวนขั้นเวลา (number of timesteps) และ ความกว้างคือจำนวนของตัวแปรของข้อมูลอนุกรมเวลาแบบหลายตัวแปร โดยคอนโวลูชันคอร์เนลมีความกว้างเท่ากับอนุกรมเวลาเสมอในขณะที่ความยาวสามารถเป็นไปได้อย่างหลายค่า ด้วยวิธีนี้คอร์เนลจะเคลื่อนที่ไปในทิศทางเดียวจากจุดเริ่มต้นของอนุกรมเวลาไปจนถึงจุดสิ้นสุด มันจะย้ายไปทางซ้ายหรือไปทางขวาเช่นเดียวกับเมื่อใช้การบิด 2 มิติแบบปกติกับภาพ องค์ประกอบของคอร์เนลจะถูกคูณด้วยองค์ประกอบที่สอดคล้องกันของอนุกรมเวลาที่ครอบคลุม ณ จุดที่กำหนด จากนั้นผลลัพธ์ของการคูณจะถูกรวมเข้าด้วยกันและฟังก์ชันกระตุ้นแบบไม่เชิงเส้นถูกนำมาใช้ในการปรับค่า ค่าผลลัพธ์ที่ได้จะกลายเป็นองค์ประกอบของอนุกรมเวลาที่ถูกรองแบบตัวแปรเดียว จากนั้นคอร์เนลจะเลื่อนไปข้างหน้าตามลำดับเวลาเพื่อสร้างค่าถัดไปจำนวนของอนุกรมเวลาใหม่ที่ถูกรอง จะเหมือนกับจำนวนของคอนโวลูชันคอร์เนล ขึ้นอยู่กับความยาวของคอร์เนลลักษณะที่แตกต่างคุณสมบัติ คุณลักษณะ ของข้อมูลอนุกรมเวลาเวลาขณะเริ่มต้นจะถูกบันทึกในชุดอนุกรมเวลาที่ถูกรองใหม่แต่ละชุด ขั้นตอนต่อไปคือการใช้การทำพูลลิ่งแบบหาค่ามากที่สุด กับเวกเตอร์อนุกรมเวลาที่กรองแต่ละตัว ค่าที่ใหญ่ที่สุดนั้นมาจากเวกเตอร์แต่ละตัว เวกเตอร์ใหม่เกิดขึ้นจากค่าเหล่านี้และเวกเตอร์สูงสุดนี้เป็นเวกเตอร์คุณลักษณะสุดท้ายที่สามารถใช้เป็นอินพุตสำหรับ ขั้นตอนการเชื่อมโยงเต็มรูปแบบ กระบวนการทั้งหมดนี้แสดงไว้ในรูปที่ 11



รูปที่ 11 กระบวนการของโครงข่ายประสาทเทียมคอนโวลูชัน 1 มิติสำหรับอนุกรมเวลา

[ที่มา <https://towardsdatascience.com/how-to-use-convolutional-neural-networks-for-time-series-classification-56b1b0a07a57>]

2.1.6 ฟังก์ชันกระตุ้น (Activation Function) [6]

สารสนเทศส่งออกไปยังแต่ละบัพจะมีการใช้ฟังก์ชันกระตุ้น $f(x)$ ที่ช่วยในการคำนวณค่าถ่วงน้ำหนักของให้โครงข่ายประสาทเทียม เพื่อให้สามารถแก้ไขปัญหาที่ไม่เป็นเชิงเส้นได้หรือสามารถเรียนรู้ความหลากหลายของรูปแบบสารสนเทศมากขึ้น ซึ่งฟังก์ชันกระตุ้นที่นิยมมีดังรายละเอียดตามตารางที่ 2

ตารางที่ 2 ประเภทของฟังก์ชันกระตุ้น ค่าผลลัพธ์ และสมการ

ฟังก์ชัน	ค่าผลลัพธ์	สมการของฟังก์ชัน	
ฟังก์ชันซิกมอยด์ (sigmoid function)	0 ถึง 1	$\sigma(x) = \frac{1}{1+e^{-x}}$	(2.13)
ฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ (hyperbolic tangent function)	-1 ถึง 1	$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	(2.14)
ฟังก์ชันค่าสูงสุดอย่างอ่อน (softmax function)	0 ถึง 1	$f(x) = \frac{e^{x_j}}{\sum_{i=1}^K e^{x_i}}$	(2.15)
ฟังก์ชันเรคตีไฟด์เชิงเส้น (Rectified Linear Unit Function - ReLU)	$+$ หรือ 0	$f(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases}$	(2.16)
ฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้น (Exponential Linear Unit Function - ELU)	$(-\infty, \infty)$	$f(\alpha, x) = \begin{cases} \alpha (e^x - 1), & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases}$	(2.17)

สำหรับค่าฟังก์ชันค่าสูงสุดอย่างอ่อนในชั้นของโครงข่ายประสาทเทียมที่ต้องการคำนวณมีผลลัพธ์ทั้งหมด K ตัว ค่าของผลลัพธ์จะแทนด้วยสัญลักษณ์ x กล่าวคือ ฟังก์ชันค่าสูงสุดอย่างอ่อนของผลลัพธ์ที่ j หรือ แทนด้วย สัญลักษณ์ f_j

2.1.7 การหาค่าเหมาะสมที่สุด (Optimization) [11]

การเรียนรู้ด้วยสโตแคสติกเกรเดียนต์เดสเซนท์ (Stochastic Gradient Descent - SGD) ถูกนำมาใช้เพื่อเป็นฟังก์ชันจุดประสงค์ ในการลดค่าของฟังก์ชันต้นทุนให้น้อยที่สุดด้วยการปรับน้ำหนักของเส้นเชื่อมในโครงข่ายประสาทเทียม โดยมีการปรับค่าน้ำหนัก โดยนำค่าน้ำหนักก่อนหน้ามาคำนวณตามสมการที่ 2.18

$$w_t = w_{t-1} - \alpha \frac{\partial J_t}{\partial w} \quad (2.18)$$

โดยที่ w คือ พารามิเตอร์น้ำหนักที่ต้องการจะปรับค่า

α คือ อัตราการเรียนรู้

$\frac{\partial J_t}{\partial w}$ คือ เกรเดียนต์ของฟังก์ชันต้นทุนเทียบกับ w

การปรับน้ำหนักให้เข้าสู่จุดเหมาะสมและเพื่อลดโอกาสการอิงอยู่กับจุดที่ดีที่สุดท้องถิ่น (local optima) ซึ่งเป็นค่าที่เหมาะสมที่สุดในช่วง ๆ หนึ่งเท่านั้น ซึ่งไม่ใช่ทั้งหมดอย่างทั้งหมดสามารถใช้โมเมนตัม (momentum) เพื่อลดปัญหาดังกล่าว ตามสมการที่ 2.19 และ 2.20

$$v_t = \gamma v_{t-1} + \alpha \frac{\partial J_t}{\partial w} \quad (2.19)$$

$$w_t = w_{t-1} - v_t \quad (2.20)$$

โดยที่ v คือ ค่าความเร็วอันมีการปรับค่าพร้อมกับ w

γ คือ ค่าสัมประสิทธิ์ของโมเมนตัม (momentum coefficient)

2.1.8 การวิเคราะห์องค์ประกอบสำคัญ (Principle Component Analysis - PCA) [12]

การวิเคราะห์องค์ประกอบสำคัญ เป็นเทคนิคหนึ่งที่ใช้ในการลดตัวแปร p ตัวโดยสร้างตัวแปรใหม่ที่เรียกว่า องค์ประกอบสำคัญ หรือ ปัจจัย k ตัวที่มีความสัมพันธ์เชิงเส้นของตัวแปร p ตัว ดังนั้นจำนวนองค์ประกอบสำคัญจะต้องมีจำนวนไม่เกินจำนวนตัวแปร กล่าวคือ $k \leq p$

ขั้นตอนการคำนวณของวิธีการวิเคราะห์องค์ประกอบสำคัญ

ให้ $X' = (x_1, x_2, \dots, x_p)$ ซึ่งมีเมทริกซ์ความแปรปรวนร่วม Σ ที่มีค่าเฉพาะ (eigenvalue) และเวกเตอร์เฉพาะ (eigenvector) เป็น (λ_p, w_p) โดยที่ค่าเฉพาะ $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ และเวกเตอร์เฉพาะ $W' = (w_1, w_2, \dots, w_p)$ สร้างสมการที่ใช้ประมาณค่าองค์ประกอบสำคัญหรือปัจจัยตัวที่ i (PC_i) เมื่อ $i = 1, \dots, p$ ดังนี้

$$\begin{aligned} PC_1 &= w'_1 X = w_{11}x_1 + w_{12}x_2 + \dots + w_{1p}x_p \\ PC_2 &= w'_2 X = w_{21}x_1 + w_{22}x_2 + \dots + w_{2p}x_p \\ &\vdots \\ PC_p &= w'_p X = w_{p1}x_1 + w_{p2}x_2 + \dots + w_{pp}x_p \end{aligned} \quad (2.21)$$

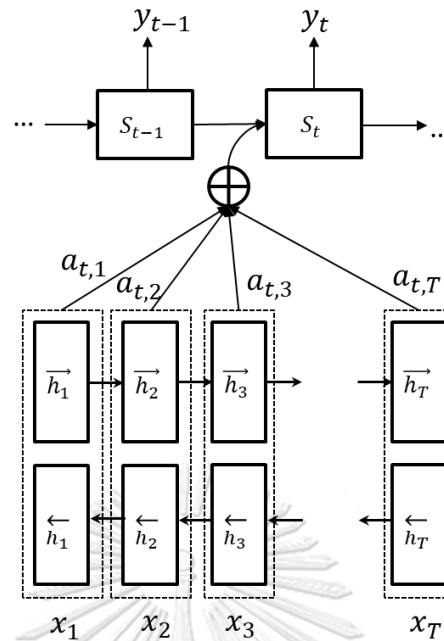
โดยเริ่มจาก (PC_1) ที่มีค่าความแปรปรวนสูงสุดไปจนถึง (PC_p) ที่มีค่าความแปรปรวนต่ำสุด หรือ $\text{Var}(PC_1) \geq \text{Var}(PC_2) \geq \dots \geq \text{Var}(PC_p)$ จะได้ว่า

$$\begin{aligned} \sum_{i=1}^p \text{Var}(x_i) &= \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \lambda_1 + \lambda_2 + \dots + \lambda_p \\ &= \sum_{i=1}^p \text{Var}(PC_i) \end{aligned} \tag{2.22}$$

$$\therefore PC_k \text{ มีสัดส่วนของค่าแปรปรวน} = \frac{\lambda_k}{\sum_{i=1}^p \lambda_i}; k=1, 2, \dots, p$$

2.1.9 กลไกจุดสนใจ (Attention Mechanism)

ในปี ค.ศ. 2014 Dzmitry B. และคณะ [13] ได้นำเสนอกลไกจุดสนใจเพื่อใช้ในการแปลภาษาผ่านเครื่องด้วยโครงข่ายประสาทเทียมการแปล (neural machine translation) โดยมีจุดมุ่งหมายเพื่อการสร้างโครงข่ายประสาทเทียมเดี่ยวที่สามารถปรับแต่งร่วมกันเพื่อเพิ่มประสิทธิภาพการแปลให้มากขึ้น ซึ่งแต่เดิมนั้น ด้วยการใช้เวกเตอร์ที่มีความยาวคงที่นั้นเป็นปัญหาคอขวดในการปรับปรุงประสิทธิภาพของสถาปัตยกรรมตัวเข้ารหัส - ถอดรหัสพื้นฐาน (basic encoder-decoder) และได้เสนอแบบจำลองการค้นหาโดยอัตโนมัติ สำหรับส่วนต่าง ๆ ของประโยคต้นฉบับ ด้วยวิธีการใหม่นี้ทำให้ได้ประสิทธิภาพการแปลที่เทียบเท่ากับระบบวลีที่มีอยู่ในงานการแปลภาษาอังกฤษเป็นภาษาฝรั่งเศสที่ทันสมัยที่สุดในขณะนั้น โดยหลักการของกลไกจะสร้างค่าความสนใจ (attention mechanism) ให้กับสมาชิกแต่ละตัวในลำดับข้อมูล แล้วนำค่าที่ได้คูณกลับไปหาสมาชิกตัวนั้น ๆ เพื่อสร้างเวกเตอร์ผลลัพธ์ที่จะนำไปใช้ในโครงข่ายประสาทเทียมต่อไป ค่าความสนใจนี้จะเปรียบเสมือนค่าน้ำหนักที่ถ่วงข้อมูลแต่ละตัวในลำดับข้อมูลนั้น ถ้าข้อมูลไหนควรสนใจมากค่าความสนใจจะมากด้วย โดยสถาปัตยกรรมนี้ประกอบด้วย โครงข่ายประสาทเทียมแบบวกกลับแบบสองทิศทาง (Bidirectional Recurrent Neural Network – BiRNN) ที่เป็นตัวเข้ารหัสและตัวถอดรหัสที่จำลองการค้นหาผ่านประโยคต้นฉบับระหว่างตัวถอดรหัสการแปล โดยกำหนดความน่าจะเป็นตามเงื่อนไข (conditional probability) ตามสมการที่ 2.23



รูปที่ 2.8 แบบจำลองกลไกความสนใจในการสร้างคำเป้าหมายที่ y_t ของประโยคต้นฉบับ (x_1, x_2, \dots, x_T)

$$p(y_i | y_i, \dots, y_{i-1}, x) = g(y_{i-1}, s_i, c_i) \quad (2.23)$$

เมื่อ S_i คือ สถานะที่ซ่อนอยู่ของอาร์เอ็นเอ็น ณ เวลา i คำนวณได้จาก สมการ 2.24

$$s_i = f(s_{i-1}, y_{i-1}, c_i) \quad (2.24)$$

y_i คือ เวกเตอร์ลำดับที่ i ของประโยคปลายทาง

f คือ โครงข่ายประสาทเทียมที่เลือกใช้ เช่น แอลเอสทีเอ็ม หรือ หน่วยประมวลผลวนกลับ (Gate Recurrent Unit – GRU)

c_i คือ เวกเตอร์บริบทของโน้ตซ์ลำดับที่ i จากประโยคปลายทาง คำนวณจาก สมการที่ 2.26

T_x คือ จำนวนโน้ตซ์ของประโยคต้นทาง x

h_j คือ เวกเตอร์ลำดับที่ j จากประโยคต้นทางได้จากชั้นซ่อนของ f

a_{ij} คือ น้ำหนักหรือค่าความสนใจของแต่ละโน้ตซ์ลำดับที่ i จากประโยคปลายทางเมื่อเทียบกับโน้ตซ์ลำดับที่ j หรือ h_j ของประโยคต้นฉบับ คำนวณจาก สมการที่ 2.27

att คือ ฟังก์ชันสำหรับคำนวณค่าความสนใจ

ความน่าจะเป็นตามเงื่อนไขอยู่บนเวกเตอร์บริบทเอกลักษณ์ c_i สำหรับ คำเป้าหมาย y_i เวกเตอร์บริบท c_i ขึ้นกับลำดับโนเตชัน (annotation) (h_1, \dots, h_{T_x}) ไปยังตัวเข้ารหัสที่จับคู่กับประโยคนำเข้าแต่ละโนเตชัน h_i ประกอบด้วยสารสนเทศเกี่ยวกับลำดับนำเข้าทั้งหมดกับการมุ่งเน้นในส่วนรอบข้างของคำที่ i ของลำดับนำเข้า เวกเตอร์บริบท c_i ถูกคำนวณเพื่อเป็นผลรวมน้ำหนักของโนเตชัน h_i

$$y_i = f(y_{i-1}, c_i) \quad (2.25)$$

$$c_i = \sum_{j=1}^{T_x} a_{ij} h_j \quad (2.26)$$

$$a_{ij} = \frac{e^{e_{ij}}}{\sum_{k=1}^{T_x} e^{e_{ik}}} \quad (2.27)$$

$$\text{หรือ } a_{ij} = \text{softmax}(\text{att}(s_{i-1}, h_j)) \quad (2.28)$$

$$\text{เมื่อ } e_{ij} = \text{att}(s_{i-1}, h_j) \quad (2.29)$$

e_{ij} คือ แบบจำลองการจัดตำแหน่งให้คะแนนความดีของตัวนำเขารอบข้างตำแหน่งที่ j และตัวนำออกตรงกับตำแหน่งที่ i คะแนนถูกอิงจากสถานะซ่อนอาร์เอ็นเอ็น s_{i-1} ก่อนการได้มาซึ่ง y_i ในสมการที่ 2.25 และโนเตชันที่ h_j ของประโยคนำเข้า เมื่อ α_{ij} เป็นความน่าจะเป็นความน่าจะเป็นที่คำเป้าหมาย y_i เป็นแนวทางเดียวกันหรือแปลจาก คำต้นทางทาง x_i ต่อจากนั้น c_i (เวกเตอร์บริบทที่ i คือโนเตชันที่ฟังประสงค์เหนือกว่าโนเตชันอื่น ๆ ด้วยความน่าจะเป็น α_{ij} และด้วยความน่าจะเป็น α_{ij} หรือ พลังงานที่เกี่ยวข้อง e_{ij} สะท้อนให้เห็นความสำคัญของโนเตชัน h_j ในส่วนของสถานะ s_{i-1} ก่อนหน้าในการตัดสินใจสถานะต่อไป s_i และการสร้าง y_i ด้วยวิธีนี้ เป็นการสร้างกลไกของความสนใจในตัวถอดรหัส ซึ่งตัวถอดรหัสจะจัดสนใจส่วนของประโยคต้นฉบับเพื่อให้ความสนใจด้วยการปล่อยให้ตัวถอดรหัสมีกลไกความสนใจ เราสามารถลดภาระการถอดรหัสสารสนเทศทั้งหมดในประโยคต้นฉบับเป็นเวกเตอร์ความยาวคงที่ให้แก่ตัวเข้ารหัส ด้วยวิธีการนี้ข้อมูลสามารถแพร่กระจายไปตามลำดับของโนเตชันซึ่งสามารถเลือกได้โดยตัวถอดรหัส

2.1.10 ไลต์แกรนเดียร์นบูตติงแมทชีน (Light Gradient Boosting Machine - LightGBM) [14]

อัลกอริทึมไลต์แกรนเดียร์นบูตติงแมทชีน ถูกพัฒนามาจากอัลกอริทึมต้นไม้การตัดสินใจแบบแกรนเดียร์นบูตติง (Gradient Boosting Decision Tree - GBDT) ที่ได้รับความนิยมและมีการใช้งานที่มีประสิทธิภาพ เช่น XGBoost และต้นไม้การตัดสินใจแบบแกรนเดียร์นบูตติงเรสชัน (Gradient Boosted Regression Tree - GBRT) โดยถูกปรับปรุงเพิ่มประสิทธิภาพและความสามารถประมวลผลข้อมูลที่มีมิติหรือมีคุณลักษณะที่มาก ๆ และเป็นข้อมูลขนาดใหญ่ โดยมีการสุ่มตัวอย่างแบบข้างเดียวแบบแกรนเดียร์น (Gradient-based on One-Side Sampling - GOSS) จีโอเอสเอส เก็บอินสแตนซ์ทั้งหมดด้วยแกรนเดียร์นขนาดใหญ่และทำการสุ่มตัวอย่างแบบสุ่มบนอินสแตนซ์ที่มีแกรนเดียร์นเล็กน้อย เพื่อลดขนาดของข้อมูลกระจายข้อมูลเมื่อคำนวณข้อมูลที่ได้รับ จีโอเอสเอส จะแนะนำตัวคุณค่งสำหรับอินสแตนซ์ข้อมูลที่มีแกรนเดียร์นขนาดเล็ก โดยเฉพาะ จีโอเอสเอสจะเรียงลำดับอินสแตนซ์ของข้อมูลเป็นอันดับแรกตามค่าสัมบูรณ์ของแกรนเดียร์นและเลือกอินสแตนซ์สูงสุด $\times 100\%$ จากนั้นสุ่มตัวอย่าง $b \times 100\%$ อินสแตนซ์จากส่วนที่เหลือของข้อมูล หลังจากนั้น จีโอเอสเอสจะขยายข้อมูลตัวอย่างด้วยแกรนเดียร์นขนาดเล็กด้วยค่าค่งที่ $(1 - a) / b$ เมื่อคำนวณการได้รับข้อมูล ด้วยการกระทำเช่นนี้เราให้ความสำคัญกับอินสแตนซ์ที่ผ่านการสอน โดยไม่เปลี่ยนแปลงการกระจายข้อมูลดั้งเดิมมาก จีบีดีที ใช้ต้นไม้ตัดสินใจเพื่อเรียนรู้ฟังก์ชันจากพื้นที่อินพุต \mathcal{X}^s ถึงพื้นที่แกรนเดียร์ \mathcal{G} โดยอนุมาณว่าเรามีชุดข้อมูลการสอนที่มีข้อมูล n อินสแตนซ์ $\{x_1, \dots, x_n\}$ เมื่อ x_i คือเวกเตอร์ที่มีมิติของคุณลักษณะ s ในพื้นที่ \mathcal{X}^s ในแต่ละรอบของ แกรนเดียร์นบูตติง แกรนเดียร์ที่เป็นลบของฟังก์ชันสูญเสียที่เกี่ยวกับการส่งออกของรูปแบบจะแสดงเป็น $\{g_1, \dots, g_n\}$ แบบจำลองต้นไม้ตัดสินใจจะแยกแต่ละโหนดที่คุณลักษณะที่มีความหมายสำคัญมากที่สุด สำหรับ จีบีดีที ข้อมูลที่ได้รับมักจะวัดจากความแปรปรวนหลังจากแยกซึ่งถูกกำหนดไว้ด้านล่างเมื่อ เรากำหนดให้ O เป็นชุดข้อมูลฝึกในโหนดที่ค่งที่ของต้นไม้ตัดสินใจ ค่าความแปรปรวนที่ได้ของการแยกคุณลักษณะที่ j ณ จุดที่ d สำหรับโหนดที่ถูกกำหนดดังนี้

$$V_{j|O}(d) = \frac{1}{n_o} \left(\frac{(\sum_{\{x_i \in O: x_{ij} \leq d\}} g_i)^2}{n_{l|O}^j(d)} + \frac{(\sum_{\{x_i \in O: x_{ij} > d\}} g_i)^2}{n_{r|O}^j(d)} \right), \quad (2.30)$$

$$\text{เมื่อ } n_o = \sum I[x_i \in O], n_{l|O}^j = \sum I[x_i \in O], n_{r|O}^j(d) = \sum I[x_i \in O: x_{ij} \leq d] \text{ and } n_{r|O}^j(d) = \sum I[x_i \in O: x_{ij} > d]$$

สำหรับคุณลักษณะที่ j อัลกอริทึมต้นไม้ตัดสินใจเลือก $d_j^* = \operatorname{argmax}_d V_j(d)$ และคำนวณอัตราการขยายที่มากที่สุด $V_j(d_j^*)$ หลังจากนั้น ข้อมูลถูกแยกตามคุณลักษณะที่ j^* ที่จุด d_{j^*} ไปยังโหนดลูกทั้งทางซ้ายและขวา วิธีการจีโอเอสเอสจะทำการเรียงลำดับอินสแตนซ์ตามค่าแแกรเดียร์นสมบูรณ์จากมากไปหาน้อย จากนั้นเก็บค่า top- $a \times 100\%$ อินสแตนซ์กับแแกรเดียร์นที่มากที่สุด และได้อินสแตนซ์เซต A และสำหรับเซตที่เหลือ A^c ประกอบด้วย $(1 - a) \times 100\%$ อินสแตนซ์กับค่าแแกรเดียร์นที่น้อยที่สุด จากนั้นสุ่มเซต B ที่มีขนาด $b \times |A^c|$ และในขั้นตอนสุดท้ายเราแยกค่าประมาณอัตราการขยายแปรปรวน $\tilde{V}_j(d)$ บนเซตย่อย $A \cup B$

$$\tilde{V}_j(d) = \frac{1}{n} \left(\frac{(\sum_{x_i \in A_l} g_i + \frac{1-a}{b} \sum_{x_i \in B_l} g_i)^2}{n_l^j(d)} + \frac{(\sum_{x_i \in A_r} g_i + \frac{1-a}{b} \sum_{x_i \in B_r} g_i)^2}{n_r^j(d)} \right), \quad (2.31)$$

เมื่อ

$$A_l = \{x_i \in A: x_{ij} \leq d\}, A_r = \{x_i \in A: x_{ij} > d\}, B_l = \{x_i \in B: x_{ij} \leq d\},$$

$$B_r = \{x_i \in B: x_{ij} > d\},$$

และสัมประสิทธิ์ $\frac{1-a}{b}$ ถูกใช้ทำการปรับเรียงผลรวมของแแกรเดียร์บน B กลับไปยังขนาดของ A^c

จุฬาลงกรณ์มหาวิทยาลัย

2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการบำรุงรักษาเชิงพยากรณ์ แบ่งเป็น 2 กลุ่มด้วยกัน คือ 1) กลุ่มการประยุกต์การสกัดคุณลักษณะในการบำรุงรักษาเชิงพยากรณ์ และ 2) กลุ่มการประยุกต์เทคนิคปัญญาประดิษฐ์ ในการบำรุงรักษาเชิงพยากรณ์[4]

1) กลุ่มการประยุกต์การสกัดคุณลักษณะในการบำรุงรักษาเชิงพยากรณ์ คือ การวิจัยและการประยุกต์ใช้การสกัดคุณลักษณะหรือการวิเคราะห์สัญญาณ โดยแบ่ง เป็น 3 โดเมน ได้แก่ โดเมนเวลา (time domain) โดเมนความถี่ (frequency domain) และโดเมนความถี่-เวลา (time-frequency domain) ซึ่งสิ่งสำคัญของวิธีทางโดเมนเวลาคือการวิเคราะห์รูปแบบของคลื่น ส่วนในด้านวิเคราะห์ความถี่ สัญญาณถูกวิเคราะห์จากความถี่ผ่านการเปลี่ยนแปลงบางอย่าง[15] ซึ่งสามารถแสดงให้เห็นถึงการกระจายของความถี่และตัวกรองสัญญาณรบกวนได้อย่างสะดวกในหลาย ๆ

กรณีที่มีข้อมูลมีความเฉพาะที่สุดได้ถูกซ่อนในเนื้อหาความถี่ของสัญญาณ ดังนั้นเทคนิคการประมวลผลสัญญาณในโดเมนเวลา ความถี่ จึงถูกนำมาใช้อย่างแพร่หลายทั้งการวินิจฉัยและทำนาย

ปี ค.ศ. 2013 Jafar Z. และคณะ [16] ได้เสนอวิธีการใช้โครงข่ายประสาทเทียมตรวจจับข้อบกพร่องของการแบริงของมอเตอร์เหนี่ยวนำ ผ่านสัญญาณของการสั่นสะเทือนโดยการออกแบบตัวกรองอัจฉริยะ โดยมีโครงข่ายประสาทเทียมกรององค์ประกอบที่ไม่ใช่ข้อบกพร่องของการแบริงออก (Removing Non-bearing Fault Component - RNFC) และใช้โครงข่ายประสาทเทียมที่สองในการจดจำรูปแบบสำหรับการจำแนกประเภทซึ่งเป็นวิธีการวิเคราะห์คุณลักษณะตามแบบเวลา ที่ต้องการการคำนวณที่ไม่ซับซ้อนมาก วิธีนี้ได้ผลลัพธ์ที่มีความแม่นยำสูงถึง 100% เมื่อเปรียบเทียบกับการใช้ระบบอนุมานฟัซซีโครงข่ายปรับตัวได้ (Adaptive Neuro-Fuzzy Inference System - ANFIS) ซึ่งเป็นวิธีที่ใช้กันโดยทั่วไป การใช้ RNFC มีความแม่นยำมากกว่าถึง 84%-30%

ปี ค.ศ. 2015 Patrick J. [17] ศึกษาการใช้การเรียนรู้ของเครื่อง เพื่อการตรวจจับประเภทของความล้มเหลวและการบำรุงรักษาเชิงพยากรณ์ ซึ่งแสดงให้เห็นถึงปัญหาต่าง ๆ ของการสกัดลักษณะสำคัญ การเลือกคุณลักษณะ และการตัดสินใจของคุณสมบัติทางกายภาพขององค์ประกอบที่มีความเกี่ยวข้องมากที่สุด ที่ยังไม่มีกรอกแบบเทคนิคการเลือกคุณลักษณะที่รู้จักกันอย่างแพร่หลาย การตัดสินใจที่ต้องคำนึงถึงความเสื่อมขององค์ประกอบทางกายภาพ ตลอดจนการออกแบบการเรียนรู้ของเครื่องแบบกระจาย (Distributed machine learning approach) ที่จะช่วยลดเวลาในกระบวนการตัดสินใจลงได้ อัลกอริทึมในการเรียนรู้ควรสามารถตรวจจับประเภทข้อบกพร่องและเรียนรู้ด้วยระบบเช่นเดียวกับการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning - RL)

2) กลุ่มการประยุกต์เทคนิคปัญญาประดิษฐ์ในการบำรุงรักษาเชิงพยากรณ์ โดยโครงข่ายประสาทเทียม (Artificial Neural Networks - ANNs) เป็นวิธีการที่ได้รับความนิยมมากที่สุด ANNs หลากหลายประเภทได้ถูกประยุกต์ใช้ เช่น โครงข่ายประสาทเทียมกระบวนการส่งค่าย้อนกลับ (Back Propagation Neural Network - BPNN)

ปี ค.ศ. 2017 Vimala M. และคณะ [15] แสดงการเปรียบเทียบ 10 อัลกอริทึม เพื่อการทำนายอายุการใช้งานที่เหลืออยู่ของเทอร์โบแฟน โดยใช้ชุดข้อมูลที่ถูกจัดทำโดยศูนย์ความเป็นเลิศทางด้านการทำนาย (The Prognostics Center of Excellence หรือ PCoE) ของศูนย์วิจัยนาซาเอมส์ (NASA Ames Research Center) โดยงานวิจัยดังกล่าวพบว่า อัลกอริทึมที่มีความแม่นยำสูงสุดคือ ป่าแบบสุ่ม

จากการศึกษาของ Zhe L. [3] พบว่าแบบจำลองการเรียนรู้ของเครื่อง เช่น โครงข่ายประสาทเทียมความน่าจะเป็น (BPNN) ซัพพอร์ตเวกเตอร์แมชชีน (SVM) โครงข่ายประสาทเทียมแบบเรเดียลเบซิสฟังก์ชัน (Radial Basis Function Network - RBNN) และ เซลฟออร์กาไนซิงแมพ (Self-Organizing Maps - SOM) มีความสามารถที่จะค้นพบข้อมูลความบกพร่องในขณะปัจจุบัน

และสามารถรองรับปัญหาที่ข้อมูลมีขนาดไม่ใหญ่มาก โดยที่คุณลักษณะสามารถที่จะถูกสกัดออกมาได้อย่างเหมาะสม อย่างไรก็ตามเมื่อวัตถุประสงค์คือข้อมูลดิบขนาดใหญ่ และเป้าหมายเพื่อประเมินความสัมพันธ์และการทำนายความล้มเหลวของอุปกรณ์ หลาย ๆ ระบบพบความท้าทายหรือปัญหาขอจุดที่การเรียนรู้เชิงลึกถูกนำเสนอเพื่อแก้ไขปัญหา

ปี ค.ศ. 2017 AbdElRahman E. และคณะ [18] ได้พัฒนาโครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network - RNN) สถาปัตยกรรมอาร์เอ็นเอ็นแบบแอลเอสทีเอ็ม เพื่อการทำนายการสั่นสะเทือนของอากาศยานและใช้วิธีข่ายประสาทวิวัฒนาการ (neuroevolution method) ตามขั้นตอนวิธีหาค่าเหมาะที่สุดด้วยระบบอาณานิคม (Ant Colony Optimization - ACO) โดยแบบจำลองได้ถูกสอนด้วยข้อมูลขนาดใหญ่จากสายการบินที่ได้รับผลกระทบจากปัญหาการสั่นสะเทือนที่มากเกินไป และแสดงให้เห็นการปรับปรุงที่ดีขึ้นโดยลดความคลาดเคลื่อนของการทำนายจาก 5.51% ถึง 4.17% และสามารถลดเส้นนำหน้จาก 21,170 เส้น เหลือ 11,810 เส้น

Liu Y. และคณะ [19] ได้นำเสนอการทำนายเวลาที่ขึ้นอยู่กับเครื่องมือการตัด (cutting tools) ของกังหันแก๊ส (gas turbine) โดยการใช้ความคล้ายของดัชนีสุขภาพ (health index similarity) โดยได้จากการทำการวิเคราะห์องค์ประกอบสำคัญ จากหลาย ๆ คุณลักษณะให้เหลือเพียงคุณลักษณะเดียวที่สามารถบ่งบอกถึงความเสื่อมประสิทธิภาพของเครื่องมือการตัดได้ จากนั้นใช้ความคล้ายตามระยะห่าง (distance similarity) มาตรฐานระยะทางแบบยูคลิเดียน (Euclidean distance) และ ความคล้ายตามระยะห่างเชิงพื้นที่ (spatial distance similarity) จากการหาค่าโคไซน์ความคล้าย ได้ค่า MAE, RMSE เท่ากับ 0.5681 และ 0.9333 ซึ่งวัดผลโดยการเปรียบเทียบกับ ซัพพอร์ตเวกเตอร์แมชชีน แบบค่ากำลังสองน้อยที่สุด (Least Square Support Vector Machine - LS-SVR) เนื่องจากจำนวนข้อมูลมีปริมาณไม่มาก ซึ่งได้ค่า MAE, RMSE เท่ากับ 2.0149 และ 2.4234 ทั้งนี้ยังมีข้อจำกัด คือ ข้อมูลมีไม่มากพอ อีกทั้งมีความไม่แน่นอนของการทำนายอายุการใช้งานที่เหลืออยู่ และยังไม่ได้ถูกประเมินกับข้อมูลชุดอื่น

ปี พ.ศ. 2559 มนัสกานต์ เสน่หา และ พีรพล เวทีกุล [7] ได้นำเสนอแบบจำลองโครงข่ายคอนโวลูชันเชิงลึกเพื่อการทำนายอายุการใช้งานคงเหลือของเครื่องจักร โดยเพิ่มประสิทธิภาพให้แก่แบบจำลองโดยใช้เทคนิคต่าง ๆ เช่น การทำเรกูลาร์ไลเซชันแบบแอลสอง การป้องกันการอิงกับข้อมูลมากเกินไปด้วยการดรอปเอาต์ การลดความซับซ้อนของแบบจำลองด้วย และใช้การหาค่าที่เหมาะสมที่สุดด้วยเกรเดียนต์แบบปรับตัวได้ ซึ่งเทคนิคดังกล่าวส่งผลให้แบบจำลองที่นำเสนอมีประสิทธิภาพดีขึ้นกว่าเดิมโดยเทียบกับโครงสร้างแบบเดิมได้แก่ LeNet-5 VGG-16 และ AlexNet

ปี ค.ศ. 2017 Rui Z. ในและคณะ [20] ได้นำเสนอการใช้แบบจำลองโครงข่ายคอนโวลูชันแอลเอสทีเอ็มแบบสองทิศทาง ได้ผลดีที่สุดเมื่อเทียบกับแบบจำลองอื่น ๆ อีกทั้งมีประสิทธิภาพใน

การตรวจสอบเครื่องจักรตามเวลาจริง ปี ค.ศ. 2014 Dzmitry B. และคณะ [13] ได้พัฒนากลไกจุดสนใจขึ้นเป็นครั้งแรก เพื่อใช้ในการแปลภาษาผ่านเครื่องด้วยโครงข่ายประสาทเทียม (Neural machine translation) ต่อมาได้รับความนิยมนำไปใช้งานด้านต่าง ๆ ที่ใช้โครงข่ายประสาทเทียมเพื่อประมวลผลข้อมูลแบบลำดับต่อลำดับ (Sequence-to-sequence model)

ปี ค.ศ.2018 Chuxu Z. และคณะ [21] ได้นำเสนอโครงข่ายประสาทเทียมเชิงลึกสำหรับการตรวจจับและวินิจฉัยความผิดปกติของอุปกรณ์ในโรงไฟฟ้าในข้อมูลอนุกรมเวลาหลายตัวแปรโดยได้ประยุกต์ใช้กลไกจุดสนใจและโครงข่ายคอนโวลูชันหน่วยความจำระยะสั้นแบบยาวมีประสิทธิภาพดีที่สุดในเมื่อเทียบกับแบบจำลองอื่นในงานวิจัยที่เกี่ยวข้อง

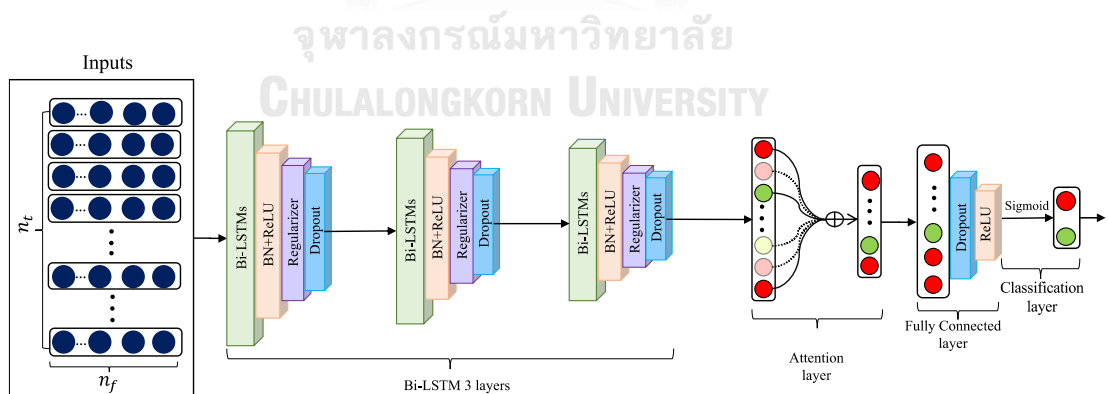
จากงานวิจัยที่ได้ศึกษามาข้างต้น ผู้วิจัยเล็งเห็นว่าควรนำหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง ซึ่งมีประสิทธิภาพเป็นอย่างมากในการเรียนรู้ข้อมูลแบบอนุกรมเวลาและเป็นข้อมูลขนาดใหญ่และโครงข่ายคอนโวลูชันในการเสริมประสิทธิภาพในการทำนาย และเสนอการประยุกต์ใช้การวิเคราะห์องค์ประกอบสำคัญในการลดมิติของข้อมูล ผสานกับการเลือกคุณลักษณะที่จะมีอิทธิพลต่อความล้มเหลวของเครื่อง เพื่อแก้ปัญหาการใช้เวลานานในการฝึกแบบจำลองที่ต้องใช้ข้อมูลขนาดใหญ่ ผู้วิจัยจึงนำเสนอเทคนิคการลดปริมาณข้อมูลโดยการสกัดลักษณะสำคัญจากฟังก์ชันทางสถิติซึ่งจะทำให้เกิดคุณลักษณะใหม่มากขึ้นและใช้กลไกจุดสนใจเพื่อให้แบบจำลองเน้นเอาคุณลักษณะที่มีความน่าสนใจ ลดเวลาในการสอนและเพิ่มประสิทธิภาพให้แก่แบบจำลองโดยจะแสดงการเปรียบเทียบประสิทธิภาพกับหน่วยความจำระยะสั้นแบบยาว อัลกอริทึมป่าแบบสุ่ม อัลกอริทึมสมการถดถอยลอจิสติก โครงข่ายประสาทเทียม และเอ็กซ์ทรีมเกรเดียนต์บูตติ้ง (Extreme Gradient Booting – xGBoot) [22] และ ไลต์จีบีเอ็ม LightGBM [14]

บทที่ 3

แบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าและเทคนิคการเพิ่มประสิทธิภาพโดยการสกัดลักษณะสำคัญร่วมกับเทคนิคการลดข้อมูล

3.1 แบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้า

งานวิจัยนี้นำเสนอการสร้างแบบจำลองโดยใช้การเรียนรู้เชิงลึกที่มีวัตถุประสงค์เพื่อทำนายว่าเครื่องรีโมทคอมเพรสเซอร์จะเกิดเหตุล้มเหลวภายใน 1 วันล่วงหน้าหรือไม่ โดยอ้างอิงแนวทางการทำวิจัยเบื้องต้นจาก [13-18] โดยใช้การจำแนกประเภทไบนารี และประยุกต์ใช้แนวคิดการเพิ่มประสิทธิภาพแบบจำลองจากงานวิจัยของ Fazle Karim และคณะ [8] ซึ่งผู้วิจัยเลือกใช้เฉพาะส่วนแนวคิดการนำโครงข่ายหน่วยความจำระยะสั้นแบบยาว ซึ่งในแต่ละชั้นจะต่อเชื่อมด้วยเทคนิคแบทชอนอร์มัลไลเซชัน ฟังก์ชันกระตุ้นรีกิริเรอ (Leaky ReLU) การทำเรกูลาร์ไรเซชันแบบแอลสอง และเทคนิคการปรับเหมาะสมเกินไป (Overfitting) โดยการลดจำนวน(Drop-out) ร่วมกับกลไกจุดสนใจที่ช่วยเสริมประสิทธิภาพ ตามด้วยชั้นการเชื่อมโยงเต็มรูปแบบและชั้นการจำแนกโดยใช้ฟังก์ชันซิกมอยด์ ดังรูปที่ 12 และจากการศึกษาของงานวิจัยที่เกี่ยวข้องเพิ่มเติม การปรับปรุงโครงสร้างแบบจำลองโดยการต่อขนานกับโครงข่ายประสาทเทียมคอนโวลูชันที่มีการทำพีเจอร์แมพช่วยลดมิติและสกัดคุณลักษณะสำคัญร่วมกับกลไกจุดสนใจ แล้วจึงรวมผลลัพธ์เข้ากับค่านำออกจากจุดสนใจที่ออกมาจากและค่านำออกจากโครงข่ายคอนโวลูชัน ตามด้วยชั้นการเชื่อมโยงเต็มรูปแบบและชั้นการจำแนกโดยใช้ฟังก์ชันซิกมอยด์ โดยมีโครงสร้างของแบบจำลองดังรูปที่ 18

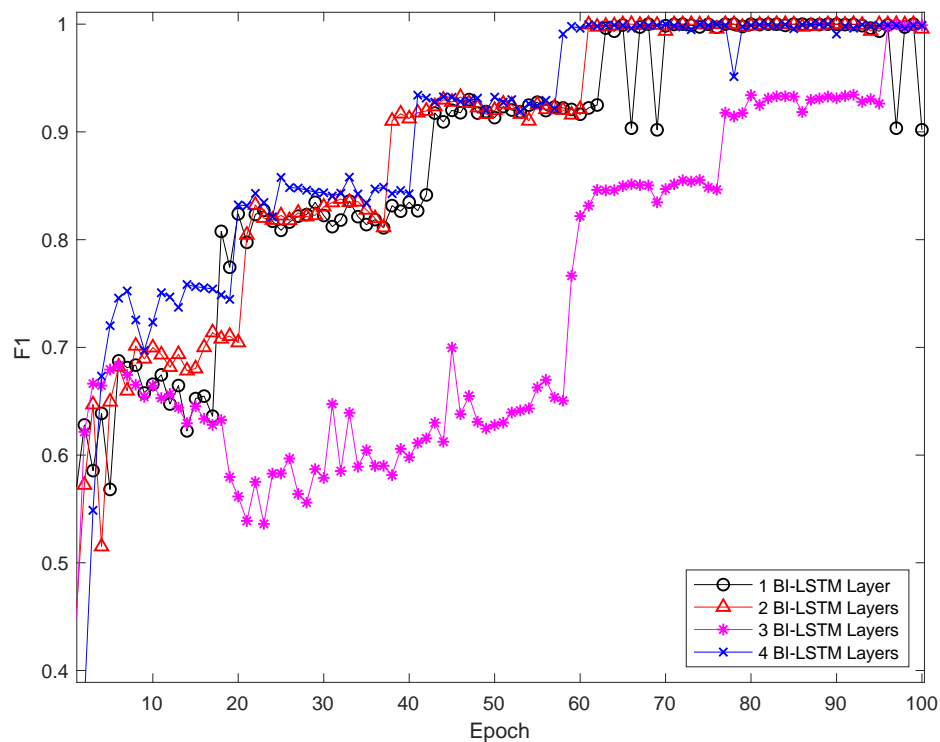


รูปที่ 12 โครงสร้างแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าโดยใช้โครงข่ายประสาทเทียมแบบสองทิศทางร่วมกับกลไกจุดสนใจ

3.2 การเพิ่มประสิทธิภาพให้กับแบบจำลอง

นอกจากการนำเสนอโครงสร้างที่เหมาะสมในการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้า และมีการเพิ่มประสิทธิภาพโดยด้วยวิธีการและเทคนิคต่าง ๆ ดังนี้

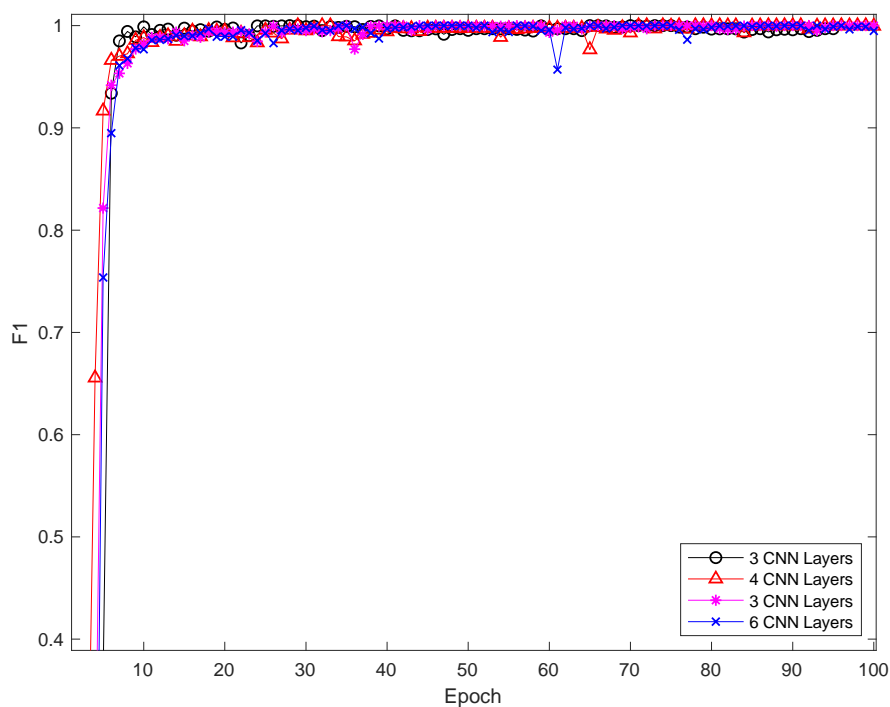
1. การปรับโครงสร้างโดยการเพิ่มจำนวนชั้นโครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง ดังรูปที่ 13 ซึ่งเป็นการทดลองให้เห็นประสิทธิภาพของการเพิ่มจำนวนชั้นหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง จาก 1 ถึง 4 ชั้นบนข้อมูลรีโมทคอมเพรสเซอร์ชนิดบี ชุดข้อมูลประเมินผล ของการเรียนรู้ 100 รอบแรก จากการทดลองพบว่าหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง 3 ชั้นให้ค่า F1 สูงที่สุด ที่ 0.94 บนข้อมูลชุดทดสอบ



รูปที่ 13 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองที่มีจำนวนชั้นหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางตั้งแต่หนึ่งถึงสี่ชั้น ทดสอบบนชุดข้อมูลประเมินผล รีโมทคอมเพรสเซอร์ชนิดบี ของการเรียนรู้ 100 รอบแรก

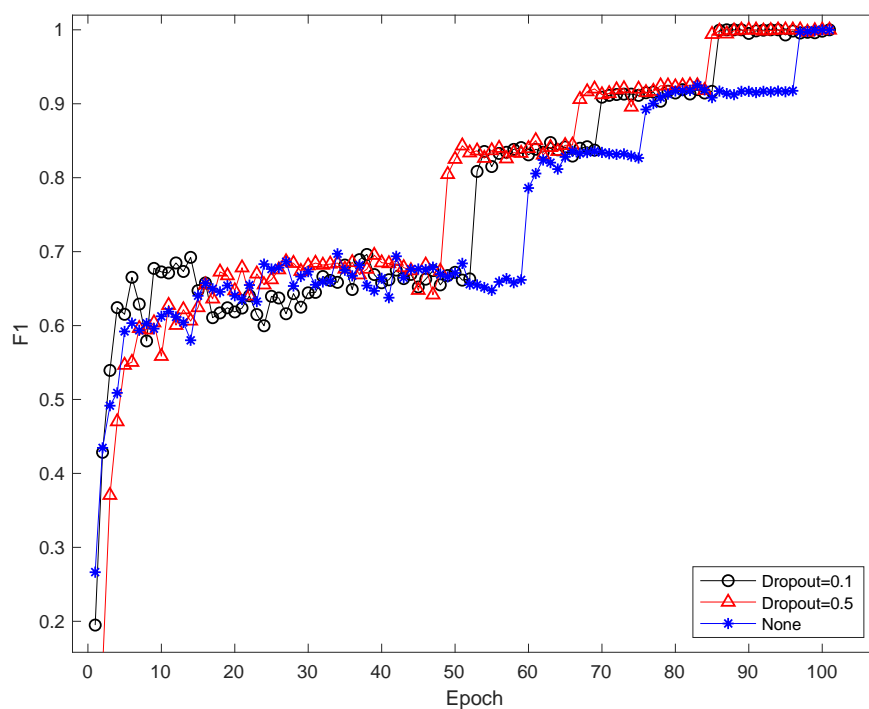
2. การปรับโครงสร้างโดยการเพิ่มจำนวนชั้นคอนโวลูชัน ดังรูปที่ 14 ซึ่งเป็นการทดลองให้เห็นประสิทธิภาพของการเพิ่มจำนวนชั้นคอนโวลูชันจากสามถึงหกชั้นในชุดข้อมูลรีโมทคอมเพรสเซอร์

ชนิดปี จากการทดลองพบว่าการใช้คอนโวลูชันสามชั้นได้ค่า F1 สูงสุดที่ 0.94 แต่ได้ค่าเฉลี่ย F1 ที่น้อยกว่าการใช้คอนโวลูชันหกชั้นดังนั้นผู้วิจัยจึงเลือกใช้คอนโวลูชันหกชั้น



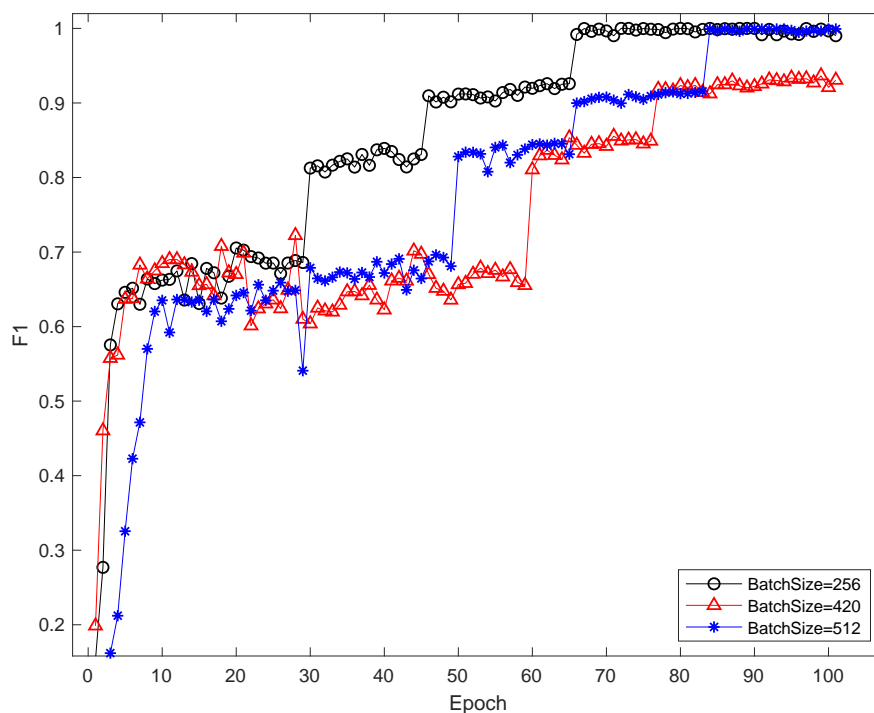
รูปที่ 14 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองที่มีจำนวนชั้นคอนโวลูชันสามถึงหกชั้น ทดสอบบนชุดข้อมูลประเมินผล ริโมทคอมแพรสเซอร์ชนิดปี ของการเรียนรู้ 100 รอบแรก

3. การนำเทคนิควิธีป้องกันการปรับเหมาะเกินไปโดยการลดจำนวน มาใช้ในการสุ่มตัดเส้นเชื่อมที่ไม่จำเป็นออกไป ทำให้หน่วยการเรียนรู้บางหน่วยไม่มีการประมวลผลข้อมูล เปรียบเสมือนการหยุดหน่วยการเรียนรู้ย่อยไม่ให้เรียนรู้กลางคัน ด้วยค่าความน่าจะเป็นค่าหนึ่ง ในวิทยานิพนธ์ได้ทำการทดลองค่าความน่าจะเป็นที่ 0.1 0.5 และ ไม่มี โดยเทคนิคนี้ทำให้ประสิทธิภาพการทำนายของแบบจำลองลู่เข้าค่าน้อยที่สุดข้างล่าง ดังรูปที่ 15 ซึ่งเป็นการทดลองให้เห็นประสิทธิภาพของการใช้การหยุดเรียนรู้กลางคัน ในชุดข้อมูลริโมทคอมแพรสเซอร์ชนิดปีในแต่ละค่าการหยุดเรียนรู้กลางคัน จากการทดลองค่าการลดจำนวนที่ดีที่สุดคือ 0.5



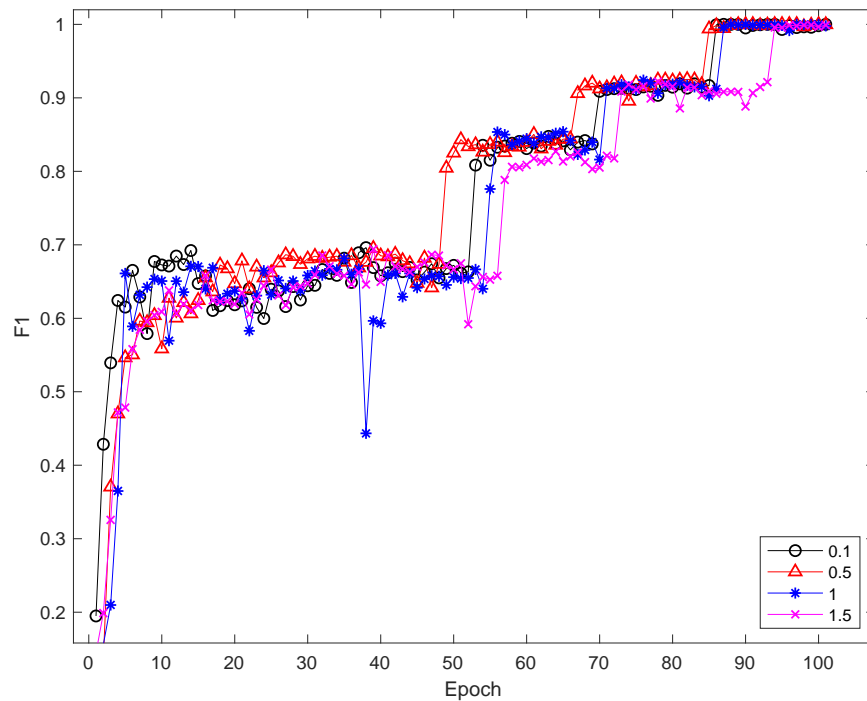
รูปที่ 15 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองด้วยการปรับค่าการลดจำนวน (Drop-out) ทดสอบบนชุดข้อมูลประเมินผล รีโมทคอมเพรสเซอร์ชนิดบีของการเรียนรู้ 100 รอบแรก

4. การปรับขนาดของแบบทชของข้อมูลฝึกมีผลต่อการปรับน้ำหนักและประสิทธิภาพการเรียนรู้ของแบบจำลอง การปรับขนาดของแบบทช สามารถเพิ่มประสิทธิภาพของการทำนายให้กับแบบจำลองได้ดังแสดงในรูปที่ 16 ซึ่งเป็นการทดลองให้เห็นถึงประสิทธิภาพของการปรับขนาดแบบทชของข้อมูลฝึก บนชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบี จากการทดลองพบว่าค่าขนาดของแบบทชที่เท่ากับ 420 ที่ได้ค่า F1 บนข้อมูลชุดทดสอบดีที่สุดที่ 0.916



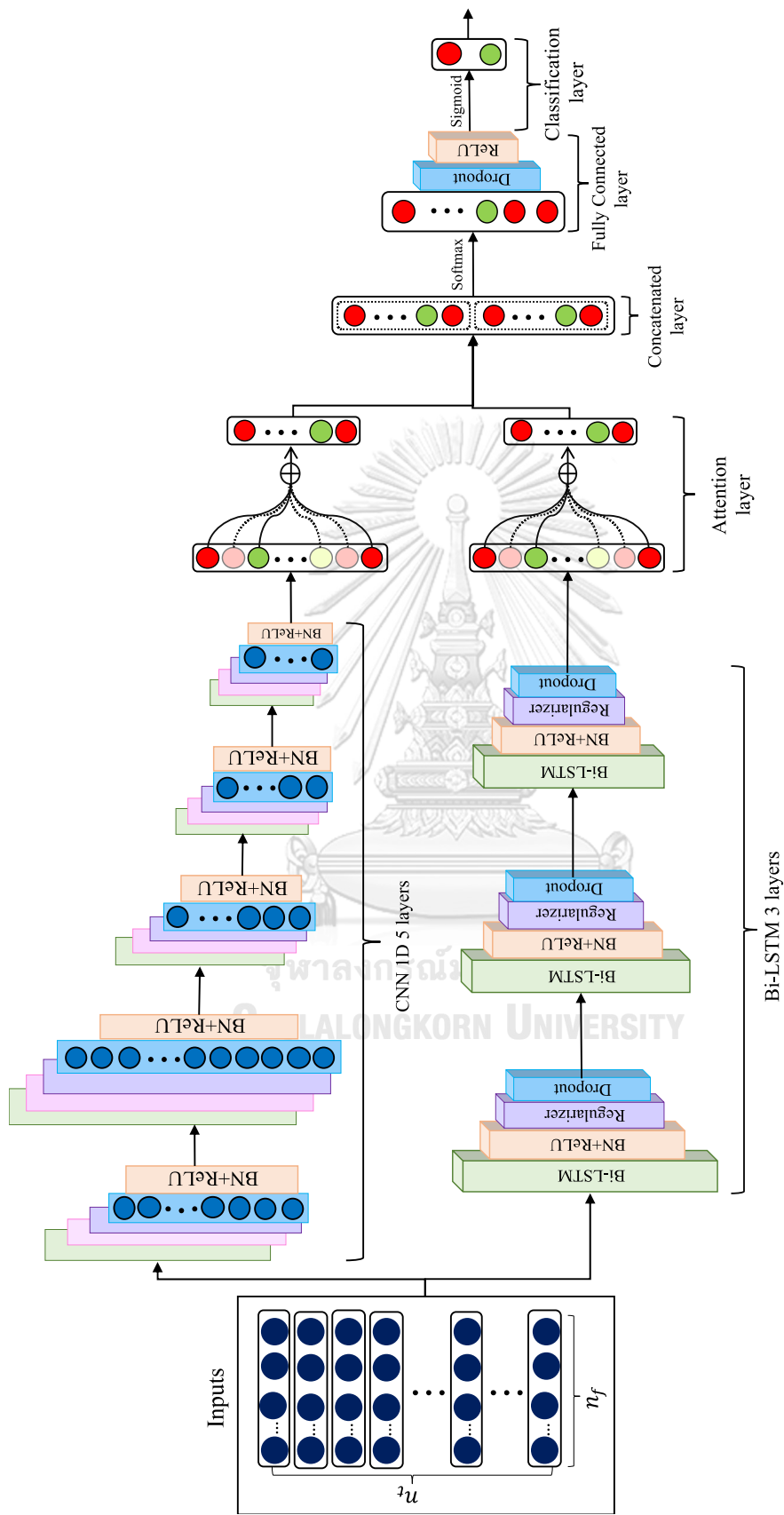
รูปที่ 16 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองด้วยการปรับขนาดแบทชของข้อมูลฝึก ทดสอบบนชุดข้อมูลประเมินผล รีโมทคอมเพรสเซอร์ชนิดบีของการเรียนรู้ 100 รอบแรก

5. การทำเรกูลาร์ไรเซชันแบบแอลสอง เพื่อลดความซับซ้อนของแบบจำลอง หลังจากการหาคุณลักษณะพิเศษโดยการเพิ่มเทคนิคนี้สามารถเพิ่มประสิทธิภาพของทำนายให้กับแบบจำลองได้ ในงานวิจัยนี้ได้มีการทดลองใช้ค่าเรกูลาร์ไรเซชันแบบแอลสองที่ 0.1 0.5 1 และ 1.5 ดังแสดงในรูปที่ 17 ซึ่งเป็นการทดลองให้เห็นถึงประสิทธิภาพของการทำเรกูลาร์ไรเซชันแบบแอลสองบนชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบีจากการทดลองค่าเรกูลาร์ไรเซชันแบบแอลสองที่เท่ากับ 1.0 ทำให้ได้ประสิทธิภาพสูงสุดที่ 0.90 บนชุดข้อมูลทดสอบ



รูปที่ 17 กราฟเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองด้วยการปรับค่าเรกูลาร์ไรเซชันแบบแอลสอง ข้อมูลบนชุดข้อมูลประเมินผลชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบีของการเรียนรู้ 100 รอบแรก

เมื่อนำเทคนิคข้างต้นมาใช้ในแบบจำลอง เพื่อช่วยเพิ่มประสิทธิภาพ โดยหลักการทำงานจะเรียงกันดังแสดงในรูปที่ 18



รูปที่ 18 โครงสร้างแบบจำลองการทำนายความล้มเหลวของรีเมทคอมเพรสเซอร์ล่วงหน้าโดยใช้โครงข่ายประสาทเทียมหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางกลไกจุดสนใจผสมกับโครงข่ายประสาทเทียมคอนโวลูชันกลไกจุดสนใจ

3.4 การแปลงเมทริกซ์ลำดับเวลาสามมิติ (Time Sequence Matrix 3D Transformation - TSMT)

จากการศึกษาวิจัยที่เกี่ยวข้อง การทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ เป็นปัญหาการทำนายอนุกรมเวลาหลายตัวแปร เมื่อกำหนดให้ ชุดข้อมูลอนุกรมเวลาหลายตัวแปร $X = \{x_1, x_2, \dots, x_n\}; x \in \mathbb{R}^{n_f \times n}$ เมื่อ n_f คือ จำนวนของคุณลักษณะ และ n คือจำนวนผลรวมของชั้นเวลาของอนุกรมเวลาหลายตัวแปร วัตถุประสงค์คือ เพื่อทำนายอนาคตที่จะเกิดเหตุการณ์ตามช่วงเวลาที่กำหนด เมื่อกำหนดให้หนึ่งอนุกรมเวลาในชุดข้อมูล X มีความยาวของชั้นเวลาเท่ากับ T แบบจำลองแทนด้วย $F(\cdot)$ โดยจะทำการป้อนข้อมูลนำเข้าหนึ่ง คือ $X_T = \{x_1, x_2, \dots, x_T\}; x \in \mathbb{R}^{n_f \times T}$ และ แบบจำลอง จะทำการสร้างลำดับผลลัพธ์เป็น ตามสมการที่ 3.1

$$\hat{Y}_{T+h} = F(X_T) \quad (3.1)$$

เมื่อ $\hat{Y}_{T+h} \in \mathbb{R}^{n_f \times T}$ เมื่อ h คือช่วงเวลาที่将会เกิดเหตุการณ์ที่แบบจำลองทำนาย ในการแปลงข้อมูลอนุกรมเวลาสองมิติให้เป็นเมทริกซ์ลำดับเวลาสามมิติหรือเทนเซอร์ที่ตี (tensor 3D) สำหรับแบบจำลองการเรียนรู้เชิงลึกเมื่อกำหนดให้ฟังก์ชันการแปลงเมทริกซ์ลำดับเวลาสามมิติ (TSMT) โดย $TSMT: X \rightarrow \tilde{X}$ เมื่อ $X \in \mathbb{R}^{n_f \times T}$ คือ การนำเข้าของอนุกรมเวลาหลายตัวแปรและผลลัพธ์จากการแปลงเมทริกซ์ลำดับเวลาสามมิติ คือ $\tilde{X} \in \mathbb{R}^{n_f \times n_t \times n_s}$ โดย n_t คือจำนวนชั้นเวลา n_s คือจำนวนของตัวอย่าง คำนวณได้จากสมการที่ 3.2

$$n_s = T - n_t + 1 \quad (3.2)$$

โดยสามารถเขียนได้เป็นอัลกอริทึมที่ 1 S คือจำนวนการเหลื่อมเมื่อกำหนดให้เท่ากับทุก ๆ 1 นาที S จึงเท่ากับ 1 และแสดงตัวอย่างในรูปที่ 20

Algorithm 1 Time Sequence Matrix 3D Transformation

Input: $X \in \mathbb{R}^{n_f \times T}$: 2D input multivariate time series

Output: $\tilde{X} \in \mathbb{R}^{n_f \times n_t \times n_s}$: 3D input temporal tensor

Init: $\tilde{X} \in \mathbb{R}^{n_f \times n_t \times n_s}$

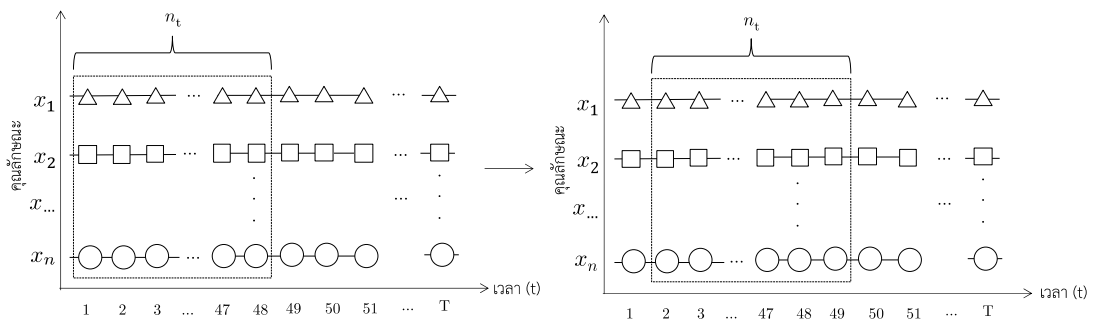
1: for $i = 1$ to n_s do

2: $\tilde{X}[:, :, i] \leftarrow X[:, i * s : i * s + n_t]$

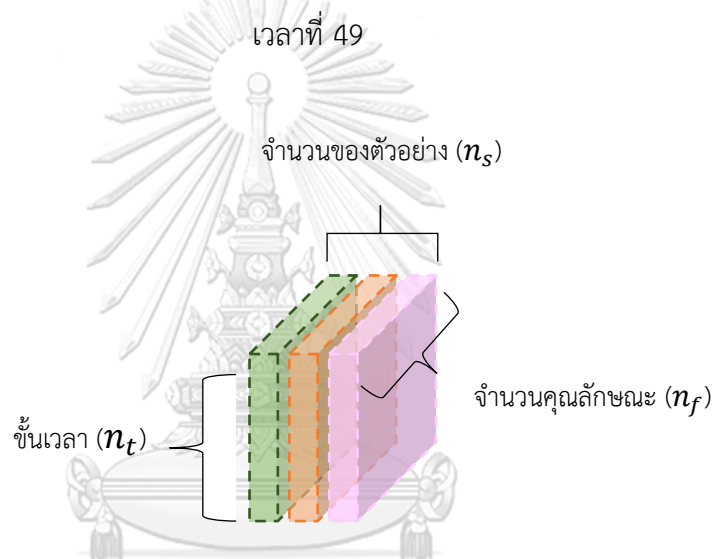
3: end for

4: return \tilde{X}

รูปที่ 19 อัลกอริทึมการแปลงเมทริกซ์ลำดับเวลาสามมิติ



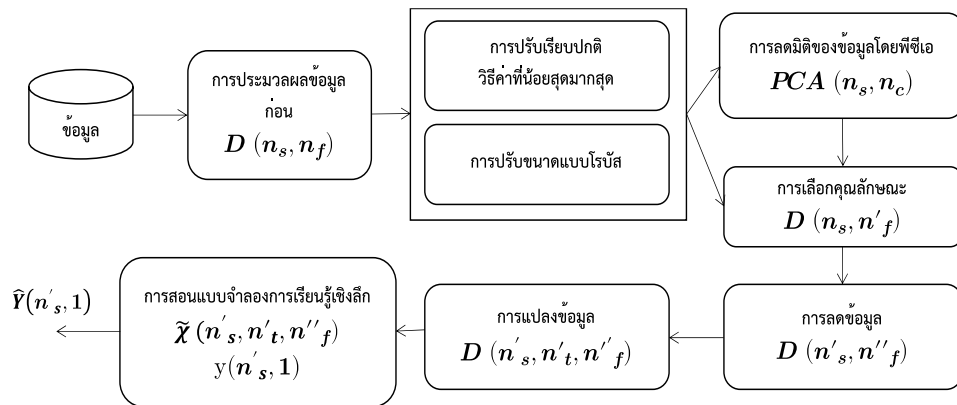
รูปที่ 20 วิธีการแปลงข้อมูลอนุกรมเวลาแบบหลายตัวแปรโดยวิธีการเมทริกซ์ลำดับเวลาสามมิติหรือการเลื่อนซ้อนทุก ๆ 48 ชั้นเวลา จากเวลาที่ 1 ถึงเวลาที่ 48 และวิธีการเลื่อนซ้อนจากเวลาที่ 2 ถึง



รูปที่ 21 เมทริกซ์ลำดับเวลาจากหนึ่งอนุกรมเวลาสามมิติ

3.5 เทคนิคการลดข้อมูล

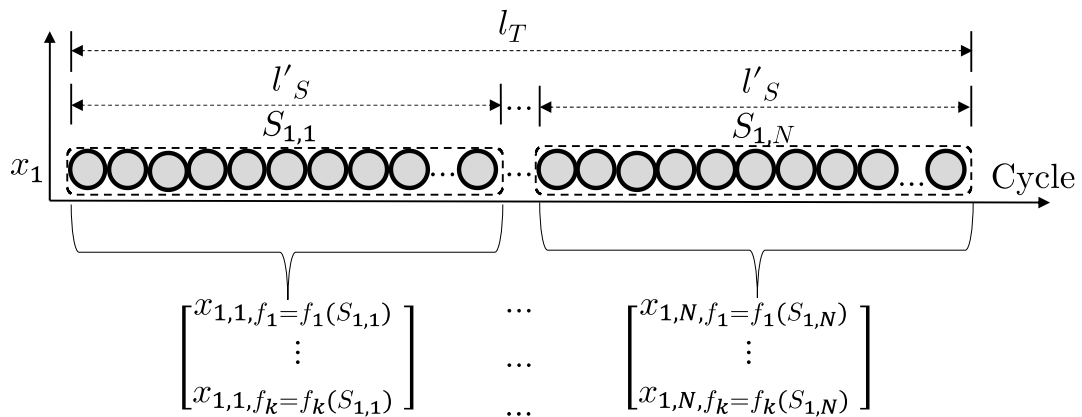
เนื่องจากข้อมูลที่มีปริมาณมากและเมื่อนำมาสร้างเป็นเมทริกซ์ลำดับเวลาสามมิติยิ่งมีปริมาณเพิ่มขึ้นอย่างมากในปัญหาที่ต้องใช้ข้อมูลขนาดใหญ่ในการสอนแบบจำลองการเรียนรู้เชิงลึก จึงมักพบปัญหาในการนำข้อมูลเข้าในหน่วยความจำหลักในครั้งเดียวได้ ดังนั้นผู้วิจัยจึงนำเสนอขั้นตอนวิธีเพื่อลดข้อมูล 3 ขั้นตอนหลักคือ 1) ลดมิติข้อมูลโดยใช้การวิเคราะห์องค์ประกอบสำคัญ 2) การเลือกคุณลักษณะสำคัญ และ 3) สกัดลักษณะสำคัญทางสถิติเพื่อลดจำนวนระเบียบข้อมูล โดยแสดงแนวคิดในรูปที่ 22



รูปที่ 22 เทคนิคการเพิ่มประสิทธิภาพโดยการสกัดลักษณะสำคัญร่วมกับเทคนิคการลดข้อมูล

ตารางที่ 3 คำอธิบายขั้นตอนการประมวลผลข้อมูลก่อนในรูปที่ 22

ขั้นตอน	คำอธิบาย
นำเข้า	นำข้อมูลเข้าจากเซ็นเซอร์และค่าคุณลักษณะต่าง ๆ ของเครื่องและฉลาก เมื่อ D คือ ชุดข้อมูลของเหตุการณ์ใด ๆ n_s คือ จำนวนของแถวของ D n_f คือ จำนวนของคุณลักษณะใน D
การประมวลผลข้อมูลก่อนหน้า	การทำความสะอาดข้อมูล การบูรณาการข้อมูล การแปลงข้อมูล โดยใช้วิธีการปรับเรียงปกติวิธีค่าน้อยสุดมากที่สุด และการปรับขนาดแบบโรบัสต์
การวิเคราะห์องค์ประกอบสำคัญ	การลดมิติของข้อมูลโดยใช้การวิเคราะห์องค์ประกอบสำคัญ ได้ข้อมูลองค์ประกอบที่สำคัญที่เลือก (n_c)
การเลือกคุณลักษณะ	เลือกคุณลักษณะที่ได้จากการแปลงข้อมูลทั้งสองวิธีร่วมกับคุณลักษณะที่ได้จากการวิเคราะห์องค์ประกอบสำคัญ (n'_f)
การลดข้อมูล	ประยุกต์ใช้ขั้นตอนทางสถิติในการลดขนาดข้อมูล โดยเมื่อผ่านขั้นตอนนี้จำนวนแถวของข้อมูล (n'_s) จะลดลงแต่จำนวนของคุณลักษณะจะมากขึ้นเท่ากับผลคูณของจำนวนคุณลักษณะที่ถูกเลือก (n'_f) กับจำนวนฟังก์ชันทางสถิติที่ใช้ในการสกัดคุณลักษณะ (k) เมื่อ $k < n'_f$
การแปลงเมทริกซ์ลำดับเวลาสามมิติ	n_t คือ ความยาวช่วงของเหตุการณ์ที่กำหนดทำการเปลี่ยนรูปแบบข้อมูลให้เป็นเมทริกซ์สามมิติ (n'_s, n'_t, n''_f) จำนวนของตัวอย่าง (n'_s) ขึ้นเวลา (n'_t) และ จำนวนคุณลักษณะ (n''_f) ผ่านกระบวนการลดข้อมูลก่อนหน้า ดังแสดงในรูปที่ 23



รูปที่ 23 แนวคิดที่นำเสนอเพื่อลดปริมาณข้อมูลด้วยการสกัดคุณลักษณะสำคัญ

โดยใช้ฟังก์ชันทางสถิติ

เมื่อ กำหนดให้ x_1 คือคุณลักษณะหนึ่ง S_1 ซึ่งเป็นข้อมูลอนุกรมเวลาของรีโมทคอมเพรสเซอร์ตั้งแต่การเริ่มทำงานจนกระทั่งก่อนหยุดการทำงานในระยะเวลาหนึ่ง (T) และ l_T คือ ความยาวของอนุกรมเวลาทั้งหมด และ (l_S) คือ จำนวนชั้นเวลาของอนุกรมเวลาที่ถูกเลือกใช้ในการสอนแบบจำลอง จากการศึกษาข้อมูลที่ใช้ในการทำวิจัยและการอ้างอิงจากวิศวกรผู้เชี่ยวชาญพบว่าเวลาที่น้อยที่สุดที่สามารถบ่งชี้ความผิดปกติหรือความถดถอยของการทำงานของเครื่องรีโมทคอมเพรสเซอร์ในกรณีที่คาดว่าระบบแจ้งเตือนสามารถทำได้อยู่ที่ประมาณ 16 ชั่วโมง หรือ 960 นาที ดังนั้นชั้นเวลา (n_t) ในการทดลอง คือ 960 เมื่อกำหนดให้การใช้จำนวนแถวของข้อมูลในการสกัดคุณลักษณะเพื่อลดข้อมูล (l'_S) อยู่ที่ 20 และเลือกใช้ฟังก์ชันทางสถิติที่ในการสกัดคุณลักษณะ (f_k) 15 ฟังก์ชัน กำหนดให้ความยาวของอนุกรมเวลา (l_S) เท่ากับความยาวของอนุกรมเวลาทั้งหมด (l_T) และเท่ากับจำนวนชั้นเวลา (n_t) ดังนั้นจำนวนของชั้นเวลาใหม่ที่ผ่านการลดข้อมูล (n'_t) โดยสามารถคำนวณได้จากสมการ 3.1

$$n'_t = \frac{l_S}{l'_S} \quad (3.1)$$

เมื่อ $f_k < n'_f$

$$n'_t = \frac{960}{20} = 48 \quad (3.2)$$

ดังนั้นจำนวนแถวของข้อมูลที่เหลือหลังจากการผ่านขั้นตอนการลดข้อมูล (n'_s) สามารถคำนวณได้จากสมการที่ 3.2

$$n'_s = \frac{l_s}{l'_s} - n'_t + 1 \quad (3.2)$$

โดยขั้นตอนวิธีในการลดปริมาณข้อมูลโดยใช้การสกัดคุณลักษณะและการแปลงข้อมูลเมทริกซ์ลำดับเวลาสามมิติแสดงตามอัลกอริทึมที่ 2 ในรูปที่ 24

Algorithm 2 TSMT with Data Reduction

Input: $X \in \mathbb{R}^{n_f \times T}$: 2D input multivariate time series

Output: $\tilde{\chi} \in \mathbb{R}^{n'_s \times n'_t \times n'_f}$: 3D input temporal tensor

Init: $\tilde{\chi} \in \mathbb{R}^{n'_s \times n'_t \times n'_f}$
 $d \in \mathbb{R}^{n_f \times T}$
 $X' \in \mathbb{R}^{n'_f \times n'_t \times n'_s}$

```

1: for  $c = 1$  to  $n_f$  do
2:    $d[0, :] \leftarrow X[c, 0, T]$ 
3:    $k[:, :] \leftarrow \text{empty}[:, :]$ 
4:   for  $i$  in range(0,  $T, l'_s$ ) do
5:      $e[:, :] \leftarrow d[i * s : i * s + n'_t]$ 
6:      $v[:, :] \leftarrow \text{empty}[:, :]$ 
7:     for  $j=1$  to  $f_k$  do
8:        $v \leftarrow \text{concatinate}(v, f_j(e))$ 
9:     end for
10:     $k \leftarrow \text{append}(k, V)$ 
11:  end for
12:   $X' \leftarrow \text{concatinate}(X', k)$ 
13: end for
14: for  $p = 1$  to  $n'_s$  do
15:   $\tilde{\chi}[:, :, p] \leftarrow X'[:, p * s : p * s + n'_t]$ 
16: end for
17: return  $\tilde{\chi}$ 

```

รูปที่ 24 อัลกอริทึมการแปลงเมทริกซ์ลำดับเวลาสามมิติโดยการลดปริมาณข้อมูล

หลักการวัดประสิทธิภาพในการลดจำนวนพื้นที่หน่วยความจำโดยประมาณ

เมื่อพื้นที่หน่วยความจำหลักโดยประมาณ (\mathcal{M}) หลังจากการแปลงข้อมูลให้อยู่ในรูปเมทริกซ์ลำดับเวลาสามมิติ ด้วยวิธีการพื้นฐานที่ไม่มีกระบวนการลดข้อมูลและไม่มีการเลือกคุณลักษณะแสดงในสมการที่ 3.3

$$\mathcal{M} = n_f \times (T - n_t + 1) \quad (3.3)$$

เมื่อมีจำนวนคุณลักษณะ (n_f) ทั้งหมดเท่ากับ 56 คุณลักษณะ มีข้อมูลทั้งหมดของชุดข้อมูลฝึกของรีโมทคอมเพรสเซอร์ชนิดปีมีทั้งหมด 2,491,560 แถว (T) เมื่อผ่านกระบวนการแปลงข้อมูลให้อยู่ในรูปเมทริกซ์ลำดับเวลาสามมิติ จะใช้พื้นที่หน่วยความจำหลักโดยประมาณ 139,473,656 หน่วย เมื่อผ่านการเลือกคุณลักษณะสำคัญและการสกัดข้อมูลโดยใช้ฟังก์ชันทางสถิติ (n'_f) จำทำให้จำนวนแถวของข้อมูลลดลง (n'_s) จึงสามารถคำนวณพื้นที่หน่วยความจำหลักโดยประมาณใหม่ จากสมการที่ 3.4

$$\tilde{M} = (n'_f \times k) \times n'_s \quad (3.4)$$

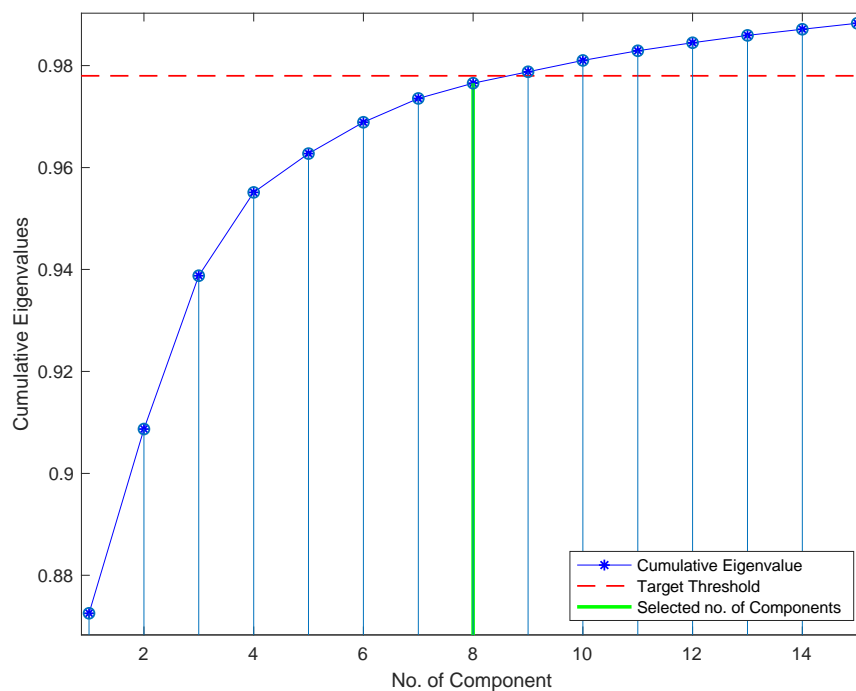
$$\begin{aligned} \tilde{M} &= n'_f \times k \times \left(\frac{T}{l'_s} - n'_t + 1 \right) \\ &= 25 \times 15 \times \left(\frac{2,491,560}{20} - 48 + 1 \right) \\ &= 66,735,438 \text{ หน่วย} \end{aligned}$$

ดังประสิทธิภาพในการลดจำนวนพื้นที่หน่วยความจำลงโดยประมาณ เท่ากับร้อยละ 66.52

3.6 การวิเคราะห์องค์ประกอบสำคัญ

การวิเคราะห์องค์ประกอบสำคัญมีวัตถุประสงค์เพื่อลดมิติของชุดข้อมูลและเพิ่มประสิทธิภาพในการแบ่งแยกข้อมูลที่อยู่คนละประเภทให้ออกจากกันได้ดียิ่งขึ้น เช่น ข้อมูลของเซ็นเซอร์ที่มีค่าอยู่ในช่วงเดียวกันจะถูกจัดให้อยู่ในประเภทเดียวกัน ในวิทยานิพนธ์นี้แบ่งข้อมูลออกเป็น 2 ประเภทหลักคือ ข้อมูลคุณลักษณะที่ก่อนเกิดความล้มเหลวและข้อมูลคุณลักษณะในช่วงเวลาการทำงานปกติ โดยใช้ข้อมูลคุณลักษณะจำนวน 84 คุณลักษณะที่แสดงในตารางที่ ก.5 – ก.6 ในภาคผนวก ก ซึ่งการวิเคราะห์องค์ประกอบสำคัญเป็นขั้นตอนย่อยเพื่อหาองค์ประกอบสำคัญหรือคุณลักษณะที่ได้จากการหาเมทริกซ์การแปลงและเวกเตอร์เฉลี่ย จากชุดข้อมูลฝึกและชุดข้อมูลประเมินผลถูกใช้ในการหาเมทริกซ์ไอเกนเวกเตอร์ (A_i) เมทริกซ์ค่าไอเกน และเมทริกซ์ค่าไอเกนและเวกเตอร์ค่าเฉลี่ย (M_{xi}) 1 ชุดและใช้ข้อมูลชุดทดสอบมาทำการแปลงเป็นเมทริกซ์เวกเตอร์ค่าเฉลี่ย โดยข้อมูลดังกล่าวจะถูกนำไปใช้ในขั้นตอนการเลือกคุณลักษณะและการทดลองในขั้นต่อไป โดยเกณฑ์การเลือกองค์ประกอบสำคัญใช้เกณฑ์เปอร์เซ็นต์ความแปรปรวน ใช้ค่าขีดแบ่ง $L = 0.975$ และ 0.965 สำหรับข้อมูลของ

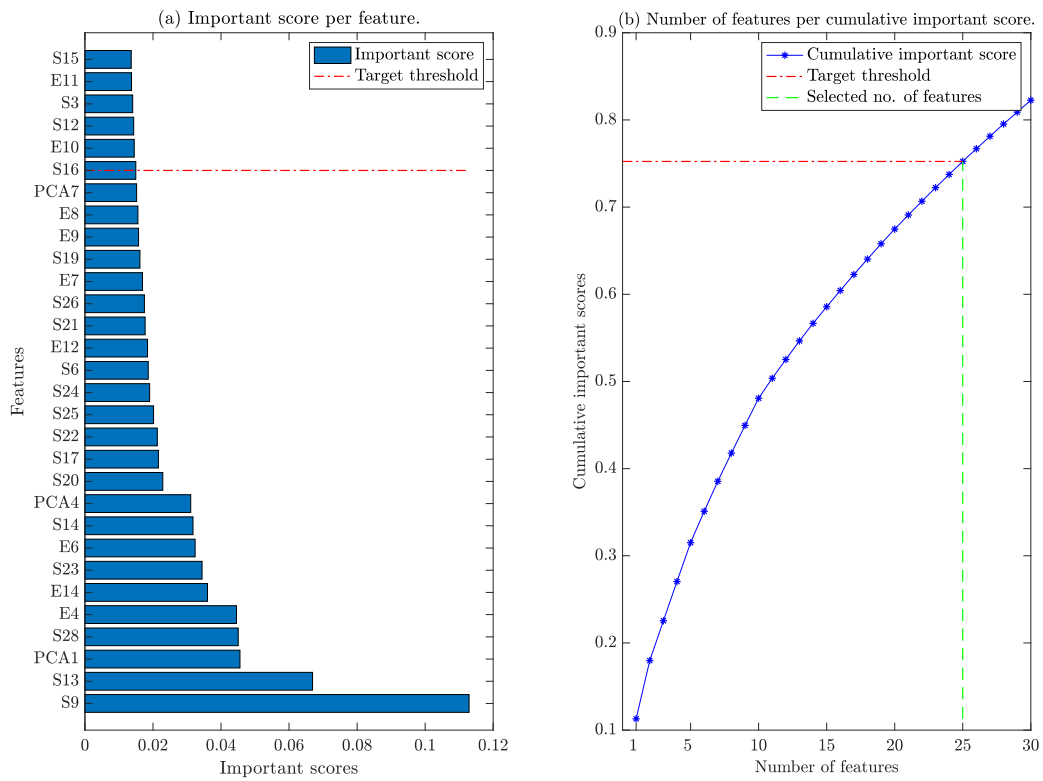
วิธีโทคมเพอร์เซฮอร์ชนิดปีและชนิดดีตามลำดับ ดังแสดงให้เห็นใน ตารางที่ ข.1 และ ข.2 ใน ภาคผนวก ข และแสดงภาพตัวอย่างของการเลือกองค์ประกอบสำคัญด้วยวิธีนี้ ในรูปที่ 25 ซึ่งสามารถ เลือกเอาองค์ประกอบสำคัญ 8 อันดับแรกจาก 15 องค์ประกอบ



รูปที่ 25 การเลือกองค์ประกอบสำคัญโดยใช้เกณฑ์เปอร์เซ็นต์ความแปรปรวน ใช้ค่าขีดแบ่ง $L = 0.975$ ซึ่งทำให้สามารถเลือก 8 อันดับแรกขององค์ประกอบสำคัญไปใช้ในกระบวนการขั้นต่อไป

3.6 การเลือกคุณลักษณะ

ขั้นตอนการเลือกคุณลักษณะนี้จะนำเอาคุณลักษณะที่ได้จากการประมวลผลก่อนหน้า 42 คุณลักษณะและคุณลักษณะที่ผ่านการแปลงจากองค์ประกอบสำคัญที่ถูกเลือกทั้ง 8 องค์ประกอบ รวมเป็น 50 คุณลักษณะมาใช้เป็นข้อมูลนำเข้าเพื่อทำการฝึกแบบจำลองไลต์แอลจีบีเอ็มและเอ็กซ์ทรีมเกรนเดียร์นบูตติงด้วยทำการฝึกแบบจำลองห้ารอบโดยใช้ข้อมูลชุดฝึก ชุดประเมินผลและชุดทดสอบเดียวกันกับการทำลองเปรียบเทียบประสิทธิภาพแบบจำลอง ซึ่งพบว่าไลต์จีบีเอ็มได้ค่า F1 เฉลี่ยมากที่สุดดังนั้นวิทยานิพนธ์นี้จึงใช้คะแนนความสำคัญเฉลี่ยที่ได้จากการฝึกแบบจำลองไลต์จีบีเอ็มในการเลือกคุณลักษณะสำคัญ โดยใช้เกณฑ์การเลือกด้วยค่าขีดแบ่งเป้าหมาย (target threshold) ของเปอร์เซ็นต์คะแนนความสำคัญสะสมเฉลี่ยที่ 75% จากการวิเคราะห์ผลพบว่า มี 25 คุณลักษณะที่ผ่านเกณฑ์ ตามที่แสดงใน รูปที่ 26 เอ ค่าเปอร์เซ็นต์คะแนนความสำคัญของ 30 คุณลักษณะแรก รูปที่ 26 บี แสดงจำนวนคุณลักษณะต่อค่าเปอร์เซ็นต์คะแนนความสำคัญสะสม



รูปที่ 26 การวิเคราะห์คะแนนความสำคัญสะสมของคุณลักษณะ

โดยใช้แบบจำลองโลทจีบีเอ็ม

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 4

แนวทางและวิธีการดำเนินงาน

วิทยานิพนธ์นี้นำเสนอการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าโดยใช้โครงข่ายประสาทเทียมหน่วยความจำสั้นระยะยาวแบบสองทิศทาง กลไกจุดสนใจ โครงข่ายประสาทเทียมคอนโวลูชัน และเทคนิคการลดข้อมูล ซึ่งบทนี้จะกล่าวถึงแหล่งข้อมูล การจัดเก็บข้อมูล การจัดเตรียมชุดข้อมูลของรีโมทคอมเพรสเซอร์ การสกัดคุณลักษณะสำคัญ การลดข้อมูล พารามิเตอร์ของแบบจำลองที่นำเสนอ การทดลองเพื่อเปรียบเทียบประสิทธิภาพวิธีการเลือกคุณลักษณะและแบบจำลองที่นำเสนอกับแบบจำลองการเรียนรู้เชิงลึกและแบบจำลองการเรียนรู้ของเครื่องที่ได้รับความนิยมงานวิจัยที่เกี่ยวข้อง

4.1 แหล่งข้อมูล

ชุดข้อมูลที่ใช้ในงานวิจัยนี้ ได้รับมาจาก บริษัท เชฟรอนประเทศไทยสำรวจและผลิต จำกัด (Chevron Thailand Exploration and Production, Ltd.) ซึ่งเป็นข้อมูลของรีโมทคอมเพรสเซอร์ ทั้งหมด 14 เครื่อง และ ข้อมูลประวัติการทำงานของของรีโมทคอมเพรสเซอร์ ตั้งแต่ปี พ.ศ. 2559 ถึง 2562 และข้อมูลอื่นที่ใช้ในการสร้างแบบจำลอง โดยข้อมูลดิบได้จาก ระบบบันทึก 2 ระบบได้แก่ ระบบสารสนเทศโรงงาน (Plant Information - PI) และ ระบบรีโมทคอมเพรสเซอร์เว็บ (remote compressor web) ระบบพายฮิสทอเรียล (PI Historian system) ได้รับข้อมูลจาก ระบบการควบคุมดูแลและการเก็บข้อมูล (Supervisory Control and Data Acquisition - SCADA) โดยระบบพายฮิสทอเรียล จัดเก็บข้อมูลลงในฐานข้อมูลเชิงสัมพันธ์ (relational database) ผู้วิจัยได้สร้างโปรแกรมสกัด-แปลง-โหลด (Extract-Transform-Load - ETL) เพื่อบูรณาการข้อมูลจากระบบพายฮิสทอเรียล และ ข้อมูลระบบรีโมทคอมเพรสเซอร์

1. ระบบพายฮิสทอเรียล ประกอบไปด้วย 2 ฐานข้อมูลหลัก ได้แก่

(1) ฐานข้อมูลระบบพายกรอบสินทรัพย์ ข้อมูลบรรยายคุณลักษณะ (meta data) ของเซ็นเซอร์ซึ่งเก็บข้อมูลในลักษณะฐานข้อมูลเชิงสัมพันธ์ (relational database) แบบลำดับชั้นความสัมพันธ์ (hierarchy relationship)

(2) ฐานข้อมูลประวัติ (PI Achieve) เป็น ข้อมูลในอดีตของเซ็นเซอร์ ผู้วิจัยเลือกใช้ข้อมูลจาก 2 องค์ประกอบสำคัญ ได้แก่ Reciprocating Compressor และ Reciprocating Engine โดยแสดงรายละเอียดในตารางที่ 4.1

2. ระบบรีโมทคอมเพรสเซอร์เว็บ

หน้าที่หลักของระบบคือเป็นระบบบันทึกประวัติการหยุดการทำงานของรีโมทคอมเพลสเซอร์ผู้วิจัยใช้เป็นข้อมูลในการติดตาม โดยเลือกเฉพาะเหตุการณ์ การบำรุงรักษาเชิงแก้ไขแบบไม่ได้วางแผนล่วงหน้า (UCM) และการหยุดการทำงานของเครื่องแบบวางแผนล่วงหน้า (Planned Downtime - PD)

4.2 การเก็บรวบรวมข้อมูลและการวิเคราะห์ข้อมูลเชิงตรวจสอบ

ในเบื้องต้นของการวิจัย ผู้วิจัยทำการดึงข้อมูลตัวอย่างเพื่อทำการศึกษาและทดลองสมมติฐาน โดยใช้ฐานข้อมูล Microsoft SQL Server 2012 โดยพบข้อมูลเบื้องต้นดังต่อไปนี้

- ขนาด (volume) ข้อมูลของกลุ่มตัวอย่างทั้งหมดที่ใช้มีขนาด 1.55 เทระไบต์ (TB) และเหลือ 13 กิกะไบต์ (GB) หลังจากการสกัดและแปลงให้อยู่ในรูปแบบที่สามารถนำไปใช้ในการวิเคราะห์และสร้างแบบจำลองได้ และมีเหตุการณ์ UCM ตั้งแต่ปี พ.ศ. 2559 ถึง 2562 มีมากกว่า 2,000 เหตุการณ์ในพื้นที่ที่ถูกเลือกในการทำวิจัย

- อัตราความเร็ว (velocity) ข้อมูลหรือความถี่ของข้อมูลที่จัดเก็บไว้ในระบบ PI Historian ถูกจัดเก็บตามอัตราการบีบอัดข้อมูลของแต่ละเซ็นเซอร์ซึ่งโดยทั่วไปจะเก็บข้อมูลทุก ๆ 1 นาที

- ความหลากหลาย (variety) ข้อมูลที่ใช้นำมาจาก 3 ฐานข้อมูล ที่มีโครงสร้างแตกต่างกัน ได้แก่ ฐานข้อมูลเชิงอนุกรมเวลา ฐานข้อมูลเชิงสัมพันธ์ และฐานข้อมูลเชิงสัมพันธ์ลำดับชั้น

จากการวิเคราะห์ข้อมูลเบื้องต้นและกลุ่มตัวอย่างยังเป็นเพียงร้อยละ 7.86 ของข้อมูลของอุปกรณ์ชนิดนี้ที่องค์กร ดังนั้นเพื่อให้การวิเคราะห์ข้อมูลเชิงตรวจสอบ การประมวลผลข้อมูลก่อนหน้า รวมถึงกระบวนการสร้างแบบจำลอง ดำเนินไปได้อย่างรวดเร็ว ผู้วิจัยจึงเลือกใช้แพลตฟอร์มการวิเคราะห์ข้อมูลขนาดใหญ่ ของบริษัทไมโครซอฟต์ ชื่อว่า Azure Databricks ซึ่งเป็นสถาปัตยกรรม Spark Cluster บนแพลตฟอร์มการประมวลผลแบบคลาวด์ โดยข้อมูลที่ใช้ในการวิจัยถูกเก็บไว้ใน Azure Data Lake

4.3 การประมวลผลข้อมูลก่อนหน้า (Data Pre-processing)

ข้อมูลที่ใช้ในการทำวิจัยเลือกใช้ข้อมูลรีโมทคอมเพลสเซอร์จำนวน 14 เครื่อง โดยแต่ละเครื่องจะมีอายุการใช้งานที่แตกต่างกันและมีการปรับค่าตั้งต้นของเครื่องที่แตกต่างกันตามมาตรฐานของรุ่นในการผลิต และปรับแต่งของวิศวกร ข้อมูลเป็นข้อมูลที่เริ่มตั้งแต่เครื่องเริ่มทำงานอย่างปกติ และจากการหยุดการทำงานครั้งก่อนหน้าไม่ว่าจะด้วยสาเหตุอะไร จนกระทั่งเครื่องหยุดการทำงาน อาจเริ่มเสื่อมถอยในระหว่างทำงานในบางจุด เกินกว่าเกณฑ์กำหนดของระบบรักษาความปลอดภัยของเครื่อง ระบบจะถือว่าการทำงานของเครื่องอยู่ในภาวะไม่ปลอดภัยจึงสั่งหยุดการทำงานโดยอัตโนมัติ ดังนั้นช่วงเวลาก่อนที่เครื่องจะถูกสั่งหยุดการทำงานโดยอัตโนมัตินี้ เป็นช่วงเวลาที่ต้องการให้

แบบจำลองสามารถการทำนายได้อย่างน้อย 1 วันล่วงหน้า โดยขั้นตอนหลักในการประมวลผลข้อมูลก่อนหน้าได้แก่ 1) การทำความสะอาดข้อมูล 2) การบูรณาการข้อมูล 3) การสกัดลักษณะสำคัญ และ 4) การแปลงข้อมูล

4.3.1 การทำความสะอาดข้อมูล(Data Cleaning)

จากการศึกษาข้อมูลเบื้องต้นพบข้อมูลอนุกรมเวลาไม่สมบูรณ์เนื่องจากสาเหตุดังต่อไปนี้

กรณีที่ 1 เนื่องจากสัญญาณของเซ็นเซอร์ของอุปกรณ์ขาดหายไปในช่วงหนึ่ง อาจเกิดจากสภาพอากาศทำให้การสื่อสารของเครื่องมือวัดขัดข้อง เรียกว่าข้อมูลขาดหาย ผู้วิจัยนำเสนอแนวทางในการใช้ค่ามัธยฐานในการวางทับค่าที่ขาดหายคำนวณตามสมการที่ 4.1

$$\text{median}(S_{m,k,l,sd}) = \begin{cases} S_{(n_t+1)/2} & : n_t \text{ เป็นเลขคี่,} \\ \frac{1}{2}(S_{(n_t/2)} + S_{(\frac{n_t}{2}+1)}) & : n_t \text{ เป็นเลขคู่} \end{cases} \quad (4.1)$$

เมื่อ

$S_{m,k,l,sd}$ คือ ข้อมูลอนุกรมเวลาของเซ็นเซอร์ m ของรีโมทคอมเพรสเซอร์ k ช่วงเวลา l และ สาเหตุการหยุดการทำงานประเภท sd

l_y คือ ช่วงเวลาที่กำหนด คือ $[t_s, t_d]$ ก่อนเกิดความล้มเหลวของปีที่ y โดยที่ t_s คือ RUL เริ่ม และ t_d คือ RUL สุดท้ายที่เลือกใช้ในการหาค่า

กรณีที่ 2 เซ็นเซอร์นั้นไม่ได้ถูกเชื่อมต่อเข้ากับระบบพายฮิสโทรเรียล หรือ รีโมทคอมเพรสเซอร์ไม่ได้ติดตั้งเซ็นเซอร์ ในตำแหน่งที่เหมือนกับรีโมทคอมเพรสเซอร์ตัวอื่น ๆ ดังนั้นผู้วิจัยจะไม่ใช้เซ็นเซอร์นั้นมาเป็นลักษณะหรือทำการทดลองโดยพิจารณาว่าส่วนใหญ่ไม่มีค่าของเซ็นเซอร์ตัวดังกล่าวและผู้เชี่ยวชาญเห็นด้วยในการตัดออก

กรณีที่ 3 การบันทึกเวลาหยุดการทำงานหรือเริ่มการทำงานของรีโมทคอมเพรสเซอร์มีความคลาดเคลื่อน ผู้วิจัยตัดเหตุการณ์นั้นออกจากการทดลอง เนื่องจากการบันทึกเวลาคลาดเคลื่อนในระบบรีโมทคอมเพรสเซอร์เว็บทำให้เกิดค่าแปลกแยก (outlier) และค่าที่ไม่สอดคล้อง (inconsistency)

เนื่องจากข้อมูลมีค่ารบกวนจากความผิดปกติของตัวรับสัญญาณทำให้เกิดค่าที่สูงหรือต่ำผิดปกติในช่วงหนึ่งซึ่งอาจไม่เกิดขึ้นจริง ผู้วิจัยจึงเสนอการแก้ไขปัญหามาโดยการปรับเรียบข้อมูลรบกวน (smoothing noisy data) โดยใช้ค่าเฉลี่ยเคลื่อนที่อย่างง่าย (Simple Moving Average - SMA) คำนวณได้จากสมการที่ 4.2

$$SMA(S) = \frac{A_1 + A_2 + \dots + A_n}{n} \quad (4.2)$$

เมื่อ

S คือ ข้อมูลอนุกรมเวลาของเซ็นเซอร์ใด ๆ

A_i คือ ค่าของเซ็นเซอร์ ณ เวลาที่ **i**

n คือ จำนวนช่วงเวลา

4.3.2 การบูรณาการข้อมูล (Data fusion)

ใช้ข้อมูลจากระบบรีโมทคอมเพรสเซอร์เว็บร่วมกับ PI Achieve และ PI AF สร้างตารางข้อมูล และค่าต่าง ๆ เช่นเวลาที่เหลืออยู่ (Remaining Useful Life – RUL) วงรอบการทำงานของเครื่อง (cycle) และประเภทเหตุการณ์ เป็นต้น และจากการศึกษาแนวทางในทำวิจัยของ [17,20]

ตารางที่ 4 โครงสร้างตารางข้อมูล

ดัชนี	แอตทริบิวต์	ประเภท	คำอธิบาย
1	Engine Code	String	
2	Cycle	Integer	time in cycles [1,n]
3	Event Code	String	Event code before failure
4	RUL	Integer	Remaining Useful Life
5	LABEL1	Integer	
6	S_1	Double	Sensor measurement 1
...
n	S_n	Double	Sensor measurement n

4.3.3 การสกัดคุณลักษณะสำคัญ (Feature Extraction)

ในขั้นตอนการประมวลผลข้อมูลก่อนหน้าเพื่อสกัดคุณลักษณะสำคัญ โดยการวัดความคล้ายคลึงกันระหว่างค่าขีดจำกัดสูงสุดและค่าขีดจำกัดต่ำสุดของแต่ละเซ็นเซอร์ ที่ระบุไว้ในระบบควบคุมความปลอดภัยของรีโมทคอมเพรสเซอร์ โดยใช้วิธีการวัดระยะทางแบบยูคลิด (Euclidean distance) เมื่อ $x^{(i,k,m)}$ คือค่าระยะทางแบบยูคลิดของข้อมูลเซ็นเซอร์ที่คำนวณได้จากสมการที่ 4.3 จากค่าขีดจำกัดสูงสุด ($H_{max}^{(i,m)}$) และ ค่าขีดจำกัดต่ำสุด ($H_{min}^{(i,m)}$) ของเซ็นเซอร์ที่ m^{th} ของรีโมทคอมเพรสเซอร์เครื่องที่ i^{th} ณ ข้อมูลจุดที่ k^{th}

$$E_{x(i,k,m)} = \sqrt{(x^{(i,k,m)} - H_{max}^{(i,m)})^2 + (x^{(i,k,m)} - H_{min}^{(i,m)})^2} \quad (4.3)$$

วิธีในการสกัดคุณลักษณะสำคัญด้วยฟังก์ชันทางสถิติเพื่อลดข้อมูลอนุกรมเวลาในหนึ่งเหตุการณ์ (S) ก่อนนำเข้าฝึกแบบจำลอง โดยสกัดคุณลักษณะสำคัญของช่วงความยาวของลำดับ (S_l) ในแต่ละคุณลักษณะที่ถูกเลือก

$$S = (s(t_1), s(t_2), \dots, s(t_v), \dots, s(t_{n_t}))^T \quad (4.4)$$

$$S = (s_1, s_2, \dots, s_{n_t})^T$$

เมื่อ s_v คือ ค่าของคุณลักษณะ ณ ลำดับที่ v โดย $v \geq 1$ และ $v \leq n_t$

S คือ เมทริกซ์สลับเปลี่ยนของข้อมูลลำดับอนุกรมเวลาของคุณลักษณะใด ๆ

n_t คือ ความยาวข้อมูลลำดับอนุกรม S

ตารางที่ 5 ฟังก์ชันทางสถิติที่ใช้ในการสกัดคุณลักษณะสำคัญ

ชื่อสมการ	สมการ	
ค่ามากที่สุด (Max)	$\text{Maximum}(S) = \max\{s_1, s_2, \dots, s_v, \dots, s_{n_t}\}$	(4.5)
ค่าน้อยสุด (Min)	$\text{Minimum}(S) = \min\{s_1, s_2, \dots, s_v, \dots, s_{n_t}\}$	(4.6)
ค่าเฉลี่ย (Mean)	$\text{mean}(S) = \bar{S} = \frac{1}{n_t} \sum_{v=1}^{n_t} s_v$	(4.7)
ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)	$\text{std}(S) = \sqrt{\frac{1}{n_t} \sum_{v=1}^{n_t} (s_v - \bar{S})^2}$	(4.8)
ค่าแปรปรวน (Varian)	$\text{var}(S) = \frac{1}{n_t} \sum_{v=1}^{n_t} (s_v - \bar{S})^2$	(4.9)
เปอร์เซ็นต์ไทล์ (Percentile) ที่ k แบบไม่แจกแจงความถี่	$\text{Percentile } k(S) = S \frac{k}{100} (n + 1)$	(4.10)

ตารางที่ 6 ฟังก์ชันทางสถิติที่ใช้ในการสกัดลักษณะสำคัญ (ต่อ)

ชื่อสมการ	สมการ	
ค่ากว้างมากที่สุด (Max Range)	$\text{Max Range} = \text{Percentile } 100 - \text{Percentile } 0$	(4.11)
ค่าความสัมพันธ์ เปอร์เซ็นต์ไทล์ (Relative Percentile)	$\text{Relative Percentile } k = \text{Percentile } k(S) - \text{mean}(S)$	(4.12)
ค่าเฉลี่ยกำลังสอง (Root Mean Square Value)	$\text{RMS}(S) = \frac{1}{n_t} \sum_{v=1}^{n_t} (S^2_v)^{\frac{1}{2}}$	(4.13)

4.3.4 การแปลงข้อมูล (Data Transformation)

เนื่องจากข้อมูลแต่ละคุณลักษณะมีค่าที่แตกต่างกันมาก จึงต้องมีการนำข้อมูลมาทำการนอร์มัลไรเซชันเพื่อให้ชุดข้อมูลมีช่วงข้อมูลใกล้เคียงกัน เหมาะแก่การนำไปหาความสัมพันธ์ หรือวิเคราะห์ร่วมกัน โดย ตัวปรับขนาดแบบแข็งแกร่ง (robust scaler) โดยใช้ค่าพิสัยระหว่างควอร์ไทล์ (interquartile range) เป็นตัวหารผลต่างของข้อมูลในแถว i กับควอร์ไทล์ที่ 1 ซึ่งวิธีนี้มีความทนทานต่อค่าแปลกแยกมาก ดังสมการที่ 4.14

$$\text{robustScaler}(x_i) = \frac{x_i - Q_1(X)}{Q_3(X) - Q_1(X)} \quad (4.14)$$

Q_1 คือ ค่าควอร์ไทล์ที่ 1

Q_3 คือ ค่าควอร์ไทล์ที่ 3

x_i คือ ค่าข้อมูลแถวที่ i

4.4 การทดลอง

วิทยานิพนธ์นี้ นำเสนอแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์โดยใช้โครงข่ายประสาทเทียมหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางร่วมกับกลไกลดจุดสนใจ ผสานกับโครงข่ายประสาทเทียมคอนโวลูชันร่วมกับกลไกลดจุดสนใจ สำหรับรีโมทคอมเพรสเซอร์ชนิดปีและชนิดดี โดยการทดลอง แบ่งออกเป็น 3 การทดลอง ได้แก่

1) การทดลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองกับวิธีการเลือกคุณลักษณะ 4 วิธีดังต่อไปนี้

1.1 การใช้คุณลักษณะที่ได้จากการประมวลผลก่อนหน้า 56 คุณลักษณะ ซึ่งได้จากการนำ 28 เซ็นเซอร์ผ่านกระบวนการนอร์มัลไรเซชันด้วยวิธีค่าพิสัยระหว่างควอร์ไทล์และมากที่สุดน้อยสุดนอร์มัลไรเซชัน ซึ่งถือเป็นการสกัดคุณลักษณะ

1.2 การใช้คุณลักษณะจากการวิเคราะห์ห้องค์ประกอบสำคัญ 8 คุณลักษณะ

1.3 การใช้คุณลักษณะที่ได้จากการประมวลผลก่อนหน้าและคุณลักษณะจากการวิเคราะห์ห้องค์ประกอบสำคัญ 64 คุณลักษณะ

1.4 การใช้คุณลักษณะที่ได้จากกระบวนการเลือกคุณลักษณะ 25 คุณลักษณะ

ในงานวิจัยนี้ได้ทำการทดลองเบื้องต้นเกี่ยวกับการใช้วิธีการนอร์มัลไรเซชันกับค่าเซ็นเซอร์เพียงวิธีใดวิธีหนึ่งแล้วพบว่าประสิทธิภาพดีกว่าการใช้สกัดคุณลักษณะเพิ่มด้วยวิธีมากที่สุดน้อยสุดนอร์มัลไรเซชันดังนั้นจึงถือว่าการใช้วิธีค่าพิสัยระหว่างควอร์ไทล์และมากที่สุดน้อยสุดในการนอร์มัลไรเซชันเป็นวิธีพื้นฐานในการทดลอง

2) การทดลองเพื่อเปรียบเทียบประสิทธิภาพแบบจำลองที่นำเสนอกับโครงข่ายประสาทเทียมเชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้นได้แก่แบบจำลองลำดับที่หนึ่งถึงเจ็ดในตารางที่ 8

3) การทดลองเพื่อเปรียบเทียบประสิทธิภาพแบบจำลองที่นำเสนอกับแบบจำลองการเรียนรู้ของเครื่องได้แก่แบบจำลองลำดับที่แปดถึงสิบเอ็ดในตารางที่ 7

ตารางที่ 7 ชื่อย่อและชื่อเต็มแบบจำลองในการทดลอง

ลำดับ	ชื่อย่อแบบจำลอง	ชื่อแบบจำลอง
1	ABDLSTM&ACNN	โครงข่ายประสาทเทียมหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางร่วมกับกลไกจุดสนใจผสานกับโครงข่ายประสาทเทียมคอนโวลูชันร่วมกับกลไกจุดสนใจ
2	ACNN	โครงข่ายประสาทเทียมคอนโวลูชันร่วมกับกลไกจุดสนใจ
3	CNN	โครงข่ายประสาทเทียมคอนโวลูชัน
4	ABD-LSTM	โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางกับกลไกจุดสนใจ
5	BD-LSTM	โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง
6	LSTM	โครงข่ายหน่วยความจำระยะสั้นแบบยาว
7	ANN-MLP	โครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้น

ตารางที่ 8 ชื่อย่อและชื่อเต็มแบบจำลองในการทดลอง (ต่อ)

ลำดับ	ชื่อย่อแบบจำลอง	ชื่อแบบจำลอง
8	RF	ป่าแบบสุ่ม
9	LR	ถดถอยโลจิสติก
10	XGBoost	เอ็กซ์ทรีมเกรเดียนต์บูตดิ้ง
11	LightGBM	ไลท์จีบีเอ็ม

4.5 การแบ่งข้อมูลในการทดลอง

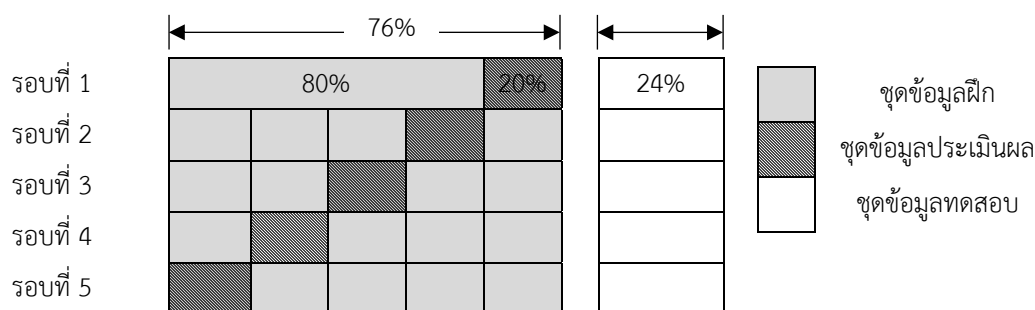
ข้อมูลที่ใช้ในการทดลองประกอบด้วย 2 ชุดข้อมูล จากรีโมทคอมเพรสเซอร์ชนิดบีและชนิดดี โดยแต่ละชุดข้อมูล ถูกแบ่งออกเป็น ชุดข้อมูลฝึก ชุดข้อมูลประเมินผล ชุดข้อมูลทดสอบ ตามสถิติในตารางที่ 8 และ 9

ตารางที่ 9 จำนวนข้อมูลที่ใช้ในการทดลองสำหรับชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบี

หน่วย	ชุดข้อมูลฝึก	ชุดข้อมูลประเมินผล	ชุดข้อมูลทดสอบ	รวม
เหตุการณ์	604	158	248	1010
แถว	2,491,560	696,040	1,014,850	4,202,450
อัตราส่วน	59%	17%	24%	100%

ตารางที่ 10 จำนวนข้อมูลที่ใช้ในการทดลองสำหรับชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดดี

หน่วย	ชุดข้อมูลฝึก	ชุดข้อมูลประเมินผล	ชุดข้อมูลทดสอบ	รวม
เหตุการณ์	273	71	111	455
แถว	1,285,110	345,510	511,230	2,141,850
อัตราส่วน	60%	16%	24%	100%



รูปที่ 27 การทดลอง 5 รอบ ในแต่ละรอบ กล่องเทาอ่อนแสดงสัดส่วนชุดข้อมูลฝึก กล่องสีเทาเข้มแสดงสัดส่วนชุดข้อมูลประเมินผล และกล่องสีขาวแสดงสัดส่วนชุดข้อมูลทดสอบ

การทดลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองบนข้อมูลที่ผ่านกระบวนการลดข้อมูลและการแปลงเมทริกซ์ลำดับเวลาสามมิติ เพื่อหลีกเลี่ยงความคลาดเคลื่อนของชุดข้อมูลประมินที่มีเพียงร้อยละ 17 และ ร้อยละ 16 ในชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบีและดีตามลำดับ ผู้วิจัยออกแบบการทดลองโดยแต่ละแบบจำลอง จะทำการทดลอง 5 รอบ โดยนำเอาชุดข้อมูลประมินผลมาต่อกับชุดข้อมูลฝึกเพื่อรักษาลำดับเวลาของข้อมูลจากนั้น ทำการแบ่งข้อมูลแบ่งข้อมูลแต่ละรอบให้มีชุดข้อมูลฝึกและชุดข้อมูลประมินผลอัตราส่วน 80:20 โดยใช้ชุดข้อมูลทดสอบที่เป็นข้อมูลอนุกรมเวลาที่เกิดจากหลังชุดข้อมูลฝึกและชุดข้อมูลประมินผล ทดสอบในการทดสอบทั้ง 5 รอบ โดยแสดงวิธีการแบ่งข้อมูลไว้ในรูปที่ 27 โดยอัตราส่วนของคลาสบวกและคลาสลบแสดงในตารางที่ 10

ตารางที่ 11 สถิติข้อมูลเมื่อแบ่งตามคลาสของชุดข้อมูลฝึก ชุดข้อมูลประมินและชุดข้อมูลทดสอบ

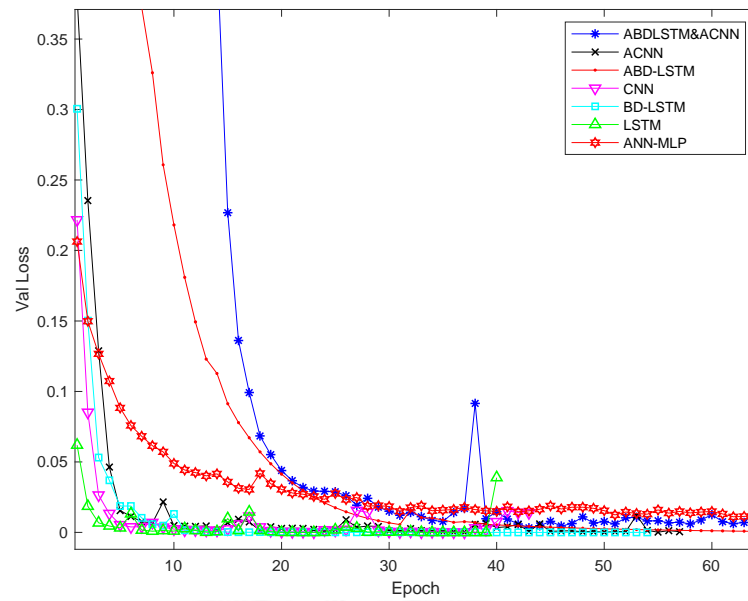
ชนิด	หน่วย	ชุดข้อมูลฝึก	ชุดข้อมูลประมิน	ชุดข้อมูลทดสอบ	รวม
บี	จำนวนเหตุการณ์	604	158	248	1010
	จำนวนแถว	2,491,560	696,040	1,014,850	4,202,450
	อัตราส่วน	59%	17%	24%	100%
	อัตราส่วนคลาสลบ	59%	17%	24%	76%
	อัตราส่วนคลาสบวก	60%	15%	25%	24%
ดี	จำนวนเหตุการณ์	273	71	111	455
	จำนวนแถว	1,285,110	345,510	511,230	2,141,850
	อัตราส่วน	60%	16%	24%	100%
	อัตราส่วนคลาสลบ	60%	16%	24%	81%
	อัตราส่วนคลาสบวก	60%	15%	25%	19%

4.6 การฝึก (Training)

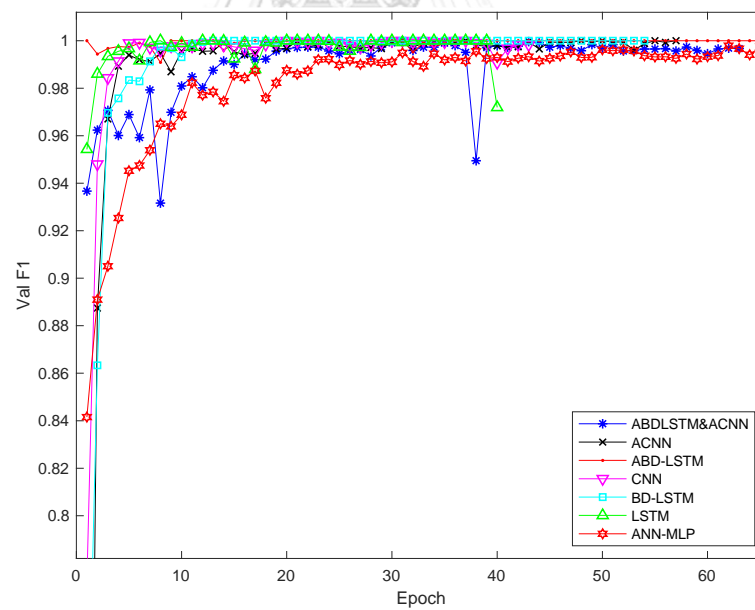
การฝึกแบบจำลองถือเป็นกระบวนการหลักที่สำคัญ วิทยานิพนธ์นี้มีการทดลองเปรียบเทียบประสิทธิภาพ แบบจำลองการเรียนรู้เชิงลึกและการเรียนรู้ของเครื่อง ดังนั้นจึงมีวิธีการฝึกแบบจำลอง 2 วิธีดังต่อไปนี้

4.6.1 การฝึกแบบจำลองการเรียนรู้เชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้น

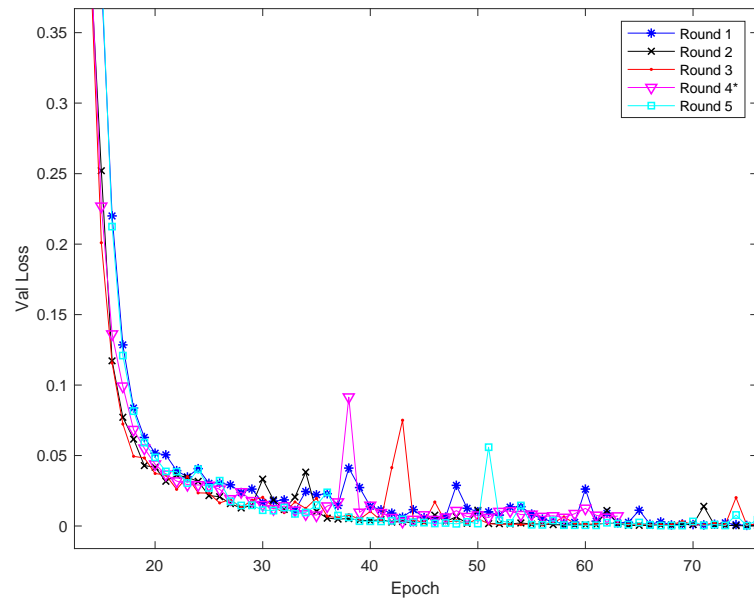
ผู้วิจัยเลือกตัดจำนวนแถวของข้อมูลนำเข้าที่ละ 48 แถวช่วงเวลา หรือ 48 ช่วงเวลา ซึ่งถูกสกัดจากข้อมูลแต่ละจุดข้อมูลผ่านกระบวนการสกัดคุณลักษณะเพื่อลดข้อมูลทุก ๆ 20 แถว โดยกำหนดขนาดแต่ละรุ่น (batch size) เป็น 420 ซึ่งได้จากการทดลองปรับปรุงพารามิเตอร์ สำหรับการเลือกแบบจำลอง ได้มีเงื่อนไขแต่ละรอบการเรียนรู้ (epoch) เพื่อเปรียบเทียบค่าจากฟังก์ชันต้นทุนบนชุดข้อมูลประเมินผล (validation loss) ในรอบการเรียนรู้ปัจจุบัน กับรอบการเรียนรู้ที่ผ่านมาโดยจะบันทึกค่าน้ำหนักที่ดีขึ้นจากค่าของฟังก์ชันต้นทุนที่ลดลง โดยกำหนดจำนวนรอบการเรียนรู้สูงสุดที่ 1,000 รอบการเรียนรู้ ตัวอย่างเช่น ในรูปที่ 28 แสดงการเปรียบเทียบค่าฟังก์ชันต้นทุนระหว่างแบบจำลองที่นำเสนอกับแบบจำลองโครงข่ายประสาทเทียมเชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้นบนชุดข้อมูลประเมินผลของรีโมทคอมเพรสเซอร์ชนิดบี รูปที่ 29 แสดงการเปรียบเทียบด้วยค่า F1 ผู้วิจัยใช้การหยุดก่อน (early stopping) เพื่อหยุดการฝึกเมื่อค่าจากฟังก์ชันต้นทุนบนชุดข้อมูลประเมินผลมีค่าไม่เปลี่ยนแปลงน้อยกว่าค่าขีดแบ่งที่กำหนด 30 ครั้ง หรือครบจำนวนรอบการเรียนรู้สูงสุด โดยทำการฝึกแบบจำลอง 5 รอบ แสดงในรูปที่ 30 โดยใช้ชุดข้อมูลฝึกชุดข้อมูลประเมินผล และชุดข้อมูลทดสอบที่อธิบายวิธีการแบ่งไว้ในหัวข้อที่ 4.5 เมื่อยุติการฝึกจะใช้ค่าน้ำหนักของแบบจำลองในรอบการเรียนรู้ที่มีค่าฟังก์ชันต้นทุนต่ำที่สุด (minimum validation loss) แสดงใน รูปที่ 31 ในการทดสอบด้วยชุดข้อมูลทดสอบ แบบจำลองที่ได้ค่า F1 มากที่สุดจะถูกเลือกไปใช้งานจริงและค่าเฉลี่ย F1 ทั้ง 5 รอบจะใช้ในการเปรียบเทียบประสิทธิภาพของแบบจำลองต่อวิธีการเลือกคุณลักษณะ



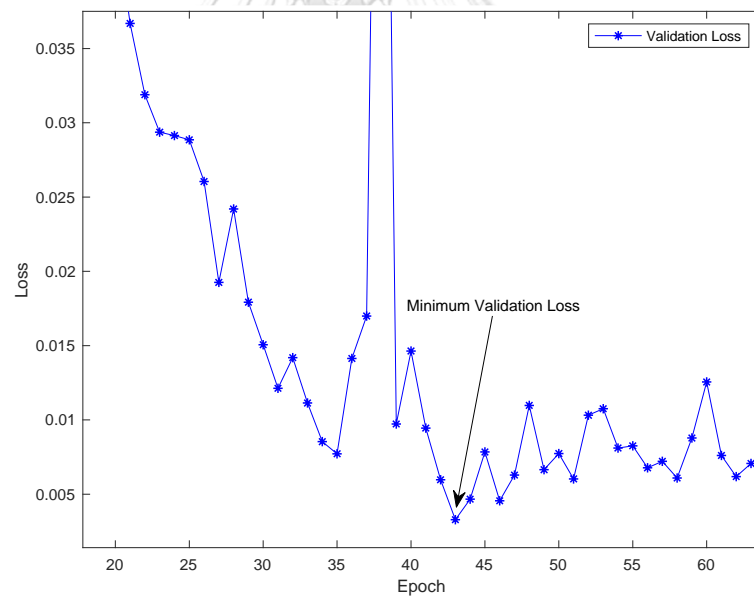
รูปที่ 28 แสดงการเปรียบเทียบค่าฟังก์ชันต้นทุนระหว่างแบบจำลองที่นำเสนอ กับแบบจำลองโครงข่ายประสาทเทียมเชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้นบนชุดข้อมูลประเมินผลของรีโมทคอมเพรสเซอร์ชนิดบี



รูปที่ 29 แสดงการเปรียบเทียบค่า F1 ระหว่างแบบจำลองที่นำเสนอ กับแบบจำลองโครงข่ายประสาทเทียมเชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้นบนชุดข้อมูลประเมินผลของรีโมทคอมเพรสเซอร์ชนิดบี



รูปที่ 30 กราฟแสดงการเปรียบเทียบค่าฟังก์ชันต้นทุนต่อรอบการเรียนรู้ระหว่างแบบจำลองที่นำเสนอในแต่ละรอบการทดลองบนชุดข้อมูลประเมินผลของริโมทคอมเพรสเซอร์ชนิดบี



รูปที่ 31 กราฟแสดงค่าฟังก์ชันต้นทุนต่อรอบการเรียนรู้ของรอบการทดลองที่ดีที่สุดโดยจะใช้ค่าน้ำหนักในรอบการเรียนรู้ที่ได้ค่าฟังก์ชันต้นทุนที่ต่ำที่สุดในการประเมินประสิทธิภาพของแบบจำลองด้วยชุดข้อมูลทดสอบ

การเก็บข้อมูลใช้ภาษา SQL บนฐานข้อมูล Microsoft SQL Server 2012 การสร้างโปรแกรมการประมวลผลข้อมูลก่อนหน้าใช้ ภาษา Python PySpark และ Pandas การลดมิติข้อมูลใช้ SparkML ซึ่งโดยการประมวลผลใช้สปาร์ตครัสเตอร์ (Spark Clusters) สำหรับการประมวลผล

ข้อมูลที่ใช้ซีพียูและหน่วยความจำหลัก ชนิดที่ 1 ในตารางที่ 10 การสร้างโปรแกรมการฝึกแบบจำลองใช้ภาษา Python และ Keras Tensorflow การประเมินผลใช้ Sklearn การบันทึกผลการทดลองใช้ MLFlow ประมวลผลข้อมูลบนสเปกครัสเตอร์ประเภทซีพียู ชนิดที่ 2 ในตารางที่ 10

4.6.2 การฝึกแบบจำลองการเรียนรู้ของเครื่อง

การฝึกแบบจำลองที่สร้างจากอัลกอริทึมป่าแบบสุ่ม ถดถอยโลจิสติก เอ็กซ์ทรีมเกรเดียนต์บูตติงและโลท์จีบีเอ็ม วิทยานิพนธ์นี้ใช้ชุดข้อมูลฝึกและชุดข้อมูลประเมินผลในการฝึกแบบจำลองโดยการใช้การตรวจสอบไขว้ 5 พับ (5-fold cross-validation) จากนั้นทำการเลือกจากนั้นทำการเลือกชุดพารามิเตอร์ที่ดีที่สุดด้วยวิธีการค้นหาแบบกริด (Grid Search) จากนั้นเลือกชุดค่าพารามิเตอร์ที่ดีที่สุดในการฝึกแบบจำลองขั้นสุดท้ายโดยจะแบ่งข้อมูลตามขั้นตอนวิธีที่แสดงในรูปที่ 27 โดยทำการฝึก 5 รอบ โดยแต่ละรอบจะใช้ชุดข้อมูลทดสอบเดียวกันกับแบบจำลองการเรียนรู้เชิงลึก

การสร้างแบบจำลองป่าแบบสุ่ม ถดถอยโลจิสติกใช้ภาษา Python และ SparkML เอ็กซ์ทรีมเกรเดียนต์บูตติงและโลท์เอ็กซ์ทรีมเกรเดียนต์บูตติง ใช้ XGBoost และ LightGBM library โดยการประมวลผลใช้สเปกครัสเตอร์ชนิดที่ 1 ในตารางที่ 11

ตารางที่ 12 รายละเอียดสเปกครัสเตอร์ที่ใช้ในการประมวลผลข้อมูลก่อน การพัฒนาขั้นตอนวิธีและการทดลอง

ชนิด	รายละเอียดสเปกครัสเตอร์
1	1 Spark Cluster (5.4 ML includes Apache Spark 2.4.0, Scala 2.11) Driver node: RAM 256 GB memory, 32 cores Worker node: RAM 256 GB memory, 32 cores, maximum 8 worker nodes
2	4 Spark Clusters (5.4 ML includes Apache Spark 2.4.0, GPU, Scala 2.11) Driver node: RAM 112 GB memory, 1 GPU (NVIDIA Tesla V100-PCIE)

4.7 การประเมินผล (Evaluation)

การประเมินความสามารถของแบบจำลองในการทำนายหรือการจำแนกคลาสบวกหรือคลาสนลบใช้ การวัดความถูกต้อง (accuracy) ค่าความแม่นยำ (precision) ค่าความระลึก (recall) และ F1 score ซึ่งคำนวณได้จากตารางที่ 11 เมทริกซ์ความสับสน (confusion matrix) ดังต่อไปนี้

เมทริกซ์ความสับสน แสดงจำนวนของการทำนายและผลเฉลยจากฉลาก ที่ทราบอยู่แล้วใส่ลงในตารางขนาด $n \times n$ โดยที่ n เป็นจำนวนของคลาสที่เป็นไปได้ ในที่นี้ใช้ $n = 2$ ดัง ตารางที่ 12 ซึ่งเป็นเมทริกซ์จัตุรัส ประกอบด้วยคอลัมน์และแถวที่แสดงจำนวนอินสแตนซ์เป็นอัตราส่วนสัมบูรณ์ "คลาสจริง" (actual class) เทียบกับ

"คลาสที่ทำนาย" (predicted class) โดยให้ positive (P) เป็นฉลากของคลาสที่ 1 หรือเรียกว่าคลาสบวก และ Negative (N) หรือเรียกว่าคลาสลบ เป็นฉลากของคลาสที่ 2 และสามารถใช้กับแบบจำลองประเภทหลายคลาสได้เช่นกัน

ตารางที่ 13 เมทริกซ์ความสับสน

		Predicted Class	
		Negative	Positive
Actual Class	Negative	True Negative (TN)	False Negative (FN)
	Positive	False Positive (FP)	True Positive (TP)

True Positive (TP) คือ จำนวนตัวอย่างที่เป็นบวก และแบบจำลองทำนายว่าเป็นบวก

False Positive (FP) คือ จำนวนตัวอย่างที่ไม่เป็นบวก และแบบจำลองทำนายว่าเป็นบวก

False Negative (FN) คือ จำนวนตัวอย่างที่ไม่เป็นลบ และแบบจำลองทำนายว่าเป็นลบ

True Negative (TN) คือ จำนวนตัวอย่างที่เป็นลบ และแบบจำลองทำนายว่าเป็นลบ

Precision ความแม่นยำของระบบโดยสนใจแค่ในส่วนของการทำนายตามสมการที่ 4.17

$$Precision = \frac{TP}{TP + FP} \quad (4.17)$$

Recall ความแม่นยำที่สนใจในส่วนของความเป็นจริงตามสมการที่ 4.18

$$Recall = \frac{TP}{TP + FN} \quad (4.18)$$

ค่าความแม่นยำ คือ การคำนวณจำนวนคำตอบที่ถูกต้องเทียบกับจำนวนคำตอบทั้งหมดที่นำไปให้ แบบจำลอง ทำการตอบ ตามสมการที่ 4.19

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (4.19)$$

F1 คือ ค่าเฉลี่ยแบบ harmonic mean ระหว่าง precision และ recall ตามสมการที่ 4.20

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4.20)$$

ตารางที่ 14 โครงสร้าง ค่าพารามิเตอร์ต่าง ๆ ของแบบจำลองในการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าในแต่ละแบบสำหรับโครงข่ายประสาทเทียมการเรียนรู้เชิงลึก และโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้น

ANN-MLP	LSTM	BD-LSTM	ABD-LSTM	CNN	ACNN	ABDLSTM&ACNN*	
DN(512) D(0.5)	L(512) D(0.5)	BL(512) D(0.5)	BL(512) BN ReLU	C(128,8) BN ReLU	C(512,16) BN ReLU	BL(512) BN ReLU	C(256,18) BN ReLU
DN(256,ReLU) D(0.5)	L(256) D(0.5)	BL(256) D(0.5)	DN(256,L2=0.5) D(0.5)	C(256,5) BN ReLU	C(256,8) BN ReLU	DN(256,L2=0.5) D(0.5)	C(512,16) BN ReLU
			BL(256) BN ReLU DN(32, L2) D(0.5)	C(256,8) BN ReLU C(64,3) BN ReLU	C(256,8) BN ReLU C(128,5) BN ReLU	BL(256) BN ReLU DN(256,L2=0.5) D(0.5)	C(128,8) BN ReLU C(64,5) BN ReLU
			ATT	M	C(64,3) BN ReLU	BL(128) BN ReLU	C(32,3) BN ReLU
			DN(32)	DN(100,ReLU)	C(32,3) BN ReLU	DN(32, L2) D(0.5)	
					ATT	ATT	ATT
					DN(32) DN(100,ReLU)	CC DN(100,softmax) D(0.5) DN(48)	
DN(1) Sigmoid	DN(1) Sigmoid	DN(1) Sigmoid	DN(1) Sigmoid	DN(1) Sigmoid	DN(1) Sigmoid	DN(1) Sigmoid	

4.8 การกำหนดค่าต่าง ๆ ของแบบจำลอง

ในการดำเนินการทดลอง ผู้วิจัยได้ทำการกำหนดโครงสร้าง ค่าพารามิเตอร์ต่าง ๆ ให้กับแบบจำลองในแต่ละแบบได้แก่ ขนาดของหน่วยชั้นซ่อน มิติ และฟังก์ชัน ต่าง ๆ สำหรับโครงข่ายประสาทเทียมการเรียนรู้เชิงลึกและโครงข่ายประสาทเทียมที่ใช้ในการทดลอง ตารางที่ 13 อธิบายถึงรายละเอียดโครงสร้างแบบจำลองนั้น ๆ โดยระบุค่าพารามิเตอร์ที่ดีที่สุดจากการทดลองในฟังก์ชันหรือองค์ประกอบนั้น ๆ โดยชื่อเต็มของอักขระย่อแต่ละฟังก์ชันหรือองค์ประกอบแสดงในตารางที่ 13 นอกจากนี้เทคนิคที่ใช้เหมือนกันทุกแบบจำลองคือ การหาค่าเหมาะสมที่สุดได้ใช้ Adam สโตแคสติกเกรเดียนต์เดสเซนส์เพื่อลดต้นทุน loss binary cross entropy ให้น้อยที่สุดโดยกำหนดค่า * คือแบบจำลองที่นำเสนอ

ตารางที่ 15 ชื่อเต็มของอักขระย่อแต่ละฟังก์ชันหรือองค์ประกอบของ โครงสร้าง ค่าพารามิเตอร์ต่าง ๆ ของแบบจำลองโครงข่ายประสาทเทียมการเรียนรู้เชิงลึกและโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้นในตารางที่ ตารางที่ 13

อักขระย่อ	ชื่อเต็ม
BL	Bidirectional LSTMs(Units)
DN	Dense (Units)
D	Dropout (value)
C	Conv1D (Filter, Kernel size)
ReLU	ReakyReLU
L	LSTM (Units)
BN	Batch Normalization
GAP	Global Average Pooling1D
GMP	GlobalMaxPooling1D
M	MaxPooling1D

โดยค่าพารามิเตอร์ที่ดีที่สุดของแบบจำลองต้นไม่ตัดสินใจเอ็กซ์ทริมเกรเดียนต์บูทดิ้ง โลทจีบีเอ็ม ถดถอยโลจิสติกและป่าแบบสุ่มแสดงในตารางที่ 15-17

ตารางที่ 16 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ โดยใช้แบบจำลองต้นไม้ตัดสินใจเอ็กซ์ทรีมเกรเดียนต์บูตติงและโลทเอ็กซ์จีบีเอ็ม

ลำดับ	ชื่อพารามิเตอร์	ค่า
1	Boosting	GBDT
2	Learning rate	0.01
3	Num Leaves	80
4	Num Threads	50
5	Metric	Binary Logloss
6	Feature fraction	0.8
7	Number Boost Round	10000
8	Estimators	500

ตารางที่ 17 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ โดยใช้แบบจำลองถดถอยโลจิสติก

ลำดับ	ชื่อพารามิเตอร์	ค่า
1	Estimator	Logistic Regression
2	Estimator Param Maps Length	9
3	Evaluator	Binary Classification Evaluator
4	Regularization	0.6
5	Max Iteration	10

ตารางที่ 18 ค่าพารามิเตอร์ที่ดีที่สุดสำหรับแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ โดยใช้แบบจำลองป่าแบบสุ่ม

ลำดับ	ชื่อพารามิเตอร์	ค่า
1	Estimator	Random Forest Classifier
2	Estimator Param Maps Length	2
3	Evaluator	Binary Classification Evaluator
4	Num of Trees	20

บทที่ 5

ผลการดำเนินงาน

จากแนวทางและวิธีการดำเนินงานที่กล่าวไปในบทที่ผ่านมา ในบทนี้จะกล่าวถึงผลการดำเนินงาน ประกอบด้วยผลการใช้เทคนิคลดข้อมูล ผลการทดลองและการวิเคราะห์ผลในแต่ละการทดลอง

5.1 ผลการใช้เทคนิคลดข้อมูล

ผลการทดลองของหน่วยความจำหลักโดยประมาณจากการใช้เทคนิคการลดข้อมูล โดยทำการเปรียบเทียบการใช้คุณลักษณะที่ได้จากกระบวนการเลือกคุณลักษณะ (LGBM) กับการใช้คุณลักษณะที่ได้จากการประมวลผลก่อนหน้า หรือวิธีพื้นฐาน (Baseline) การใช้คุณลักษณะจากการวิเคราะห์องค์ประกอบสำคัญ (PCA) 8 องค์ประกอบ การทดลองใช้คุณลักษณะที่ได้จากการประมวลผลก่อนหน้าและจากการวิเคราะห์องค์ประกอบสำคัญ (Baseline+PCA) พบว่าสามารถลดพื้นที่หน่วยความจำโดยประมาณที่ 66.52% และ 33.04% สำหรับการประมวลผลข้อมูลชุดฝึกของชุดข้อมูลรีโมทคอทเพรสเซอร์ชนิดบีและดี แสดงในตารางที่ 18 และ 19 ตามลำดับ

ตารางที่ 19 ผลการใช้เทคนิคการลดข้อมูลและการเลือกคุณลักษณะด้วยไลทจีบีเอ็ม การวิเคราะห์องค์ประกอบสำคัญ และวิธีพื้นฐานร่วมกับการวิเคราะห์องค์ประกอบสำคัญ เปรียบเทียบกับวิธีพื้นฐานบนชุดข้อมูลฝึกกรีมทคอมเพรสเซอร์ชนิดบี โดย ‘*’ หมายถึงวิธีการเลือกคุณลักษณะที่นำเสนอ

ชื่อวิธี	วิธีการเลือกคุณลักษณะ/จำนวนคุณลักษณะ			
	Baseline	PCA	Baseline+PCA	LGBM*
จำนวนคุณลักษณะ	56	8	64	25
หน่วยความจำก่อนลดข้อมูล	139,473,656	19,924,808	159,398,464	62,265,025
หน่วยความจำหลังลดข้อมูล	104,606,040	14,943,720	119,549,760	46,699,125
เปอร์เซ็นต์การปรับปรุงจากก่อนลดข้อมูลของวิธีพื้นฐาน		89.29%	14.29%	66.52%

ตารางที่ 20 ผลการใช้เทคนิคการลดข้อมูลและการเลือกคุณลักษณะด้วยไลท์จีบีเอ็ม การวิเคราะห์องค์ประกอบสำคัญ และวิธีพื้นฐานร่วมกับการวิเคราะห์องค์ประกอบสำคัญ เปรียบเทียบกับวิธีพื้นฐานบนชุดข้อมูลฝักริโมทคอมเพรสเซอร์ชนิดดี

ชื่อวิธี	วิธีการเลือกคุณลักษณะ/จำนวนคุณลักษณะ			
	Baseline	PCA	Baseline+PCA	LGBM*
จำนวนคุณลักษณะ	56	8	64	25
หน่วยความจำก่อนลดข้อมูล	71,912,456	10,273,208	82,185,664	32,103,775
หน่วยความจำหลังลดข้อมูล	107,869,440	10,273,280	82,186,240	32,104,000
เปอร์เซ็นต์การปรับปรุงจากก่อนลดข้อมูลของวิธีพื้นฐาน		85.71%	-14.29%	55.36%

คุณลักษณะของวิธี Baseline แสดงในภาคผนวก ค ตารางที่ ค.1-ค.2 คุณลักษณะของวิธี PCA แสดงในภาคผนวก ค ตารางที่ ค.4 ลำดับที่ 29-36 คุณลักษณะของวิธี Baseline+PCA แสดงในภาคผนวก ค ตารางที่ 3-4 คุณลักษณะของวิธี LGBM* แสดงใน ภาคผนวก ข ตารางที่ ข.3 ลำดับที่ 1-25 เมื่อ ‘*’ คือ วิธีการเลือกคุณลักษณะที่นำเสนอ

5.2 การทดลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองและวิธีการเลือกคุณลักษณะ

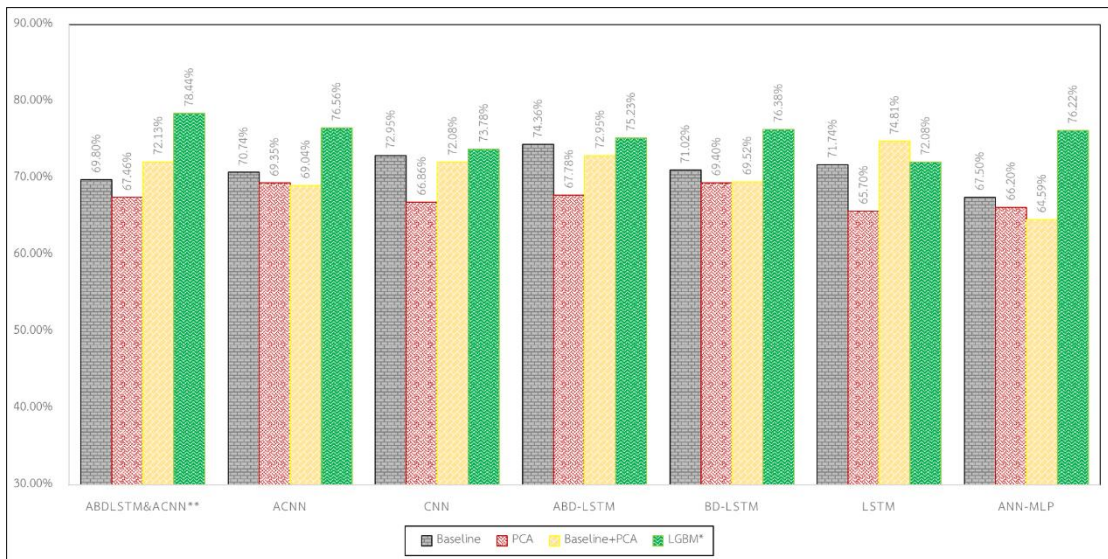
การทดลองนี้มีวัตถุประสงค์ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองด้วยค่า F1 เฉลี่ยต่อวิธีการเลือกคุณลักษณะ ซึ่งผลการทดลองแสดงในตารางที่ 20 โดยเทคนิคการเลือกคุณลักษณะที่นำเสนอสามารถเพิ่มประสิทธิภาพทั้งการลดมิติของข้อมูลและความแม่นยำของแบบจำลองที่นำเสนอและกลุ่มแบบจำลองการเรียนรู้เชิงลึกยกเว้นแบบจำลองหน่วยความจำระยะสั้นแบบยาวที่ได้ค่า F1 เฉลี่ยที่น้อยกว่าการเลือกคุณลักษณะแบบ Baseline+PCA

การวิเคราะห์องค์ประกอบสำคัญโดยมีสามองค์ประกอบสำคัญถูกเลือกในกระบวนการเลือกคุณลักษณะ แสดงถึงความสอดคล้องกันของทั้งสองวิธี ถึงแม้ว่าการทำพีซีเอสามารถลดมิติของข้อมูลได้อย่างมาก ช่วยประหยัดเวลาในการฝึกแบบจำลอง แต่ผลจากการทดลองพบว่า ได้ประสิทธิภาพที่ดีกว่าวิธีการเลือกคุณลักษณะแบบอื่นยกเว้นในแบบจำลอง RF และ LR ที่ทำการทดสอบบนชุดข้อมูลริโมทคอมเพรสเซอร์ชนิดดี

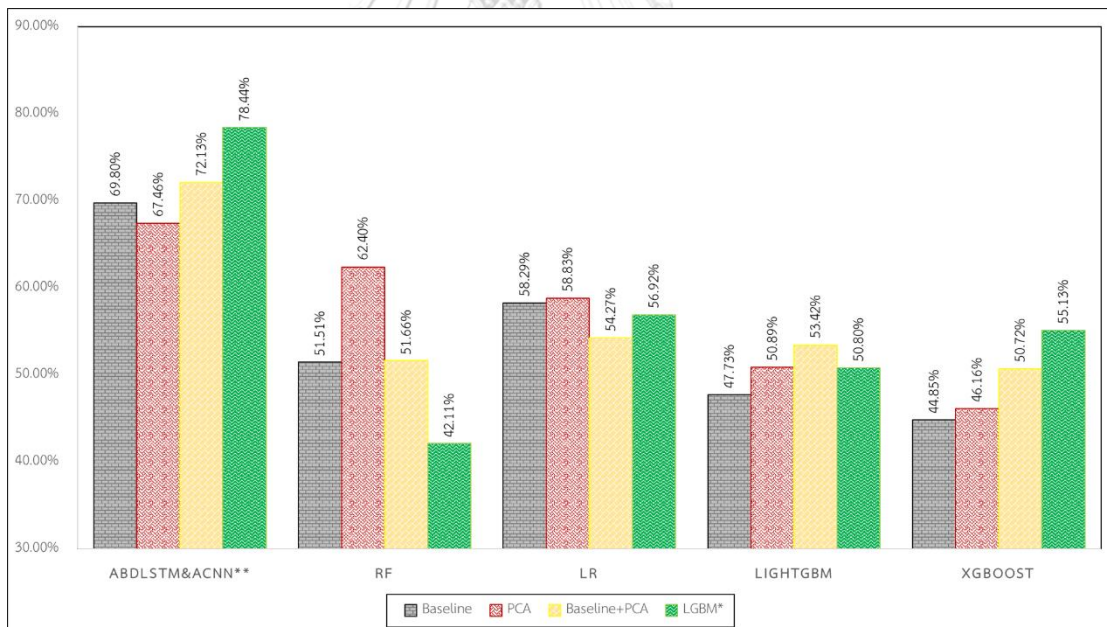
ตารางที่ 21 F1 เฉลี่ยของแบบจำลองวิธีการเลือกคุณลักษณะบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบีและดี โดย ‘*’ หมายถึงวิธีการเลือกคุณลักษณะที่นำเสนอหรือค่า F1 สูงสุด และ ‘**’ หมายถึงแบบจำลองที่นำเสนอ

Model/RC Type	Baseline		PCA		Baseline+PCA		LGBM*	
	B	D	B	D	B	D	B	D
ABDLSTM&ACNN**	89.26%	69.80%	65.93%	67.46%	90.57%	72.13%	93.80%	78.44%
ACNN	90.41%	70.74%	67.90%	69.35%	89.45%	69.04%	90.54%	76.56%
CNN	86.41%	72.95%	66.71%	66.86%	88.55%	72.08%	88.69%	73.78%
ABD-LSTM	86.16%	74.36%	66.96%	67.78%	84.93%	72.95%	89.58%	75.23%
BD-LSTM	80.60%	71.02%	65.20%	69.40%	80.31%	69.52%	85.50%	76.38%
LSTM	82.13%	71.74%	65.34%	65.70%	82.01%	74.81%	84.78%	72.08%
ANN-MLP	89.31%	67.50%	65.27%	66.20%	88.25%	64.59%	88.41%	76.22%
RF	78.77%	51.51%	60.33%	62.40%	78.80%	51.66%	74.27%	42.11%
LR	85.39%	58.29%	58.69%	58.83%	84.04%	54.27%	80.53%	56.92%
LightGBM	85.16%	47.73%	53.40%	50.89%	85.94%	53.42%	82.57%	50.80%
XGBoost	84.64%	44.85%	58.14%	46.16%	85.07%	50.72%	80.21%	55.13%

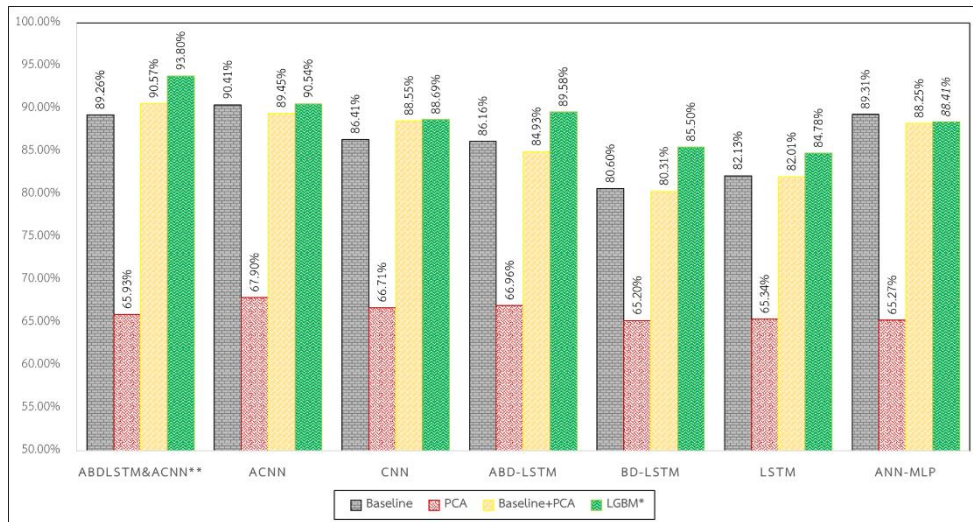
วิทยานิพนธ์นี้แยกการวิเคราะห์ประสิทธิภาพของแบบจำลองและวิธีการเลือกคุณลักษณะเป็นสองกลุ่มคือการเปรียบเทียบประสิทธิภาพของแบบจำลอง ABDLSTM&ACNN กับ ACNN CNN ABD-LSTM BD-LSTM LSTM และ ANN-MLP จากกราฟแท่งในรูปที่ 32 - 33 บนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบีและดี และกลุ่มที่สองคือการเปรียบเทียบประสิทธิภาพแบบจำลอง ABDLSTM&ACNN กับแบบจำลองการเรียนรู้ของเครื่อง RS LR XGBoost และ LightGBM ในรูปที่ 34 - 35 บนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบีและดี



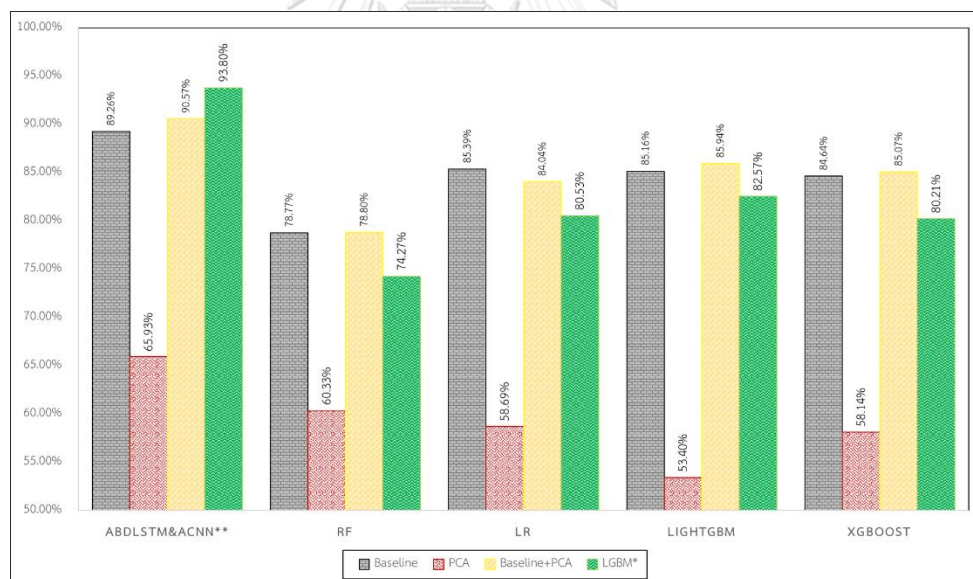
รูปที่ 32 ค่า F1 เฉลี่ยบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบี ในรูปแบบกราฟแท่งจากการทดลองแบบจำลองที่นำเสนอเกี่ยวกับวิธีการเลือกคุณลักษณะกับแบบจำลองการเรียนรู้เชิงลึกและแบบจำลองโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้น โดย ‘**’ หมายถึงแบบจำลองที่นำเสนอ



รูปที่ 33 ค่า F1 เฉลี่ยบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดดี ในรูปแบบกราฟแท่งจากการทดลองแบบจำลองที่นำเสนอเกี่ยวกับวิธีการเลือกคุณลักษณะกับแบบจำลองการเรียนรู้เชิงลึกและแบบจำลองโครงข่ายประสาทเทียมวิธีเพอร์เซ็ปตรอนหลายชั้น โดย ‘**’ หมายถึงแบบจำลองที่นำเสนอ



รูปที่ 34 ค่า F1 เฉลี่ยบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบี ในรูปแบบกราฟแท่งจากการทดลองแบบจำลองที่นำเสนอกับวิธีการเลือกคุณลักษณะกับแบบจำลองการเรียนรู้เชิงลึกที่นำเสนอกับแบบจำลองการเรียนรู้ของเครื่อง โดย “**” หมายถึงแบบจำลองที่นำเสนอ



รูปที่ 35 ค่า F1 เฉลี่ยบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดดี ในรูปแบบกราฟแท่งจากการทดลองแบบจำลองที่นำเสนอกับวิธีการเลือกคุณลักษณะกับแบบจำลองการเรียนรู้เชิงลึกที่นำเสนอกับแบบจำลองการเรียนรู้ของเครื่อง โดย “**” หมายถึงแบบจำลองที่นำเสนอ

วิธีการเลือกคุณลักษณะ LGBM และแบบจำลองที่นำเสนอ ABDLSTM&ACNN มีประสิทธิภาพสูงที่สุดโดยได้ค่า F1 เฉลี่ย 93.80% และ 78.44% บนชุดข้อมูลทดสอบรีโมทคอมเพรสเซอร์ชนิดบีและดีตามลำดับ

ดังนั้นในการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าโดยมีข้อมูลนำเข้ามีลักษณะเป็นอนุกรมเวลาแบบหลายตัวแปรจึงเหมาะกับแบบจำลอง ABDLSTM&ACNN และวิธีการเลือกคุณลักษณะ LGBM เมื่อเทียบกับแบบจำลองพื้นฐานจากการเรียนรู้ของเครื่องและการเรียนรู้เชิงลึกทั้ง 10 แบบจำลองและวิธีการเลือกคุณลักษณะวิธีอื่น

5.3 การเปรียบเทียบประสิทธิภาพของแบบจำลองและการเลือกแบบจำลอง

จากผลการทดลองพบว่าการเพิ่มประสิทธิภาพของแบบจำลองโครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางด้วยกลไกจุดสนใจสามารถเพิ่มค่า F1 ถึง 4.08% บนชุดข้อมูลบีแต่ไม่สามารถเพิ่มประสิทธิภาพให้กับแบบจำลองสำหรับชุดข้อมูลดีและการใช้เพิ่มกลไกจุดสนใจกับโครงข่ายประสาทเทียมคอนโวลูชันสามารถเพิ่มประสิทธิภาพได้ 1.8% และ 2.77% บนชุดข้อมูลบีและดีตามลำดับการต่อขนานของโครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางด้วยกลไกจุดสนใจและโครงข่ายประสาทเทียมคอนโวลูชันกับกลไกจุดสนใจสามารถเพิ่มประสิทธิภาพที่ F1 เพิ่มขึ้น 8.3% และ 1.88% บนชุดข้อมูลบีและดี ซึ่งแสดงการเปรียบเทียบประสิทธิภาพที่เพิ่มขึ้นระหว่างแบบจำลองที่นำเสนอกับแบบจำลองอื่น ๆ ที่ใช้ในการทดลอง ในตารางที่ 21-22

ตารางที่ 22 ค่า F1 เฉลี่ยของแต่ละแบบจำลองที่นำเสนอโดยใช้การเลือกคุณลักษณะบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดดีและดี และค่า F1 เฉลี่ยที่เพิ่มขึ้นเมื่อเทียบกับแบบจำลองที่นำเสนอ

	ชุดข้อมูลบี		ชุดข้อมูลดี	
	F1	F1 ที่เพิ่มขึ้น	F1	F1 ที่เพิ่มขึ้น
ABDLSTM&ACNN**	93.80%	-	78.44%	-
ACNN	90.54%	3.26%	76.56%	1.88%
CNN	88.69%	5.11%	73.78%	4.66%
ABD-LSTM	89.58%	4.22%	75.23%	3.21%
BD-LSTM	85.50%	8.30%	76.38%	2.06%
LSTM	84.78%	9.02%	72.08%	6.36%
ANN-MLP	88.41%	5.39%	76.22%	2.22%

ตารางที่ 23 ค่า F1 เฉลี่ยของแต่ละแบบจำลองที่นำเสนอโดยใช้การเลือกคุณลักษณะบนชุดข้อมูลทดสอบของรีโมทคอมเพรสเซอร์ชนิดบีและดี และค่า F1 เฉลี่ยที่เพิ่มขึ้นเมื่อเทียบกับแบบจำลองที่นำเสนอ (ต่อ)

	ชุดข้อมูลบี		ชุดข้อมูลดี	
	F1	F1 ที่เพิ่มขึ้น	F1	F1 ที่เพิ่มขึ้น
RF	74.27%	19.53%	42.11%	36.33%
LR	80.53%	13.27%	56.92%	21.52%
LightGBM	82.57%	11.23%	50.80%	27.64%
XGBoost	80.21%	13.59%	55.13%	23.31%

ตารางที่ 23 แสดงค่า F1 ในแต่ละรอบการทดลอง ของแบบจำลอง ABDLSTM&ACNN โดยรอบที่สี่ในรอบการเรียนรู้ที่ 43 ได้ค่า F1 สูงสุดที่ 97.23% ถูกเลือกให้นำไปใช้ จากตารางที่ 24 สำหรับรีโมทคอมเพรสเซอร์ชนิดดีจะเลือกใช้ค่านี้จากในรอบการทดลองที่สองในรอบการเรียนรู้ที่ 30 ที่ได้ค่า F1 สูงสุดที่ 85.48%

ตารางที่ 24 ค่า F1 ของแบบจำลอง ABDLSTM&ACNN ที่ดีที่สุดบนชุดข้อมูลทดสอบของชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดบี

	รอบที่ 1	รอบที่ 2	รอบที่ 3	รอบที่ 4	รอบที่ 5	ค่าเฉลี่ย
ค่าเปอร์เซ็นต์ Precision	91.71%	93.36%	91.15%	97.23%	95.61%	93.81%
ค่าเปอร์เซ็นต์ Recall	91.99%	93.36%	91.00%	97.26%	95.63%	93.85%
ค่าเปอร์เซ็นต์ F1	91.81%	93.36%	91.04%	97.19%*	95.62%	93.80%

ตารางที่ 25 ค่า F1 ของแบบจำลอง ABDLSTM&ACNN ที่ดีที่สุด บนชุดข้อมูลทดสอบของชุดข้อมูลรีโมทคอมเพรสเซอร์ชนิดดี

	รอบที่ 1	รอบที่ 2	รอบที่ 3	รอบที่ 4	รอบที่ 5	ค่าเฉลี่ย
ค่าเปอร์เซ็นต์ Precision	75.24%	86.19%	81.61%	79.88%	75.51%	79.69%
ค่าเปอร์เซ็นต์ Recall	72.90%	85.73%	81.16%	74.87%	74.87%	77.91%
ค่าเปอร์เซ็นต์ F1	73.49%	85.48%*	81.35%	76.74%	75.16%	78.44%

ตารางที่ 26 การเปรียบเทียบค่า F1 เฉลี่ยของของคลาส 0 และ คลาส 1 ของชุดข้อมูลบี

	Baseline		PCA		Baseline+PCA		LGBM	
	F1_0	F1_1	F1_0	F1_1	F1_0	F1_1	F1_0	F1_1
ABDLSTM&ACNN	92.56%	79.17%	83.89%	11.04%	93.61%	81.28%	95.90%	87.37%
ACNN	93.60%	80.65%	83.12%	21.37%	92.89%	78.93%	93.85%	80.45%
CNN	90.72%	73.23%	83.91%	14.16%	92.38%	76.83%	92.71%	76.40%
ADB-LSTM	90.91%	71.62%	82.56%	19.28%	89.70%	70.37%	93.13%	78.70%
BLSTM	86.18%	63.56%	83.61%	8.93%	86.45%	61.54%	90.42%	70.45%
LSTM	87.94%	64.35%	84.75%	6.02%	87.38%	65.58%	89.47%	70.43%
ANN-MLP	93.24%	77.30%	85.54%	3.31%	92.15%	76.33%	92.81%	74.97%
RF	67.73%	84.57%	44.46%	68.67%	67.83%	84.57%	61.22%	81.13%
LR	78.04%	89.25%	19.18%	79.45%	81.88%	88.19%	70.60%	85.75%
LightGBM	85.12%	85.14%	52.81%	53.10%	85.87%	85.91%	82.56%	82.57%
XGBoost	84.62%	84.63%	57.81%	57.98%	84.98%	85.03%	80.20%	80.21%

ตารางที่ 27 การเปรียบเทียบค่า F1 เฉลี่ยของของคลาส 0 และ คลาส 1 ของชุดข้อมูลดี

	Baseline		PCA		Baseline+PCA		LGBM	
	F1_0	F1_1	F1_0	F1_1	F1_0	F1_1	F1_0	F1_1
ABDLSTM&ACNN	72.91%	72.87%	68.82%	66.66%	68.64%	68.76%	79.19%	78.89%
ACNN	71.66%	71.19%	72.70%	70.98%	70.29%	69.66%	76.88%	76.72%
CNN	72.37%	72.66%	70.52%	68.64%	72.63%	72.35%	73.39%	73.58%
ADB-LSTM	75.07%	74.72%	71.30%	69.50%	73.37%	73.16%	76.59%	75.91%
BLSTM	71.41%	71.22%	72.09%	70.72%	70.64%	70.07%	75.99%	76.18%
LSTM	72.46%	72.10%	69.34%	67.47%	75.01%	74.91%	71.94%	72.01%
ANN-MLP	67.65%	67.58%	69.25%	67.69%	66.81%	65.68%	76.51%	76.36%
RF	52.65%	52.07%	44.44%	43.24%	52.43%	52.04%	44.44%	43.24%
LR	61.77%	59.98%	59.68%	58.27%	57.13%	55.66%	59.68%	58.27%
LightGBM	49.68%	48.69%	51.94%	51.41%	55.47%	54.43%	50.10%	49.09%
XGBoost	48.99%	46.83%	48.99%	46.83%	51.97%	51.34%	48.92%	47.07%

ในตารางที่ 25 แสดงการเปรียบเทียบประสิทธิภาพด้วย F1 เฉลี่ยของ คลาส 1 พบว่าแบบจำลอง LR ได้ค่า F1 สูงสุดด้วยวิธีการเลือกคุณลักษณะพื้นฐาน แต่อย่างไรก็ตามแบบจำลอง LR ค่า F1 คลาส 0 น้อยกว่าวิธีที่นำเสนอสำหรับข้อมูลชุดบีแต่สำหรับชุดข้อมูลดีแบบจำลองที่นำเสนอได้ค่า F1 มากที่สุด ดังแสดงในตารางที่ 26 และตัวอย่างการทำนายแสดงในภาคผนวก ค ในรูปที่ ค.6 - 8

บทที่ 6

สรุปการดำเนินงาน

ในบทนี้จะกล่าวถึงผลการดำเนินงานวิทยานิพนธ์ ประกอบด้วยผลการทดลองขั้นตอนวิธีที่นำเสนอ ปัญหาที่พบในการจัดทำวิทยานิพนธ์ และงานวิจัยในอนาคต

6.1 สรุปการวิจัย

ในวิทยานิพนธ์นี้ได้มุ่งเน้นการปรับโครงสร้างประสาทเทียมเชิงลึกให้สามารถทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าได้ โดยใช้ข้อมูลอนุกรมเวลาหลายตัวแปรโดยได้ดำเนินงานตามขอบเขตดังนี้

1) ทำการทำนายความล้มเหลวล่วงหน้าของรีโมทคอมเพรสเซอร์โดยใช้ข้อมูลอนุกรมเวลาหลายตัวแปรเป็นข้อมูลนำเข้าแบบจำลองโครงข่ายประสาทเทียมหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางร่วมกับกลไกจุดสนใจ ผสานกับโครงข่ายประสาทเทียมคอนโวลูชันร่วมกับกลไกจุดสนใจ

2) การลดปริมาณข้อมูลอนุกรมเวลาโดยใช้กระบวนการลดมิติข้อมูลโดยการวิเคราะห์องค์ประกอบสำคัญ การเลือกคุณลักษณะโดยใช้อัลกอริทึมไลทจีบีเอ็ม และการลดปริมาณของข้อมูลด้วยการสกัดคุณลักษณะสำคัญจากฟังก์ชันทางสถิติ

3) แสดงการเปรียบเทียบประสิทธิภาพการเลือกคุณลักษณะและประสิทธิภาพของแบบจำลองที่นำเสนอกับแบบจำลองโครงข่ายประสาทเทียมเชิงลึกและแบบจำลองการเรียนรู้ของเครื่องอื่น ๆ ที่ได้รับความนิยม

โดยใช้เทคนิคการเพิ่มประสิทธิภาพการทำนาย เช่น โดยการปรับจำนวนชั้น หน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางให้เหมาะสม รวมทั้งการปรับจำนวนชั้นของโครงข่ายประสาทเทียมคอนโวลูชัน การใช้กลไกจุดสนใจ การทำเรกูลาร์ไลเซชันแบบแอลสอง การป้องกันการอิงกับข้อมูลมากเกินไปด้วยการทำดรอปเอาต์ การปรับอัตราการเรียนรู้และขนาดเบตของข้อมูลฝึก ล้วนส่งผลทำให้แบบจำลองที่ได้มีประสิทธิภาพดีขึ้น

สำหรับการทดลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองกับวิธีการเลือกคุณลักษณะ 4 วิธีได้แก่การใช้คุณลักษณะที่ได้จากการประมวลผลก่อนหน้า การใช้คุณลักษณะจากการวิเคราะห์องค์ประกอบสำคัญ การใช้คุณลักษณะที่ได้จากการประมวลผลก่อนหน้าและคุณลักษณะจากการวิเคราะห์องค์ประกอบสำคัญ และการใช้คุณลักษณะที่ได้จากกระบวนการเลือกคุณลักษณะ ผลการทดลองแสดงให้เห็นว่าวิธีการเลือกคุณลักษณะช่วยให้แบบจำลองการเรียนรู้เชิงลึกมีประสิทธิภาพที่ดีกว่าวิธีพื้นฐาน

การทดลองเพื่อเปรียบเทียบประสิทธิภาพของโครงข่ายประสาทเทียมหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางร่วมกับกลไกจุดสนใจ ผสานกับโครงข่ายประสาทเทียมคอนโวลูชันร่วมกับกลไกจุดสนใจ กับโครงข่ายประสาทเทียมเชิงลึกและการเรียนรู้ของเครื่อง ได้แก่

- 1) โครงข่ายประสาทเทียมคอนโวลูชันร่วมกับกลไกจุดสนใจ
- 2) โครงข่ายประสาทเทียมคอนโวลูชัน
- 3) โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทางที่ใช้กลไกจุดสนใจ
- 4) โครงข่ายหน่วยความจำระยะสั้นแบบยาวแบบสองทิศทาง
- 5) โครงข่ายหน่วยความจำระยะสั้นแบบยาว
- 6) โครงข่ายประสาทเทียมวิธีเฟอร์เซ็ปตรอนหลายชั้น
- 7) ป่าแบบสุ่ม
- 8) ถดถอยโลจิสติก
- 9) เอ็กซ์ทรีมเกรเดียนต์บูตดิ้ง
- 10) ไลท์จีบีเอ็ม

โดยการเปรียบเทียบประสิทธิภาพแบบจำลองที่นำเสนอได้ค่า F1 สูงสุดที่ 97.23% และ 85.48% สำหรับแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ล่วงหน้าชนิดปีและติดตามลำดับ สำหรับแบบจำลองและวิธีการที่นำเสนอนี้สามารถนำไปประยุกต์ใช้ในการสร้างแบบจำลองการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ชนิดอื่นและอุปกรณ์ชนิดอื่น ๆ ได้

6.2 ข้อจำกัดของงานวิจัย

- 1) ชุดข้อมูลจริงของอุปกรณ์ประเภทนี้ไม่มีการเผยแพร่ดังนั้นจึงไม่สามารถหางานวิจัยที่สามารถใช้ชุดข้อมูลเดียวกันในการเปรียบเทียบประสิทธิภาพของแบบจำลองที่ปรับปรุงขึ้นได้
- 2) ข้อมูลของรีโมทคอมเพรสเซอร์ชนิดดีมีจำนวนน้อยและข้อมูลของเซ็นเซอร์ส่วนใหญ่มีความไม่สมบูรณ์
- 3) ข้อจำกัดด้านฮาร์ดแวร์ เนื่องจากแบบจำลองมีความโคงสร้างค่อนข้างซับซ้อนและกระบวนการวิธีประมวลผลข้อมูลก่อนมีหลายขั้นตอนเพื่อลดระยะเวลาในการประมวลผลจึงจำเป็นต้องใช้ระบบการประมวลผลข้อมูลขนาดใหญ่ในการทดลอง

6.3 แนวทางในการวิจัยในอนาคต

เนื่องจากลักษณะข้อมูลของรีโมทคอมเพรสเซอร์มีการปรับค่าควบคุมต่าง ๆ หรือการเปลี่ยนค่าตั้งต้นอุปกรณ์และปัจจัยภายนอกที่ควรนำมาพิจารณา เช่น สภาพอากาศและปริมาณน้ำฝนที่ตกในช่วงเกิดลมพายุหรือมารสุมในทะเลและคุณสมบัติของปิโตรเลียมที่เปลี่ยนแปลงตามธรรมชาติ อีก

ทั้งการทำนายนี้ไม่สามารถระบุได้ว่าเกิดความล้มเหลวของอุปกรณ์ที่จุดใดที่จะสามารถช่วยให้วิศวกรซ่อมบำรุงดำเนินการตรวจสอบได้อย่างรวดเร็วมากขึ้น

การเปรียบเทียบประสิทธิภาพการทำนายควรมีการให้ค่าน้ำหนักความสำคัญกับคลาสบวกมากกว่าโดยพิจารณาถึงค่าความสูญเสียของเครื่องที่จะเกิดกับโอกาสของการเกิดเหตุเสียเป็นต้น

ในการวิจัยควรเพิ่มการตีความหมายจากผลการทำนายเพื่อให้เข้าใจสาเหตุการทำงานของเครื่องที่มีความผิดปกติ ณ จุดใด ๆ หลังจากการวิเคราะห์โดยผู้เชี่ยวชาญทางวิศวกรรมเครื่องยนต์แล้วนั้นเพื่อให้มาซึ่งองค์ความรู้ในการแปลงผลที่ดียิ่งขึ้น



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY





ตารางที่ ก.1 คุณลักษณะที่เลือกนำมาใช้ในกระบวนการวิเคราะห์เบื้องต้น

ลำดับ	ชื่อเซ็นเซอร์	ประเภทของรีโมทคอมเพรสเซอร์/เลขที่													
		ปี										ดี			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Lube Oil Pressure		1	1								1	1	1	1
2	Lube Oil Tank Level	1		1	1	1	1	1	1	1	1	1	1	1	1
3	Lube Oil Temperature	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	Stage 1 Discharge Temperature Difference	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	Stage 2 Discharge Temperature Difference	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	Throw #1 Discharge Pressure	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	Throw #1 Discharge Temperature	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	Throw #1 Suction Pressure	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	Throw #1 Suction Temperature	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	Throw #2 Discharge Pressure	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	Throw #2 Discharge Temperature	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	Throw #2 Suction Pressure	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	Throw #2 Suction Temperature	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	Throw #3 Discharge Pressure	1	1		1	1	1	1	1	1	1	1	1	1	1
15	Throw #3 Suction Pressure			1											
16	Throw #3 Suction Temperature	1	1		1	1	1	1	1	1	1	1	1		1
17	Throw #4 Discharge Temperature											1	1	1	1
18	Throw #4 Suction Pressure	1	1		1	1	1		1	1	1	1	1	1	1

ตารางที่ ก.1 คุณลักษณะที่เลือกนำมาใช้ในกระบวนการวิเคราะห์เบื้องต้น (ต่อ)

ลำดับ	ชื่อเซ็นเซอร์	ประเภทของรีโมทคอมเพรสเซอร์/เลขที่													
		ปี										ดี			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
42	Jacket Water Temperature	1		1	1	1	1		1	1	1	1	1	1	1
43	Left Bank Exhaust Temperature	1	1	1	1	1	1	1	1	1	1				
44	Lube Oil Pressure - Engine	1	1		1	1	1		1	1	1	1	1	1	1
45	Lube Oil Tank Level			1								1			1
46	Lube Oil Temperature - Engine														1
47	Manifold Pressure		1	1								1	1	1	1
48	Right Bank Exhaust Temperature	1	1	1	1	1	1	1	1	1	1				
49	Running Status	1	1	1	1	1	1	1	1	1	1				
50	Speed	1	1		1	1	1	1	1	1	1	1	1	1	1
51	Vibration - Engine	1		1	1	1	1	1	1	1	1	1	1	1	1
	รวมจำนวนเซ็นเซอร์	41	36	37	41	41	41	34	41	41	41	44	41	40	45

1 หมายถึง รีโมทคอมเพรสเซอร์ มีเซ็นเซอร์หรือแอตทริบิวต์ ตัวนี้ในระบบ

ตารางที่ ก.2 คุณลักษณะที่ถูกเลือกนำมาใช้ในการสร้างแบบจำลอง

แอตทริบิวต์ย่อ	ประเภทข้อมูล	หน่วยวัด	ชื่อแอตทริบิวต์
S1	Double	psi	Lube Oil Pressure
S2	Double	°F	Lube Oil Temperature
S3	Double	psi	Throw #1 Discharge Pressure
S4	Double	°F	Throw #1 Discharge Temperature
S5	Double	psi	Throw #1 Suction Pressure

ตารางที่ ก.2 คุณลักษณะที่ถูกเลือกนำมาใช้ในการสร้างแบบจำลอง (ต่อ)

แอดทริบิวต์ย่อ	ประเภทข้อมูล	หน่วยวัด	ชื่อแอดทริบิวต์
S6	Double	°F	Throw #1 Suction Temperature
S7	Double	psi	Throw #2 Discharge Pressure
S8	Double	°F	Throw #2 Discharge Temperature
S9	Double	psi	Throw #2 Suction Pressure
S10	Double	°F	Throw #2 Suction Temperature
S11	Double	psi	Throw #3 Discharge Pressure
S12	Double	°F	Throw #3 Suction Temperature
S13	Double	in/s	Vibration
S14	Double	°F	Cylinder #01 Temperature
S15	Double	°F	Cylinder #02 Temperature
S16	Double	°F	Cylinder #03 Temperature
S17	Double	°F	Cylinder #04 Temperature
S18	Double	°F	Cylinder #05 Temperature
S19	Double	°F	Cylinder #06 Temperature
S20	Double	°F	Cylinder #07 Temperature
S21	Double	°F	Cylinder #08 Temperature
S22	Double	°F	Cylinder #09 Temperature
S23	Double	°F	Cylinder #10 Temperature
S24	Double	°F	Cylinder #11 Temperature
S25	Double	°F	Cylinder #12 Temperature
S26	Double	psi	Lube Oil Pressure Engine
S27	Double	rpm	Speed
S28	Double	in/s	Vibration Engine

หน่วยวัดตัวย่อ	ชื่อหน่วยวัด
°F	องศาฟาเรนไฮต์
%	เปอร์เซ็นต์
in/s	นิ้วต่อวินาที
rpm	รอบต่อนาที (Revolutions per minute)
psi	ปอนด์ต่อตารางนิ้ว

ตารางที่ ก.3 คุณลักษณะที่ใช้ในการวิเคราะห์ห้องค์ประกอบสำคัญ

ลำดับ	แอตทริบิวต์ย่อ	ชื่อแอตทริบิว	แอตทริบิวต์ย่อ	
			Normalized	Robust Scaler
1	S1	Lube Oil Pressure	S1_NORM	S1_RBS
2	S2	Lube Oil Temperature	S2_NORM	S2_RBS
3	S3	Throw #1 Discharge Pressure	S3_NORM	S3_RBS
4	S4	Throw #1 Discharge Temperature	S4_NORM	S4_RBS
5	S5	Throw #1 Suction Pressure	S5_NORM	S5_RBS
6	S6	Throw #1 Suction Temperature	S6_NORM	S6_RBS
7	S7	Throw #2 Discharge Pressure	S7_NORM	S7_RBS
8	S8	Throw #2 Discharge Temperature	S8_NORM	S8_RBS
9	S9	Throw #2 Suction Pressure	S9_NORM	S9_RBS
10	S10	Throw #2 Suction Temperature	S10_NORM	S10_RBS
11	S11	Throw #3 Discharge Pressure	S11_NORM	S11_RBS
12	S12	Throw #3 Suction Temperature	S12_NORM	S12_RBS
13	S13	Vibration	S13_NORM	S13_RBS
14	S14	Cylinder #01 Temperature	S14_NORM	S14_RBS
15	S15	Cylinder #02 Temperature	S15_NORM	S15_RBS
16	S16	Cylinder #03 Temperature	S16_NORM	S16_RBS
17	S17	Cylinder #04 Temperature	S17_NORM	S17_RBS
18	S18	Cylinder #05 Temperature	S18_NORM	S18_RBS
19	S19	Cylinder #06 Temperature	S19_NORM	S19_RBS
20	S20	Cylinder #07 Temperature	S20_NORM	S20_RBS
21	S21	Cylinder #08 Temperature	S21_NORM	S21_RBS
22	S22	Cylinder #09 Temperature	S22_NORM	S22_RBS
23	S23	Cylinder #10 Temperature	S23_NORM	S23_RBS
24	S24	Cylinder #11 Temperature	S24_NORM	S24_RBS
25	S25	Cylinder #12 Temperature	S25_NORM	S25_RBS

ตารางที่ ก.3 คุณลักษณะที่ใช้ในการวิเคราะห์ห้องค์ประกอบสำคัญ(ต่อ)

ลำดับ	แอตทริบิวต์ย่อ	ชื่อแอตทริบิว	แอตทริบิวต์ย่อ	
			Normalized	Robust Scaler
26	S26	Lube Oil Pressure Engine	S26_NORM	S26_RBS
27	S27	Speed	S27_NORM	S27_RBS
28	S28	Vibration Engine	S28_NORM	S28_RBS
29	E1	Euclidean Distance of Low Low Lube Oil Pressure	E1_NORM	E1_RBS
30	E2	Eucli. Dist. of High High Lube Oil Temperatire	E2_NORM	E2_RBS
31	E3	Eucli. Dist. of High High Low Low Compresor Speed	E3_NORM	E3_RBS
32	E4	Eucli. Dist. of High High Low Low Vibration Engine	E4_NORM	E4_RBS
33	E5	Eucli. Dist. of Throw 1st Discharge Temperature	E5_NORM	E5_RBS
34	E6	Eucli. Dist. of Throw #1 Suction temp.	E6_NORM	E6_RBS
35	E7	Eucli. Dist. of Lube Oil Temp.	E7_NORM	E7_RBS
36	E8	Eucli. Dist. of Lube Oil Temp.	E8_NORM	E8_RBS
37	E9	Eucli. Dist. of Throw Pressure	E9_NORM	E9_RBS
38	E10	Eucli. Dist. of Lube Oil Temp.	E10_NORM	E10_RBS
39	E11	Eucli. Dist. of Cylinder Temp.	E11_NORM	E11_RBS
40	E12	Eucli. Dist. of Lube Oil Press Engine	E12_NORM	E12_RBS
41	E13	Eucli. Dist. of Engine Speed	E13_NORM	E13_RBS
42	E14	Eucli. Dist. of Vibration Engine	E14_NORM	E14_RBS

รวม

42

42

รวมทั้งหมด

84 คุณลักษณะ

ตารางที่ ก.4 คุณลักษณะที่ใช้ในการฝึกแบบจำลองไถ่จีบีเอ็ม

ลำดับ	แอตทริบิวต์ย่อ	ชื่อแอตทริบิว	ชื่อแอตทริบิว
1	S1	Lube Oil Pressure	S1_RBS
2	S2	Lube Oil Temperature	S2_RBS
3	S3	Throw #1 Discharge Pressure	S3_RBS
4	S4	Throw #1 Discharge Temperature	S4_RBS
5	S5	Throw #1 Suction Pressure	S5_RBS
6	S6	Throw #1 Suction Temperature	S6_RBS
7	S7	Throw #2 Discharge Pressure	S7_RBS
8	S8	Throw #2 Discharge Temperature	S8_RBS
9	S9	Throw #2 Suction Pressure	S9_RBS
10	S10	Throw #2 Suction Temperature	S10_RBS
11	S11	Throw #3 Discharge Pressure	S11_RBS
12	S12	Throw #3 Suction Temperature	S12_RBS
13	S13	Vibration	S13_RBS
14	S14	Cylinder #01 Temperature	S14_RBS
15	S15	Cylinder #02 Temperature	S15_RBS
16	S16	Cylinder #03 Temperature	S16_RBS
17	S17	Cylinder #04 Temperature	S17_RBS
18	S18	Cylinder #05 Temperature	S18_RBS
19	S19	Cylinder #06 Temperature	S19_RBS
20	S20	Cylinder #07 Temperature	S20_RBS
21	S21	Cylinder #08 Temperature	S21_RBS
22	S22	Cylinder #09 Temperature	S22_RBS
23	S23	Cylinder #10 Temperature	S23_RBS
24	S24	Cylinder #11 Temperature	S24_RBS
25	S25	Cylinder #12 Temperature	S25_RBS

ตารางที่ ก.4 คุณลักษณะที่ใช้ในการฝึกแบบจำลองไคท์จีบีเอ็ม(ต่อ)

ลำดับ	แอตทริบิวต์ย่อ	ชื่อแอตทริบิว	ชื่อแอตทริบิว
26	S26	Lube Oil Pressure Engine	S26_RBS
27	S27	Speed	S27_RBS
28	S28	Vibration Engine	S28_RBS
29	E1	Euclidean Distance of Low Low Lube Oil Pressure	E1_RBS
30	E2	Eucli. Dist. of High High Lube Oil Temperature	E2_RBS
31	E3	Eucli. Dist. of High High Low Low Compressor Speed	E3_RBS
32	E4	Eucli. Dist. of High High Low Low Vibration Engine	E4_RBS
33	E5	Eucli. Dist. of Throw 1st Discharge Temperature	E5_RBS
34	E6	Eucli. Dist. of Throw #1 Suction temp.	E6_RBS
35	E7	Eucli. Dist. of Lube Oil Temp.	E7_RBS
36	E8	Eucli. Dist. of Lube Oil Temp.	E8_RBS
37	E9	Eucli. Dist. of Throw Pressure	E9_RBS
38	E10	Eucli. Dist. of Lube Oil Temp.	E10_RBS
39	E11	Eucli. Dist. of Cylinder Temp.	E11_RBS
40	E12	Eucli. Dist. of Lube Oil Press Engine	E12_RBS
41	E13	Eucli. Dist. of Engine Speed	E13_RBS
42	E14	Eucli. Dist. of Vibration Engine	E14_RBS
43	PCA1	PCA Component 1st	PCA1_RBS
44	PCA2	PCA Component 2nd	PCA2_RBS
45	PCA3	PCA Component 3rd	PCA3_RBS
46	PCA4	PCA Component 4th	PCA4_RBS
47	PCA5	PCA Component 5th	PCA5_RBS
48	PCA6	PCA Component 6th	PCA6_RBS
49	PCA7	PCA Component 7th	PCA7_RBS
50	PCA8	PCA Component 8th	PCA8_RBS



ตารางที่ ข.1 ค่าไอเกนและผลรวมสะสมค่าไอเกนของชุดข้อมูลฝึกและชุดข้อมูลประเมินผลของข้อมูล
รีโมทคอมเพรสเซอร์ชนิดบี

อันดับขององค์ประกอบสำคัญ	ค่าของไอเกน	ผลรวมสะสมค่าของไอเกน
1	0.872560179	0.87256
2	0.036136937	0.90870
3	0.03007669	0.93877
4	0.016344218	0.95512
5	0.007634108	0.96275
6	0.006135656	0.96889
7	0.004682658	0.97357
8	0.002979257	0.97655
9	0.002258365	0.97881
10	0.002205131	0.98101
11	0.001896036	0.98291
12	0.001578733	0.98449
13	0.001455905	0.98594
14	0.001199664	0.98714
15	0.001163336	0.98831

ตารางที่ ข.2 ค่าไอเกนและผลรวมสะสมค่าไอเกนของชุดข้อมูลฝึกและชุดข้อมูลประเมินผลของข้อมูล
รีโมทคอมเพรสเซอร์ชนิดดี

อันดับขององค์ประกอบสำคัญ	ค่าของไอเกน	ผลรวมสะสมค่าของไอเกน
1	0.732815	0.732815
2	0.154407	0.887222
3	0.027918	0.915140
4	0.016275	0.931415
5	0.011095	0.942510
6	0.008986	0.951497
7	0.008290	0.959786
8	0.005257	0.965044

ตารางที่ ข.2 ค่าไอเกนและผลรวมสะสมค่าไอเกนของชุดข้อมูลฝึกและชุดข้อมูลประเมินผลของข้อมูล
รีโมทคอมเพรสเซอร์ชนิดดี (ต่อ)

อันดับขององค์ประกอบสำคัญ	ค่าของไอเกน	ผลรวมสะสมค่าของไอเกน
9	0.004991	0.970035
10	0.003844	0.973878
11	0.003461	0.977339
12	0.002576	0.979915
13	0.002427	0.982343
14	0.002053	0.984396
15	0.001819	0.986215

ตารางที่ ข.3 คะแนนความสำคัญเฉลี่ย ค่าเปอร์เซ็นต์คะแนนความสำคัญเฉลี่ย และผลรวมเปอร์เซ็นต์
คะแนนความสำคัญเฉลี่ยของคุณลักษณะ จากการเรียนรู้ของแบบจำลองต้นไม้ตัดสินใจ
ไลท์แกรนเดียร์นบูตติงบนชุดข้อมูลฝึกของรีโมทคอมเพรสเซอร์ชนิดบี

ลำดับ	แอตทริบิวต์ย่อย	คะแนน ความสำคัญเฉลี่ย	เปอร์เซ็นต์คะแนน ความสำคัญเฉลี่ย	ผลรวมสะสมเปอร์เซ็นต์ คะแนนความสำคัญเฉลี่ย
1	S9	85.37	11.29%	11.29%
2	S13	50.58	6.69%	17.98%
3	PCA_1	34.46	4.56%	22.54%
4	S28	34.07	4.51%	27.05%
5	E4	33.68	4.46%	31.50%
6	E14	27.23	3.60%	35.10%
7	S23	26.03	3.44%	38.55%
8	E6	24.48	3.24%	41.79%
9	S14	24.02	3.18%	44.96%
10	PCA_4	23.51	3.11%	48.07%
11	S20	17.31	2.29%	50.36%
12	S17	16.33	2.16%	52.52%
13	S22	16.09	2.13%	54.65%
14	S25	15.26	2.02%	56.67%

ตารางที่ ข.3 คะแนนความสำคัญเฉลี่ย ค่าเปอร์เซ็นต์คะแนนความสำคัญเฉลี่ย และผลรวมเปอร์เซ็นต์
คะแนนความสำคัญเฉลี่ยของคุณลักษณะ จากการเรียนรู้ของแบบจำลองต้นไม้ตัดสินใจ
ไลท์เกรนเดียร์บนชุดข้อมูลฝึกของริโมทคอมเพรสเซอร์ชนิดบี (ต่อ)

ลำดับ	แอตทริบิวต์ย่อย	คะแนน ความสำคัญเฉลี่ย	เปอร์เซ็นต์คะแนน ความสำคัญเฉลี่ย	ผลรวมสะสมเปอร์เซ็นต์ คะแนนความสำคัญเฉลี่ย
15	S24	14.37	1.90%	58.57%
16	S6	14.07	1.86%	60.43%
17	E12	13.91	1.84%	62.27%
18	S21	13.37	1.77%	64.04%
19	S26	13.22	1.75%	65.79%
20	E7	12.78	1.69%	67.48%
21	S19	12.22	1.62%	69.10%
22	E9	11.92	1.58%	70.67%
23	E8	11.78	1.56%	72.23%
24	PCA_7	11.48	1.52%	73.75%
25	S16	11.31	1.50%	75.24%
26	E10	10.95	1.45%	76.69%
27	S12	10.84	1.43%	78.13%
28	S3	10.62	1.40%	79.53%
29	E11	10.34	1.37%	80.90%
30	S15	10.29	1.36%	82.26%
31	PCA_8	10.23	1.35%	83.61%
32	S5	9.01	1.19%	84.81%
33	PCA_6	8.72	1.15%	85.96%
34	S10	8.6	1.14%	87.10%
35	E5	8.6	1.14%	88.23%
36	S4	8.47	1.12%	89.35%
37	S8	7.74	1.02%	90.38%
38	PCA_2	7.56	1.00%	91.38%
39	S11	6.51	0.86%	92.24%

ตารางที่ ข.3 คะแนนความสำคัญเฉลี่ย ค่าเปอร์เซ็นต์คะแนนความสำคัญเฉลี่ย และผลรวมเปอร์เซ็นต์
คะแนนความสำคัญเฉลี่ยของคุณลักษณะ จากการเรียนรู้ของแบบจำลองต้นไม้ตัดสินใจ
ไลท์เกรนเดียร์นชุดตึงบนชุดข้อมูลฝึกของริโมทคอมเพรสเซอร์ชนิดปี(ต่อ)

ลำดับ	แอตทริบิวต์ย่อย	คะแนน ความสำคัญเฉลี่ย	เปอร์เซ็นต์คะแนน ความสำคัญเฉลี่ย	ผลรวมสะสมเปอร์เซ็นต์ คะแนนความสำคัญเฉลี่ย
40	E13	6.36	0.84%	93.08%
41	S7	6.14	0.81%	93.89%
42	S1	6.13	0.81%	94.70%
43	PCA_3	6.04	0.80%	95.50%
44	E1	5.73	0.76%	96.26%
45	E3	5.59	0.74%	97.00%
46	S27	5.51	0.73%	97.73%
47	S18	5.4	0.71%	98.44%
48	PCA_5	5.36	0.71%	99.15%
49	E2	3.94	0.52%	99.67%
50	S2	2.47	0.33%	100.00%



ภาคผนวก ค

ตารางแสดงรายละเอียดของคุณลักษณะที่ใช้ในการเลือกคุณลักษณะแต่ละการทดลอง



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ตารางที่ ค.1 การใช้คุณลักษณะที่ได้จากการประมวลผลก่อนหน้า 56 คุณลักษณะ

ลำดับ	แอตทริบิวต์ย่อ	ชื่อแอตทริบิว	แอตทริบิวต์ย่อ	
			Normalized	Robust Scaler
1	S1	Lube Oil Pressure	S1_NORM	S1_RBS
2	S2	Lube Oil Temperature	S2_NORM	S2_RBS
3	S3	Throw #1 Discharge Pressure	S3_NORM	S3_RBS
4	S4	Throw #1 Discharge Temperature	S4_NORM	S4_RBS
5	S5	Throw #1 Suction Pressure	S5_NORM	S5_RBS
6	S6	Throw #1 Suction Temperature	S6_NORM	S6_RBS
7	S7	Throw #2 Discharge Pressure	S7_NORM	S7_RBS
8	S8	Throw #2 Discharge Temperature	S8_NORM	S8_RBS
9	S9	Throw #2 Suction Pressure	S9_NORM	S9_RBS
10	S10	Throw #2 Suction Temperature	S10_NORM	S10_RBS
11	S11	Throw #3 Discharge Pressure	S11_NORM	S11_RBS
12	S12	Throw #3 Suction Temperature	S12_NORM	S12_RBS
13	S13	Vibration	S13_NORM	S13_RBS
14	S14	Cylinder #01 Temperature	S14_NORM	S14_RBS
15	S15	Cylinder #02 Temperature	S15_NORM	S15_RBS
16	S16	Cylinder #03 Temperature	S16_NORM	S16_RBS
17	S17	Cylinder #04 Temperature	S17_NORM	S17_RBS
18	S18	Cylinder #05 Temperature	S18_NORM	S18_RBS
19	S19	Cylinder #06 Temperature	S19_NORM	S19_RBS
20	S20	Cylinder #07 Temperature	S20_NORM	S20_RBS
21	S21	Cylinder #08 Temperature	S21_NORM	S21_RBS
22	S22	Cylinder #09 Temperature	S22_NORM	S22_RBS
23	S23	Cylinder #10 Temperature	S23_NORM	S23_RBS
24	S24	Cylinder #11 Temperature	S24_NORM	S24_RBS
25	S25	Cylinder #12 Temperature	S25_NORM	S25_RBS

ตารางที่ ค.1 การใช้คุณลักษณะที่ได้จากการประมวลผลก่อนหน้า 56 คุณลักษณะ (ต่อ)

ลำดับ	แอตทริบิวต์ย่อ	ชื่อแอตทริบิวต์	แอตทริบิวต์ย่อ	
			Normalized	Robust Scaler
26	S26	Lube Oil Pressure Engine	S26_NORM	S26_RBS
27	S27	Speed	S27_NORM	S27_RBS
28	S28	Vibration Engine	S28_NORM	S28_RBS
รวม			28	28
รวมทั้งหมด			56 คุณลักษณะ	

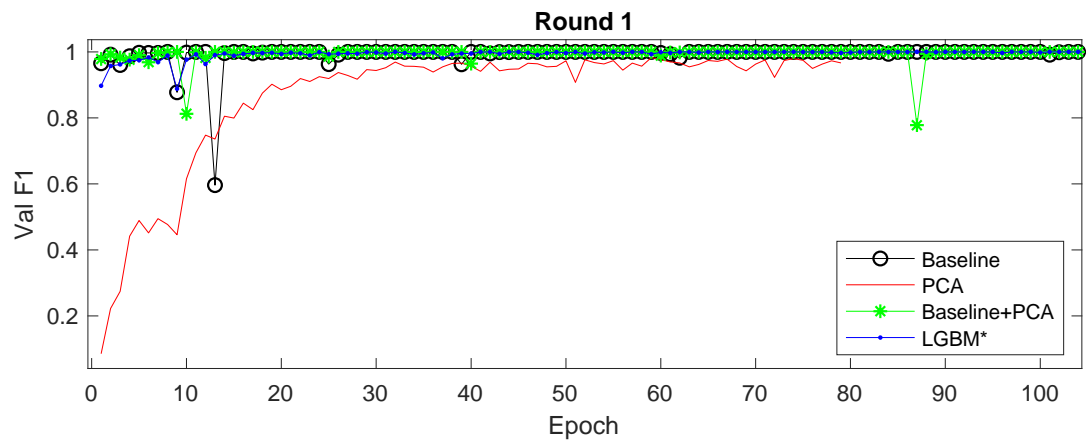
ตารางที่ ค.2 คุณลักษณะที่ได้จากการประมวลผลก่อนหน้าและองค์ประกอบสำคัญที่ถูกเลือก 64 คุณลักษณะ

ลำดับ	แอตทริบิวต์ย่อ	ชื่อแอตทริบิวต์	แอตทริบิวต์ย่อ	
			Normalized	Robust Scaler
1	S1	Lube Oil Pressure	S1_NORM	S1_RBS
2	S2	Lube Oil Temperature	S2_NORM	S2_RBS
3	S3	Throw #1 Discharge Pressure	S3_NORM	S3_RBS
4	S4	Throw #1 Discharge Temperature	S4_NORM	S4_RBS
5	S5	Throw #1 Suction Pressure	S5_NORM	S5_RBS
6	S6	Throw #1 Suction Temperature	S6_NORM	S6_RBS
7	S7	Throw #2 Discharge Pressure	S7_NORM	S7_RBS
8	S8	Throw #2 Discharge Temperature	S8_NORM	S8_RBS
9	S9	Throw #2 Suction Pressure	S9_NORM	S9_RBS
10	S10	Throw #2 Suction Temperature	S10_NORM	S10_RBS
11	S11	Throw #3 Discharge Pressure	S11_NORM	S11_RBS
12	S12	Throw #3 Suction Temperature	S12_NORM	S12_RBS
13	S13	Vibration	S13_NORM	S13_RBS
14	S14	Cylinder #01 Temperature	S14_NORM	S14_RBS
15	S15	Cylinder #02 Temperature	S15_NORM	S15_RBS
16	S16	Cylinder #03 Temperature	S16_NORM	S16_RBS
17	S17	Cylinder #04 Temperature	S17_NORM	S17_RBS
18	S18	Cylinder #05 Temperature	S18_NORM	S18_RBS
19	S19	Cylinder #06 Temperature	S19_NORM	S19_RBS
20	S20	Cylinder #07 Temperature	S20_NORM	S20_RBS

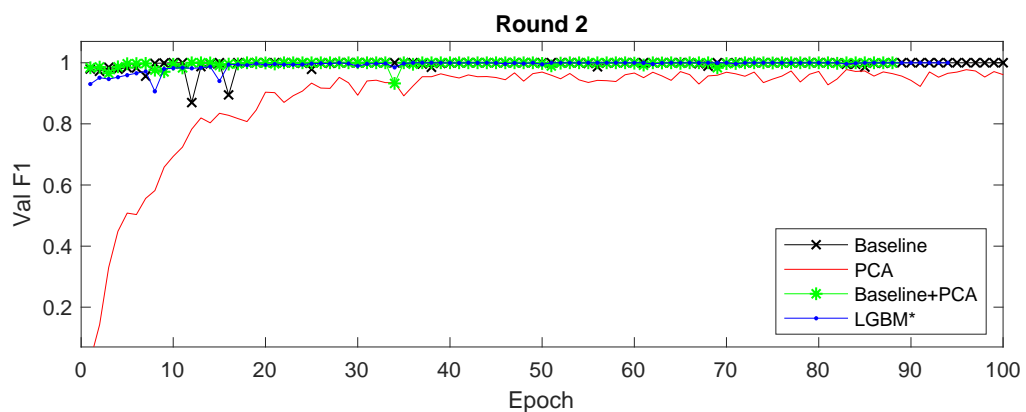
ตารางที่ ค.2 คุณลักษณะที่ได้จากการประมวลผลก่อนหน้าและองค์ประกอบสำคัญที่ถูกเลือก 64
คุณลักษณะ (ต่อ)

ลำดับ	แอตทริบิวต์ย่อ	ชื่อแอตทริบิวต์	แอตทริบิวต์ย่อ	
			Normalized	Robust Scaler
21	S21	Cylinder #08 Temperature	S21_NORM	S21_RBS
22	S22	Cylinder #09 Temperature	S22_NORM	S22_RBS
23	S23	Cylinder #10 Temperature	S23_NORM	S23_RBS
24	S24	Cylinder #11 Temperature	S24_NORM	S24_RBS
25	S25	Cylinder #12 Temperature	S25_NORM	S25_RBS
26	S26	Lube Oil Pressure Engine	S26_NORM	S26_RBS
27	S27	Speed	S27_NORM	S27_RBS
28	S28	Vibration Engine	S28_NORM	S28_RBS
29	PCA1	PCA Component 1 st		PCA1_RBS
30	PCA2	PCA Component 2 nd		PCA2_RBS
31	PCA3	PCA Component 3 rd		PCA3_RBS
32	PCA4	PCA Component 4 th		PCA4_RBS
33	PCA5	PCA Component 5 th		PCA5_RBS
34	PCA6	PCA Component 6 th		PCA6_RBS
35	PCA7	PCA Component 7 th		PCA7_RBS
36	PCA8	PCA Component 8 th		PCA8_RBS

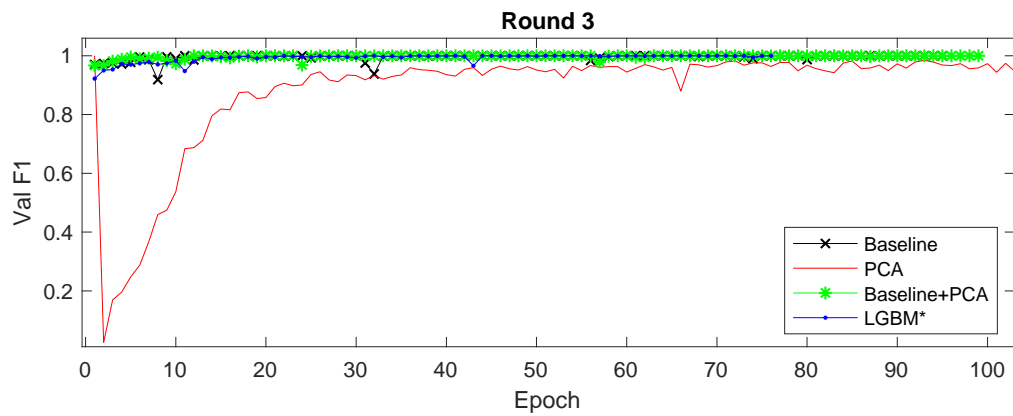
รวม 28 36
รวมทั้งหมด 64 คุณลักษณะ



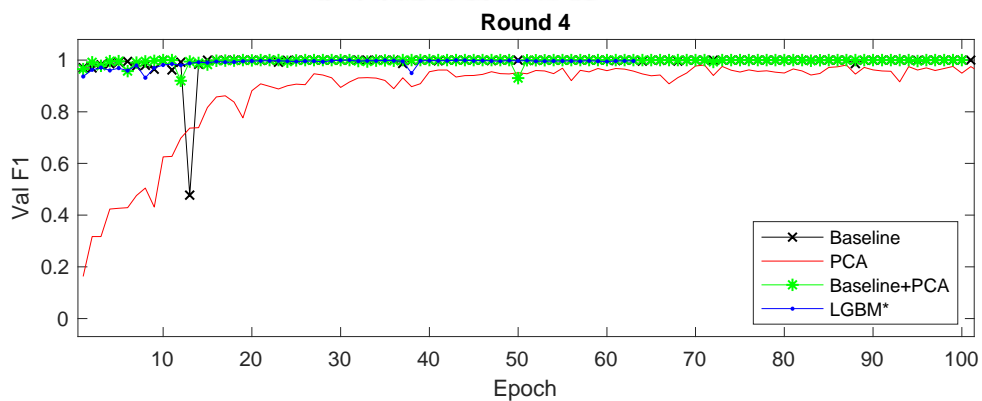
รูปที่ ค.1 กราฟแสดงค่า F1 บทชุดข้อมูลประเมินผลของการทดลองแบบจำลองที่นำเสนอ ABDLSTM&ACNN* กับการใช้คุณลักษณะจากการเลือกวิธี Baseline PCA Baseline+PCA และ LGBM ใน 100 รอบการเรียนรู้สำหรับการทดลองครั้งที่ 1



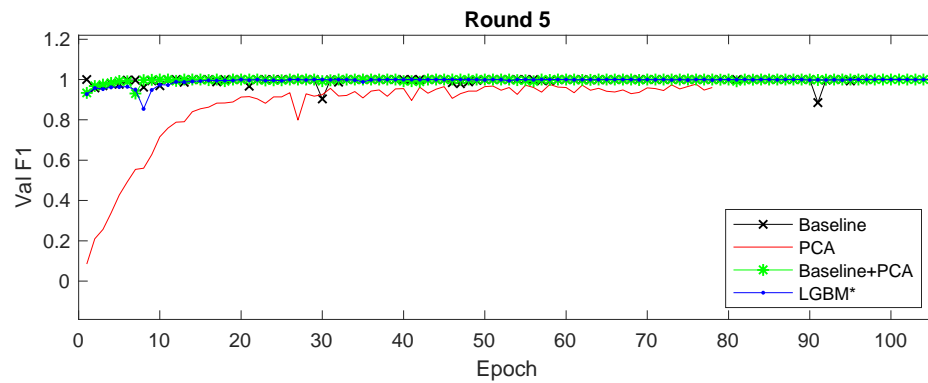
รูปที่ ค.2 กราฟแสดงค่า F1 บทชุดข้อมูลประเมินผลของการทดลองแบบจำลองที่นำเสนอ ABDLSTM&ACNN* กับการใช้คุณลักษณะจากการเลือกวิธี Baseline PCA Baseline+PCA และ LGBM ใน 100 รอบการเรียนรู้สำหรับการทดลองครั้งที่ 2



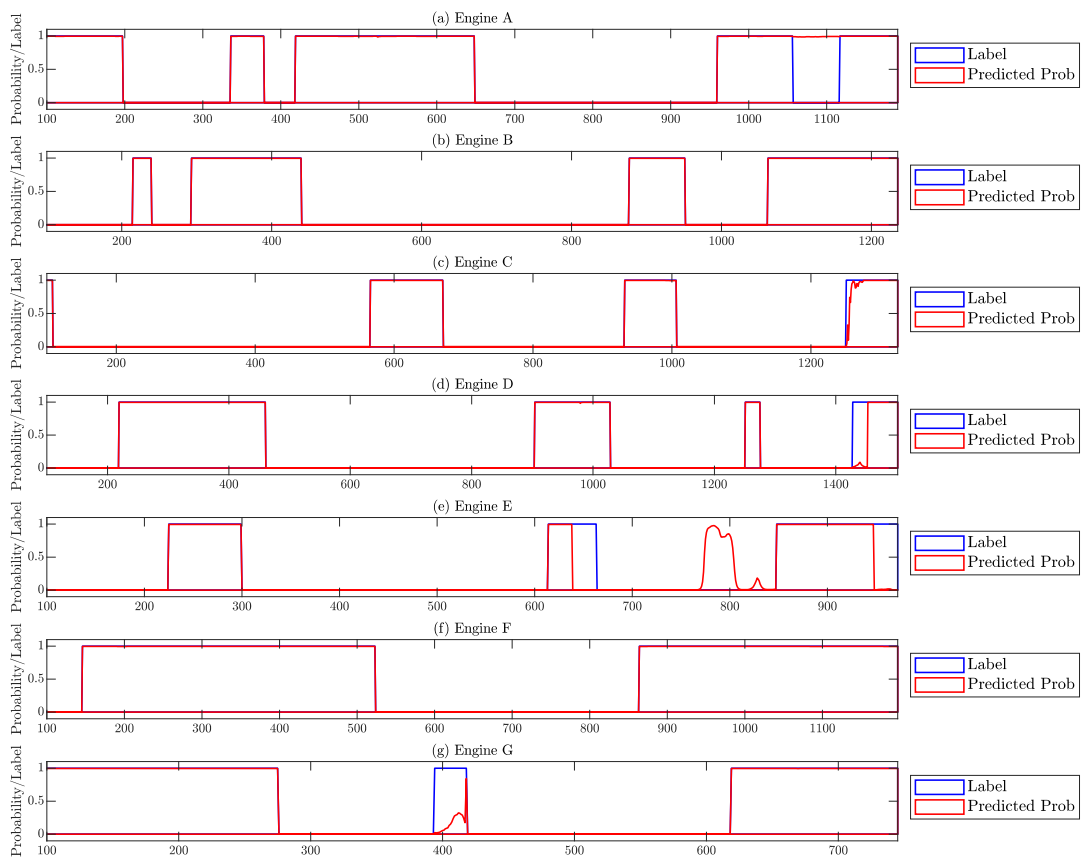
รูปที่ ค.3 กราฟแสดงค่า F1 บทชุดข้อมูลประเมินผลของการทดลองแบบจำลองที่นำเสนอ ABDLSTM&ACNN* กับการใช้คุณลักษณะจากการเลือกวิธี Baseline PCA Baseline+PCA และ LGBM ใน 100 รอบการเรียนรู้สำหรับการทดลองครั้งที่ 3



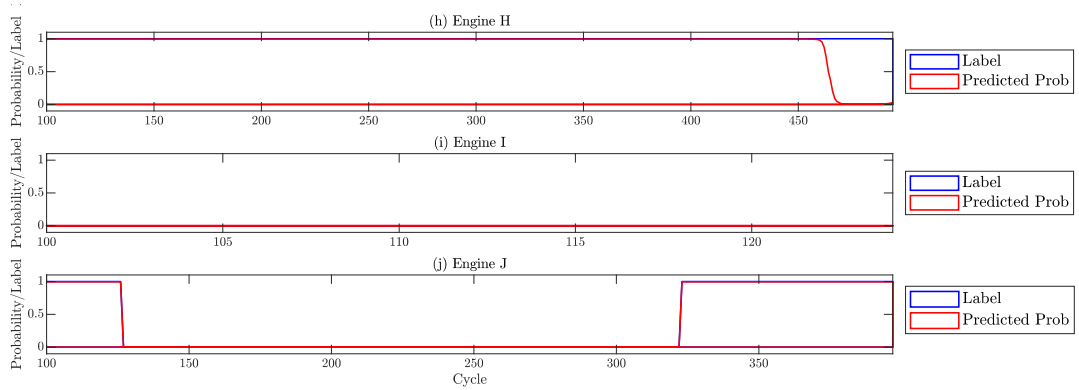
รูปที่ ค.4 กราฟแสดงค่า F1 บทชุดข้อมูลประเมินผลของการทดลองแบบจำลองที่นำเสนอ ABDLSTM&ACNN* กับการใช้คุณลักษณะจากการเลือกวิธี Baseline PCA Baseline+PCA และ LGBM ใน 100 รอบการเรียนรู้สำหรับการทดลองครั้งที่ 4



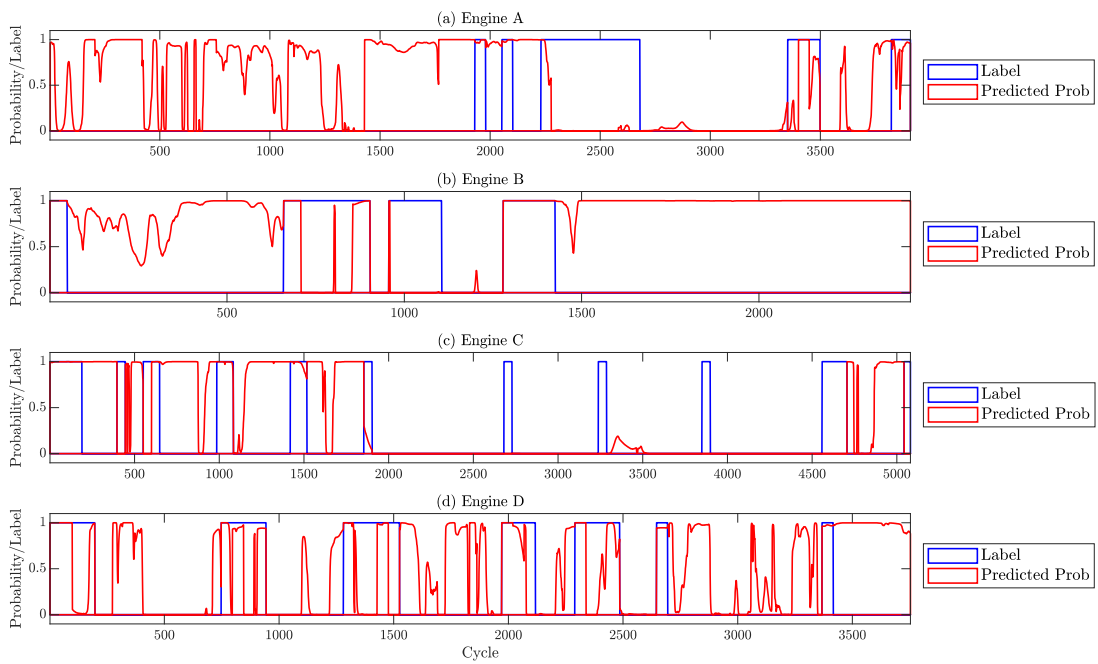
รูปที่ ค.5 กราฟแสดงค่า F1 บทชุดข้อมูลประเมินผลของการทดลองแบบจำลองที่นำเสนอ ABDLSTM&ACNN* กับการใช้คุณลักษณะจากการเลือกวิธี Baseline PCA Baseline+PCA และ LGBM ใน 100 รอบการเรียนรู้สำหรับการทดลองครั้งที่ 5



รูปที่ ค.6 ตัวอย่างการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ชนิดบี เครื่อง A ถึง G



รูปที่ ค.7 ตัวอย่างการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ชนิดบี เครื่อง H ถึง J



รูปที่ ค.8 ตัวอย่างการทำนายความล้มเหลวของรีโมทคอมเพรสเซอร์ชนิดดี เครื่อง A ถึง D

บรรณานุกรม

1. Benefits, K., *C-SERIES 3516 Compression Package Proven Quality and Reliability, Delivered Fast*. 2018.
2. Cho, S., et al., *A Study on Estimating the Next Failure Time of Compressor Equipment in an Offshore Plant*. *Mathematical Problems in Engineering*, 2016. **2016**: p. 1-14.
3. Canizo, M., et al., *Real-time predictive maintenance for wind turbines using Big Data frameworks*, in *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*. 2017, IEEE. p. 70-77.
4. Li, Z., *Deep Learning Driven Approaches for Predictive Maintenance: A Framework of Intelligent Fault Diagnosis and Prognosis in the Industry 4.0 Era*. 2018.
5. Su, C.-J. and S.-F. Huang, *Real-time big data analytics for hard disk drive predictive maintenance*. *Computers & Electrical Engineering*, 2018. **71**: p. 93-101.
6. Graves, A., A.-r. Mohamed, and G. Hinton, *Speech Recognition with Deep Recurrent Neural Networks*. 2013.
7. มนัสกันต์ เสนหา และ พีรพล เวทีกุล, “การทำนายอายุการใช้งานคงเหลือของเครื่องจักรด้วยโครงข่ายคอนโวลูชันเชิงลึกที่เพิ่มประสิทธิภาพ”, *วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย*, 2559.
8. Karim, F., et al. *Multivariate LSTM-FCNs for Time Series Classification*. 2019.
9. ทวีศักดิ์ เอี่ยมสวัสดิ์ และ บุญเสริม กิจศิริกุล, “การรู้จำตัวอักษรพิมพ์ภาษาไทยโดยใช้หน่วยความจำระยะสั้นแบบยาว”, *วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ สาขาวิทยาศาสตร์คอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย*, 2559.
10. O'Shea, K. and R. Nash, *An Introduction to Convolutional Neural Networks*. 2015.
11. Léon Bottou(NEC laboratories of America, P., NJ), Olivier Bousquet (Google Zürich, Zurich, Switzerland), *The Tradeoffs of Large Scale Learning*, in *NIPS'07 Proceedings of the 20th International Conference on Neural Information Processing Systems*. 2007, Curran Associates Inc. , USA ©2007. p. 161-168.

12. Jolliffe, I.T. and J. Cadima, *Principal component analysis: a review and recent developments*. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 2016. **374**: p. 20150202.
13. Bahdanau, D., K. Cho, and Y. Bengio, *Neural Machine Translation by Jointly Learning to Align and Translate*. 2014.
14. Ke, G., et al. *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*. 2017.
15. Mathew, V., et al., *Prediction of Remaining Useful Lifetime (RUL) of turbofan engine using machine learning*, in *2017 IEEE International Conference on Circuits and Systems (ICCS)*. 2017, IEEE. p. 306-311.
16. Zarei, J., M.A. Tajeddini, and H.R. Karimi, *Vibration analysis for bearing fault detection and classification using an intelligent filter*. Mechatronics, 2014. **24**: p. 151-157.
17. Jahnke, P., *Machine Learning Approaches for Failure Type Detection and Predictive Maintenance*. 2015.
18. ElSaid, A., et al., *Optimizing Long Short-Term Memory Recurrent Neural Networks Using Ant Colony Optimization to Predict Turbine Engine Vibration*. ArXiv, 2017.
19. Liu, Y., X. Hu, and W. Zhang, *Remaining useful life prediction based on health index similarity*. Reliability Engineering & System Safety, 2019. **185**: p. 502-510.
20. Zhao, R., et al., *Learning to Monitor Machine Health with Convolutional Bi-Directional LSTM Networks*. Sensors, 2017. **17**: p. 273.
21. Zhang, C., et al. *A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data*. 2018.
22. Chen, T. and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining – KDD'16*. 2016. p. 785-794.



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ประวัติผู้เขียน

ชื่อ-สกุล	นายวิระศักดิ์ ชมภู
วัน เดือน ปี เกิด	3 มิถุนายน 2529
สถานที่เกิด	อำเภอ สว่างแดนดิน จังหวัด สกลนคร
วุฒิการศึกษา	มัธยมศึกษา โรงเรียนหนองบัวพิทยาคาร จังหวัด หนองบัวลำภู วท.บ. วิทยาการคอมพิวเตอร์ประยุกต์-มัลติมีเดีย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้า ธนบุรี กรุงเทพมหานคร วท.ม. เทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี กรุงเทพมหานคร กำลังศึกษา วท.ม. วิทยาศาสตร์คอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร
ที่อยู่ปัจจุบัน	33 ซอย พหลโยธิน 29 เขต แขวงจตุจักร กรุงเทพมหานคร
ผลงานตีพิมพ์	W. Chomphu, B. Kijirikul, Wellhead Compressor Failure Prediction Using Attention-based Bidirectional LSTMs with Data Reduction Techniques, The 4th International Conference on Compute and Data Analysis, ACM, Silicon Valley, CA, USA, 2020: pp. 16–22. https://doi.org/10.1145/3388142.3388154 .