การเปรียบเทียบประสิทธิภาพเทคนิคการเรียนรู้ของเครื่องสำหรับการบำรุงรักษาเชิงคาดการณ์ของเครื่องยนต์อากาศยาน

วิรากานต์ กิตติบวรกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
บัณฑิตวิทยาลัย สาขาเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีไทย-ญี่ปุ่น
ปีการศึกษา 2565

THE COMPARISON OF MACHINE LEARNING ALGORITHMS FOR PREDICTIVE MAINTENANCE OF AIRCRAFT ENGINE

Wirakarn Kittibawornkul

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Information Technology

Graduate School

Thai-Nichi Institute of Technology

Academic Year 2022

หัวข้อวิทยานิพนธ์	การเปรียบเทียบประสิทธิภาพเทคนิคการเรียนรู้ของเครื่องสำหรับการบำรุงรักษาเชิงคาดการณ์ของเครื่องยนต์อากาศยาน
โดย	วิรากานต์ กิตติบวรกุล
สาขาวิชา	เทคโนโลยีสารสนเทศ
อาจารย์ที่ปรึกษา	ดร.ณปภัช วิชัยดิษฐ์

บัณฑิตวิทยาลัย สถาบันเทคโนโลยีไทย-ญี่ปุ่น อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีบัณฑิตวิทยาลัย

(รองศาสตราจารย์ ดร.พิชิต สุขเจริญพงษ์)

วันที่ ..... เดือน ..... พ.ศ. .....

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(รองศาสตราจารย์ ดร.อรรณพ หมั่นสกุล)

.....กรรมการ

(ดร.ศรายุทธ นนท์ศิริ)

.....กรรมการ

(ดร.ประมุข บุญเสี่ยง)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์

(ดร.ณปภัช วิชัยดิษฐ์)

วิรากานต์ กิตติบวรกุล : การเปรียบเทียบประสิทธิภาพเทคนิคการเรียนรู้ของเครื่องสำหรับการบำรุงรักษาเชิงคาดการณ์ของเครื่องยนต์อากาศยาน. อาจารย์ที่ปรึกษา : ดร.ณปภัช วิชัยดิษฐ์, 65 หน้า.

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อสร้างโมเดลการบำรุงรักษาเชิงคาดการณ์ของเครื่องยนต์อากาศยานด้วยเทคนิคการเรียนรู้ของเครื่องและเปรียบเทียบประสิทธิภาพในการคาดการณ์ความเสียหายของแต่ละเทคนิคเพื่อหาเทคนิคที่มีความเหมาะสม โดยใช้ชุดข้อมูลจากโมเดลใหม่ของ Commercial Modular Aero-Propulsion System Simulation (N-CMAPSS) ในการฝึกฝนและทดสอบโมเดลสำหรับการคาดการณ์ความเสียหาย ชุดข้อมูลดังกล่าวเป็นชุดข้อมูลเกี่ยวกับความเสียหายที่เกิดขึ้นกับเครื่องยนต์อากาศยานแบบ Turbofan จากศูนย์ความเป็นเลิศด้านการทำนาย (The Prognostics Center of Excellence) ของศูนย์วิจัยนาซาเอมส์ (NASA Ames Research Center) เทคนิคการเรียนรู้ของเครื่องทั้งหมด 8 เทคนิคได้แก่ ต้นไม้ตัดสินใจ ป่าแบบสุ่ม Extreme Gradient Boosting ขั้นตอนวิธีเพื่อนบ้านใกล้สุด โครงข่ายประสาทเทียม โครงข่ายประสาทเทียมคอนโวลูชัน โครงข่ายประสาทเทียมวกกลับแบบ Simple และแบบ LSTM ถูกนำมาใช้สำหรับสร้างโมเดลเพื่อคาดการณ์ความเสียหาย ในส่วนของการประเมินโมเดลใช้ค่าสัมประสิทธิ์การกำหนด (R^2) และค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ย (RMSE) จากการฝึกฝนและทดสอบโมเดลทั้งหมด 8 เทคนิคพบว่าเทคนิคป่าแบบสุ่มให้ผลลัพธ์ที่ดีที่สุดด้วยค่าสัมประสิทธิ์การกำหนดเท่ากับ 0.7780 และค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ยเท่ากับ 11.2844

บัณฑิตวิทยาลัย ลายมือชื่อนักศึกษา.....

สาขาวิชา เทคโนโลยีสารสนเทศ ลายมือชื่ออาจารย์ที่ปรึกษา.....

ปีการศึกษา 2565

WIRAKARN KITTIBAWORNKUL : THE COMPARISON OF MACHINE LEARNING ALGORITHMS FOR PREDICTIVE MAINTENANCE OF AIRCRAFT ENGINE. ADVISOR : DR.NAPAPHAT VICHAIDIS, 65 PP.

This paper aimed to create the predictive maintenance model of aircraft engine by implementing the machine learning techniques and to compare the prediction performance of each technique to determine the most suitable technique. Datasets for this study were collected from new model of Commercial Modular Aero-Propulsion System Simulation (N-CMAPSS) used for training and testing for the predictive maintenance model. N-CMAPSS was a series of data on the damage to turbofan engines from The Prognostics Center of Excellence (PCoE), NASA Ames Research Center. There were 8 machine learning techniques including (1) Decision Tree (2) Random Forest (3) Extreme Gradient Boosting (4) k-Nearest Neighbors (5) Artificial Neural Network (6) Convolutional Neural Network (7) Simple Recurrent Network and (8) LSTM Recurrent Network. They were used to implement the predictive maintenance model. In addition, the designation Coefficient of Determination or R-square (R^2) and Root Mean Square Error (RMSE) were used to evaluate the model. The analysis of the training and testing 8 techniques concluded that Random Forest was the suitable technique with 0.7780 of R^2 and 11.2844 of RMSE.

Graduate School Student's Signature.....

Field of Study Information Technology Advisor's Signature.....

Academic Year 2022

กิตติกรรมประกาศ

ในการจัดทำวิทยานิพนธ์ฉบับนี้ ทางผู้วิจัยได้พบกับอุปสรรคมากมายตลอดการดำเนินงานวิจัย การเรียนรู้ในสิ่งต่าง ๆ ที่เกี่ยวข้องเกิดขึ้นได้ด้วยความกรุณาจาก ดร.ณปภัช วิชัยดิษฐ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้ให้ความรู้และคำแนะนำตลอดจนกำลังใจให้ผู้วิจัยสามารถทำงานวิจัยนี้ได้สำเร็จลุล่วงไปได้ด้วยดี รวมถึงขอขอบพระคุณทางสถาบันเทคโนโลยีไทย-ญี่ปุ่น ที่ให้การสนับสนุนค่าเล่าเรียนแก่ผู้วิจัย ทางผู้วิจัยรู้สึกซาบซึ้งและขอกล่าวขอบพระคุณมา ณ ที่นี้

วิทยานิพนธ์ฉบับนี้จะมีความสมบูรณ์ไม่ได้หากขาดคำชี้แนะแนวทางอันเป็นประโยชน์ยิ่งจากคณะกรรมการสอบวิทยานิพนธ์อันประกอบไปด้วย รองศาสตราจารย์ ดร.อรรณพ หมั่นสกุล ดร.ศรายุทธ นนท์ศิริ และดร.ประมุข บุญเสี่ยง ที่ให้เกียรติเป็นคณะกรรมการคอยชี้แนะสิ่งที่ก่อให้เกิดประโยชน์ในการดำเนินงานวิจัยและจัดทำวิทยานิพนธ์ฉบับนี้ให้มีความสมบูรณ์ยิ่ง

ขอขอบพระคุณ คุณพรรษชล แสงอรุณ บุคลากรคณะเทคโนโลยีสารสนเทศ ผู้ให้คำแนะนำในการจัดทำเอกสารที่เกี่ยวข้องและความช่วยเหลือในการดำเนินการขั้นตอนต่าง ๆ ของทางบัณฑิตวิทยาลัย ตลอดจนช่วยอำนวยความสะดวกด้านอาคารสถานที่ให้แก่ผู้วิจัย

สุดท้ายนี้ขอขอบพระคุณครอบครัวและบุคคลอันเป็นที่รักยิ่งของทางผู้วิจัย ที่ให้การสนับสนุนทั้งแรงกายแรงใจ เป็นแรงผลักดันสำคัญให้ผู้วิจัยสามารถดำเนินงานวิจัยและจัดทำวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี

ทางผู้วิจัยหวังเป็นอย่างยิ่งว่าวิทยานิพนธ์ฉบับนี้จะเป็นประโยชน์ให้กับผู้ที่มีความสนใจในด้านปัญญาประดิษฐ์และการเรียนรู้ของเครื่อง รวมถึงการนำไปประยุกต์ใช้ให้เกิดประโยชน์ในด้านอุตสาหกรรมและงานที่เกี่ยวข้อง เพื่อพัฒนาองค์ความรู้ด้านเทคโนโลยีและอุตสาหกรรมให้มีความก้าวหน้ามากยิ่งขึ้น

วิรากานต์ กิตติบวรกุล

สารบัญ

สารบัญ (ต่อ)

สารบัญ (ต่อ)

บทที่ หน้า

สารบัญตาราง

สารบัญรูป

สารบัญรูป (ต่อ)

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญ

เครื่องยนต์เป็นระบบกลไกที่มีความสำคัญต่อการขับเคลื่อนอากาศยาน ด้วยโครงสร้างที่มีความซับซ้อนและจำเป็นที่จะต้องทำงานอยู่ในสภาวะที่ไม่เอื้ออำนวย ส่งผลให้มีโอกาสที่ประสิทธิภาพของเครื่องยนต์จะลดลงซึ่งหมายถึงค่าใช้จ่ายที่เพิ่มสูงขึ้นในหลายด้าน อัตราการสิ้นเปลืองเชื้อเพลิงที่มากขึ้น รวมถึงค่าบำรุงรักษาที่อาจตามมาในกรณีที่เกิดการเสื่อมสภาพและอาจนำไปสู่ความเสียหายที่เป็นอันตรายอย่างยิ่ง ดังนั้นความสมบูรณ์ของชิ้นส่วนเครื่องยนต์และการทำงานที่มีความเชื่อถือได้สูงจึงเป็นสิ่งที่สำคัญยิ่งของความปลอดภัยขณะทำการบิน [1, 2] ด้วยเหตุนี้ความสำคัญของการบำรุงรักษาจึงเพิ่มขึ้นและส่งผลให้แนวโน้มความสนใจในการพัฒนา รวมถึงการปรับใช้กลยุทธ์การบำรุงรักษาที่มีประสิทธิภาพเพิ่มสูงขึ้นเช่นกัน การบำรุงรักษาที่มีประสิทธิภาพจะสามารถปรับปรุงความน่าเชื่อถือของระบบ ป้องกันความล้มเหลวของระบบ และลดต้นทุนการบำรุงรักษาได้ [3] อีกทั้งในช่วงที่ผ่านมาการบำรุงรักษาเชิงคาดการณ์ (Predictive Maintenance) ได้รับความนิยมสูงขึ้น การปฏิวัติอุตสาหกรรม 4.0 ก็ให้การสนับสนุนการบำรุงรักษาประเภทนี้จากความต้องการในภาคปฏิบัติที่เพิ่มสูงขึ้นเช่นกัน เนื่องจากมีความได้เปรียบในเรื่องของความน่าเชื่อถือ ความพร้อมในการใช้งาน และความปลอดภัยที่มากกว่าการบำรุงรักษาแบบเดิม [4] ซึ่งหมายถึงการบำรุงรักษาเชิงรับ (Corrective Maintenance) ที่จะทำการซ่อมบำรุงเมื่อเกิดความเสียหายขึ้น เป็นวิธีการที่ใช้มาอย่างยาวนานและเป็นที่นิยมมากจากความง่ายในการดำเนินการแต่ก็แฝงไปด้วยค่าใช้จ่ายที่มากมาย ไม่ว่าจะเป็นค่าใช้จ่ายที่ต้องเสียไปในขณะที่กระบวนการผลิตหยุดชะงักหรือค่าดำเนินงานต่าง ๆ รวมถึงการซ่อมบำรุงเชิงป้องกัน (Preventive Maintenance) ที่เป็นการซ่อมบำรุงตามแผนที่วางไว้ แนวทางที่จะสามารถป้องกันความเสียหายที่อาจเกิดขึ้นได้แต่ก็จะเป็นการใช้ทรัพยากรสิ้นเปลืองโดยใช่เหตุและเสียค่าใช้จ่ายที่ไม่จำเป็นไปอย่างมหาศาลเช่นกัน [5] จากการศึกษาพบว่ามีสามแนวทางหลักที่ใช้สำหรับคาดการณ์ความเสียหายที่อาจเกิดขึ้น ได้แก่ แนวทางโมเดลทางกายภาพ (Physical Model-Based) ซึ่งจะใช้โมเดลทางคณิตศาสตร์และวิธีการทางสถิติเป็นหลัก แนวทางความรู้ (Knowledge-Based) เป็นแนวทางที่จะช่วยลดความซับซ้อนของโมเดลทางกายภาพจึงนิยมใช้ควบคู่กันแบบผสมผสาน (Hybrid) และแนวทางการใช้ข้อมูล (Data-Driven) มีการนำโมเดลที่หลากหลายมาปรับใช้ในการพัฒนาการบำรุงรักษาเชิงคาดการณ์ โดยพื้นฐานทางสถิติ (Statistics-Based) ปัญญาประดิษฐ์

(Artificial Intelligence: AI) และโมเดลที่อ้างอิงจากเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ถือเป็นที่นิยมอย่างมากในปัจจุบัน [6] การเรียนรู้ด้วยเครื่องถือเป็นเครื่องมือที่มีสมรรถภาพสูงในการพัฒนาอัลกอริทึมเพื่อประยุกต์ใช้สำหรับการคาดการณ์มากมาย โดยประสิทธิภาพของการนำไปประยุกต์ใช้นั้นขึ้นอยู่กับการเลือกเทคนิคการเรียนรู้ของเครื่องให้มีความเหมาะสม [7] จากเหตุผลข้างต้น การเปรียบเทียบเทคนิคการเรียนรู้ของเครื่องจะสามารถชี้ให้เห็นว่าชุดข้อมูล (Dataset) ที่มีอยู่นั้นเหมาะสมที่จะใช้เทคนิคไหนนำมาคาดการณ์ความเสียหายได้ เนื่องด้วยข้อจำกัดในทางปฏิบัติแล้ว การที่จะรวบรวมข้อมูลความเสียหายประเภทที่มีการใช้งานจนกระทั่งเกิดความเสียหาย (Run-To-Failure) ของเครื่องยนต์อากาศยานนั้นเป็นไปได้ยากเนื่องจากต้องใช้ระยะเวลานาน มีค่าใช้จ่ายสูงมาก [8] และเหตุผลสำคัญด้านความปลอดภัยที่ส่งผลกระทบอย่างร้ายแรงหากเกิดเหตุการณ์เช่นนั้นขึ้น จึงจะมีการซ่อมบำรุงก่อนเสมอ ดังนั้นงานวิจัยนี้จะนำชุดข้อมูลที่สร้างขึ้นจากโมเดลใหม่ของ Commercial Modular Aero-Propulsion System Simulation (N-CMAPSS) [9] มาใช้สำหรับฝึกฝนโมเดลที่มีเทคนิคการเรียนรู้ของเครื่องที่แตกต่างกัน รวมถึงทำการเปรียบเทียบผลลัพธ์ที่ได้จากการทดสอบในแต่ละโมเดล

1.2 วัตถุประสงค์

1. เพื่อสร้างโมเดลการบำรุงรักษาเชิงคาดการณ์ของเครื่องยนต์อากาศยานด้วยเทคนิคการเรียนรู้ของเครื่อง

2. เพื่อเปรียบเทียบประสิทธิภาพด้านความแม่นยำในการคาดการณ์ของแต่ละเทคนิคและกาเทคนิคที่มีความเหมาะสมที่สุด

1.3 ขอบเขตงานวิจัย

1. ฝึกฝนและทดสอบโมเดลที่ใช้ในการคาดการณ์ความเสียหายของเครื่องยนต์อากาศยานด้วยชุดข้อมูลที่ได้จากโมเดลใหม่ของ Commercial Modular Aero-Propulsion System Simulation (N-CMAPSS)

2. เปรียบเทียบประสิทธิภาพด้านความแม่นยำของโมเดลการบำรุงรักษาเชิงคาดการณ์ของเครื่องยนต์อากาศยาน โดยใช้เทคนิคดังต่อไปนี้

(1) ต้นไม้ตัดสินใจ

(2) ป่าแบบสุ่ม

(3) Extreme Gradient Boosting

(4) ขั้นตอนวิธีเพื่อนบ้านใกล้สุด

(5) โครงข่ายประสาทเทียม

(6) โครงข่ายประสาทเทียมคอนโวลูชัน

(7) โครงข่ายประสาทเทียมวกกลับแบบ Simple

(8) โครงข่ายประสาทเทียมวกกลับแบบ LSTM

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. โมเดลสำหรับการคาดการณ์ความเสียหายของเครื่องยนต์อากาศยานในแต่ละเทคนิค
2. การตัดสินใจเลือกใช้เทคนิคการเรียนรู้ของเครื่องให้มีความเหมาะสมกับชุดข้อมูล
3. นำความรู้ด้านการเรียนรู้ของเครื่องมาประยุกต์ใช้ในงานการบำรุงรักษาเชิงคาดการณ์อื่น
4. นำความรู้ที่ได้ไปพัฒนาการบำรุงรักษาเชิงคาดการณ์ให้มีประสิทธิภาพมากยิ่งขึ้น

บทที่ 2
การทบทวนวรรณกรรม

2.1 อุตสาหกรรม 4.0

หลังจากการปฏิวัติอุตสาหกรรมสามครั้งที่ผ่านมาในอดีต แนวคิดของการปฏิวัติอุตสาหกรรมครั้งที่สี่ในชื่อ อุตสาหกรรม 4.0 (Industry 4.0) ได้ถูกกล่าวถึงเป็นครั้งแรกในเดือนพฤศจิกายนปี พ.ศ. 2554 ผ่านบทความที่ตีพิมพ์โดยรัฐบาลเยอรมันและได้ถูกนำเสนออย่างเป็นทางการในปี พ.ศ. 2556 [10] อุตสาหกรรม 4.0 ได้แสดงให้เห็นถึงการปฏิวัตอุตสาหกรรมที่มีมาอย่างยาวนานตั้งแต่อดีตจนกระทั่งปัจจุบันดังรูปที่ 2.1 ซึ่งหมายถึงการเปลี่ยนแปลงเทคโนโลยีที่นำมาใช้ในอุตสาหกรรมอย่างรวดเร็ว ก่อให้เกิดการเปลี่ยนแปลงรูปแบบในการดำเนินงานภาคอุตสาหกรรมดังต่อไปนี้ [11]

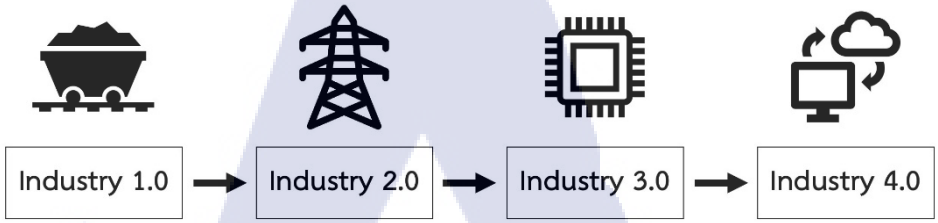
1. การปฏิวัติอุตสาหกรรมครั้งที่ 1 เกิดขึ้นเมื่อช่วงปลายศตวรรษที่ 18 ที่น้ำและไอน้ำถูกใช้เป็นพื้นฐานของการผลิตเครื่องจักรกลในยุคสมัยนั้น รวมถึงถ่านหินที่เป็นแหล่งพลังงานสำคัญ ส่งผลให้รถไฟซึ่งเป็นการผสมผสานระหว่างไอน้ำ เหล็ก และถ่านหิน กลายเป็นยานพาหนะที่ได้รับความนิยมสูงขึ้นในยุคนั้นเช่นกัน [12, 13]

2. การปฏิวัติอุตสาหกรรมครั้งที่ 2 ถึงแม้ว่าเหล็กจะถูกนำมาใช้ตั้งแต่ช่วงของการปฏิวัติครั้งแรก แต่ก็ไม่ได้ถูกใช้อย่างแพร่หลายเหมือนช่วงของการปฏิวัติอุตสาหกรรมครั้งนี้ในศตวรรษที่ 19 ที่นับเป็นจุดเริ่มต้นของการใช้ไฟฟ้าในอุตสาหกรรมและการผลิตแบบจำนวนมาก (Mass Production) [10]

3. การปฏิวัติอุตสาหกรรมครั้งที่ 3 หรือที่เรียกว่ายุคดิจิทัล ถือเป็นยุคที่สิ่งประดิษฐ์อย่างคอมพิวเตอร์นับเป็นปัจจัยสำคัญที่ก่อให้เกิดการปฏิวัติอุตสาหกรรมในครั้งนี้ เนื่องจากมีการนำเทคโนโลยีที่เกี่ยวข้องกันนี้เข้าไปใช้ในอุตสาหกรรมอย่างแพร่หลาย ระบบอิเล็กทรอนิกส์ก็ได้รับความนิยมอย่างมากในกระบวนการผลิตแบบอัตโนมัติ [12, 13]

4. การปฏิวัติอุตสาหกรรมครั้งที่ 4 หรือที่เรียกว่า อุตสาหกรรม 4.0 เป็นการผสมผสานระหว่างจุดแข็งของกระบวนการผลิตในกลุ่มอุตสาหกรรมและเทคโนโลยีอินเตอร์เน็ตเข้าด้วยกัน รวมถึงการจัดการและกลยุทธ์ด้านการบำรุงรักษาที่จะมีการเปลี่ยนแปลงเช่นกัน [14] ถือได้ว่าเป็นยุคที่อินเตอร์เน็ตของสรรพสิ่ง (Internet of Things: IoT) และระบบไซเบอร์-กายภาพ (Cyber-Physical System: CPS) ทำงานร่วมกันกับซอฟต์แวร์ (Software) เซ็นเซอร์ (Sensor) หน่วยประมวลผล

(Processor) และเทคโนโลยีการสื่อสาร ซึ่งจะมีบทบาทอย่างมากในการเพิ่มศักยภาพการทำงานของเครื่องจักรในกระบวนการผลิต [13]



รูปที่ 2.1 การปฏิวัติอุตสาหกรรม

ถึงแม้ว่าการเปลี่ยนแปลงที่เกิดขึ้นจะค่อยเป็นค่อยไปตามความพร้อมของแต่ละอุตสาหกรรมแต่ก็หลีกเลี่ยงไม่ได้ที่จะต้องพัฒนาอุตสาหกรรมให้ก้าวทันยุคสมัยที่เปลี่ยนไป ความก้าวหน้าในแง่มุมที่เกี่ยวข้องจะมีผลกระทบอย่างมากต่อชีวิตทางสังคมและจะกระตุ้นให้อุตสาหกรรมการผลิตพัฒนาขีดความสามารถเพื่อตอบสนองความต้องการของผู้บริโภคมากยิ่งขึ้นด้วยเหตุผลที่ว่าความได้เปรียบในการแข่งขันจะขึ้นอยู่กับความเชี่ยวชาญระดับสูงของเทคโนโลยีเหล่านี้ [12, 15]

2.2 ข้อมูลขนาดใหญ่

คำว่าข้อมูลขนาดใหญ่ (Big Data) ถูกกล่าวถึงครั้งแรกโดย L. Doug [16] ซึ่งหากจะกล่าวถึงความหมายหรือคำจำกัดความของข้อมูลขนาดใหญ่ ก็อาจแตกต่างกันตามแต่ละบุคคลหรือองค์กร รวมถึงการนำไปใช้งานและการสร้างมูลค่าจากข้อมูลที่มี ซึ่งอาจหมายถึงชุดข้อมูลที่ไม่สามารถจัดการได้ด้วยการจัดการฐานข้อมูลแบบเดิม เนื่องจากข้อมูลมีขนาดใหญ่และได้มาอย่างรวดเร็วมากจนกระทั่งระบบการประมวลผลแบบเดิมไม่สามารถประมวลผลได้อย่างครอบคลุม [17] แต่เดิมในอดีตการที่มนุษย์จัดเก็บข้อมูลลงบนหินเป็นสิ่งที่บ่งบอกได้ว่ามนุษย์นั้นทราบถึงวิธีจัดเก็บข้อมูลและประมวลผลมาเป็นระยะเวลานานจนกระทั่งมีคอมพิวเตอร์เหมือนในปัจจุบัน พัฒนาการด้านการสื่อสารมีความก้าวหน้าอย่างรวดเร็วเมื่อคอมพิวเตอร์ได้เข้ามามีบทบาทสำคัญกับมนุษย์ ส่งผลให้การจัดเก็บข้อมูล การประมวลผล และเทคนิคที่ใช้นำมาวิเคราะห์มีความทันสมัยมากขึ้น ปัจจุบันในส่วนของกระบวนการผลิตมีการเก็บข้อมูลด้วยการนำไมโครโปรเซสเซอร์ (Microprocessors) มาติดตั้งไว้

บนเครื่องจักรจำนวนมาก อุปกรณ์เหล่านี้ได้สร้างแหล่งข้อมูลขนาดใหญ่มากเมื่อเปรียบเทียบกับสองทศวรรษที่ผ่านมา [10, 18] นอกเหนือจากการเก็บข้อมูลในกระบวนการผลิตแล้ว ข้อมูลการออกแบบ ข้อมูลพฤติกรรมพนักงาน ข้อมูลค่าใช้จ่าย ข้อมูลด้านโลจิสติกส์ ข้อมูลปัจจัยสภาพแวดล้อม สถานะของระบบภายในองค์กร การตรวจจับข้อผิดพลาด ข้อมูลคุณภาพสินค้า และข้อมูลลูกค้า ล้วนถูกจัดเก็บในรูปแบบของข้อมูลขนาดใหญ่ทั้งสิ้น [19]

การที่จะตัดสินใจว่าข้อมูลประเภทไหนจะถูกจัดว่าเป็นข้อมูลขนาดใหญ่นั้นจำเป็นที่จะต้องเข้าใจถึงลักษณะเฉพาะ (Characteristics) ของข้อมูลขนาดใหญ่เช่นกัน แต่เดิมลักษณะเฉพาะของข้อมูลขนาดใหญ่มีเพียงสามข้อเท่านั้น ได้แก่ ความหลากหลายของข้อมูล ความเร็วของข้อมูล และขนาดของข้อมูล [16] ซึ่งต่อมาได้มีการเพิ่มความไม่ชัดเจนเป็นลักษณะเฉพาะข้อที่สี่ของข้อมูลขนาดใหญ่ด้วย

1. ขนาดของข้อมูล (Volume) ข้อมูลขนาดใหญ่จะมีปริมาณข้อมูลจำนวนมากและจะเพิ่มจำนวนขึ้นอย่างต่อเนื่องไม่มีที่สิ้นสุด การนำข้อมูลที่มีอยู่อย่างมหาศาลนี้มาทำการวิเคราะห์อย่างรอบคอบจะช่วยให้สามารถตัดสินใจได้อย่างถูกต้องและเป็นประโยชน์กับอุตสาหกรรม [20, 21]

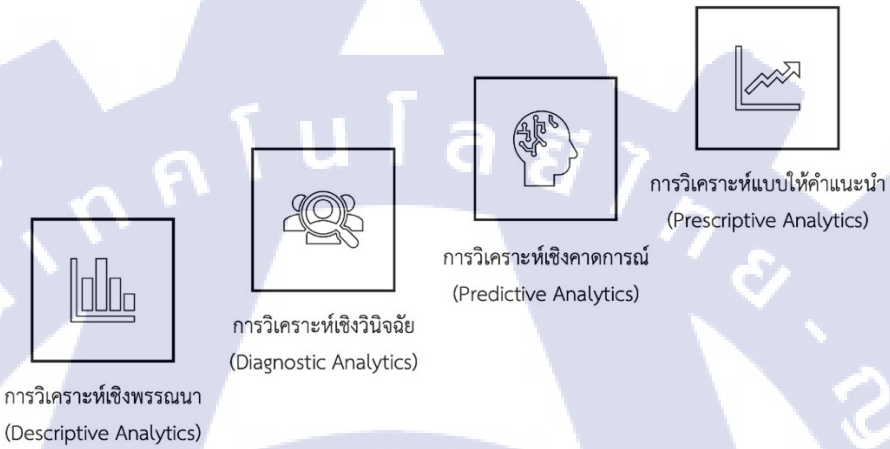
2. ความหลากหลายของข้อมูล (Variety) ข้อมูลที่ได้มาจากแหล่งข้อมูลที่แตกต่างกันล้วนส่งผลให้มีลักษณะรูปแบบของข้อมูลที่แตกต่างกัน เมื่อข้อมูลมีความหลากหลายจึงมีการจัดจำแนกข้อมูลออกเป็น 3 ประเภทหลักได้แก่ แบบโครงสร้าง (Structured) ซึ่งหมายถึงข้อมูลที่มีรูปแบบเฉพาะและเป็นระเบียบ แบบกึ่งโครงสร้าง (Semi-Structured) หมายถึงข้อมูลที่คงรูปแบบไว้บางส่วน และแบบไม่มีโครงสร้าง (Unstructured) คือข้อมูลที่ไม่มีรูปแบบเฉพาะ [22, 23]

3. ความเร็วของข้อมูล (Velocity) หรืออัตราที่ข้อมูลไหลเข้าสู่ระบบจากแหล่งที่มาของข้อมูลซึ่งจะมีความเร็วสูงมาก ดังนั้นการประมวลผลที่รวดเร็วจึงมีความจำเป็นอย่างยิ่งกับข้อมูลขนาดใหญ่ เพื่อให้สามารถใช้ข้อมูลได้อย่างมีประสิทธิภาพในเวลาที่เหมาะสม [20]

4. ความน่าเชื่อถือ (Veracity) เมื่อมีปริมาณข้อมูลมากย่อมมีทั้งข้อมูลที่เป็นจริงและเป็นเท็จเสมอ ดังนั้นความน่าเชื่อถือจึงเป็นลักษณะเฉพาะของข้อมูลขนาดใหญ่ที่ถูกเพิ่มเข้ามาเพื่อแสดงให้เห็นว่าความน่าเชื่อถือหรือความไม่แน่นอนเป็นส่วนสำคัญสำหรับการนำข้อมูลขนาดใหญ่ไปใช้ หากข้อมูลมีปริมาณมากแต่ไม่มีความน่าเชื่อถือหรือไม่ถูกต้องก็อาจไม่เป็นประโยชน์ [24]

ในยุคของอุตสาหกรรม 4.0 การนำคอมพิวเตอร์มาประยุกต์ใช้ในการวิเคราะห์ข้อมูลหรือที่เรียกกันว่า Data Analytics เข้ามามีบทบาทอย่างเห็นได้ชัด ความรู้จากหลากหลายสาขาถูกนำมาประยุกต์ใช้ร่วมกัน [25] โดยส่วนใหญ่แล้วการวิเคราะห์ข้อมูลแบบนี้จะถูกนำมาใช้ร่วมกับข้อมูลขนาด

ใหญ่ เนื่องจากคุณลักษณะเฉพาะของข้อมูลขนาดใหญ่ไม่สามารถใช้การวิเคราะห์และประมวลผลแบบเดิมได้ เพื่อความเข้าใจในการวิเคราะห์ข้อมูลที่มากขึ้นซึ่งจะช่วยให้สามารถวิเคราะห์ข้อมูลได้ตรงตามวัตถุประสงค์ที่ต้องการ ก็จำเป็นจะต้องศึกษาถึงประเภทของการวิเคราะห์ข้อมูลทั้งสี่ประเภท (รูปที่ 2.2) ได้แก่ การวิเคราะห์เชิงพรรณนา การวิเคราะห์เชิงวินิจฉัย การวิเคราะห์เชิงคาดการณ์ และการวิเคราะห์แบบให้คำแนะนำ [18]



รูปที่ 2.2 ประเภทของการวิเคราะห์

1. การวิเคราะห์เชิงพรรณนา (Descriptive Analytics) เป็นการวิเคราะห์ที่อาศัยข้อมูลที่มีอยู่โดยไม่จำเป็นต้องสร้างแบบจำลอง ซึ่งหมายถึงการนำรายงานต่าง ๆ ที่มีอยู่ในอดีตมาทำการวิเคราะห์ถึงสิ่งที่ต้องการ เช่นการวิเคราะห์ความสัมพันธ์ ความเชื่อมโยง หรือรูปแบบ เป็นต้น [26]

2. การวิเคราะห์เชิงวินิจฉัย (Diagnostic Analytics) เป็นการวิเคราะห์ถึงสาเหตุ เหตุผล หรือปัจจัยต่าง ๆ ที่ส่งผลกระทบหรือก่อให้เกิดเหตุการณ์นั้นขึ้น รวมถึงการคิดเชื่อมโยงถึงเหตุและผลด้วยเช่นกัน สามารถนำข้อมูลที่มีอยู่มาใช้สำหรับการวิเคราะห์ประเภทนี้ได้ [27]

3. การวิเคราะห์เชิงคาดการณ์ (Predictive Analytics) เป็นการนำความรู้ที่เกี่ยวข้องและข้อมูลในอดีตมานำเสนอในรูปแบบของโมเดลเพื่อวิเคราะห์ข้อมูลและคาดการณ์ถึงสิ่งที่จำเกิดขึ้นในอนาคต การสร้างโมเดลจากเทคนิคการเรียนรู้ของเครื่องแบบมีผู้สอนถือเป็นส่วนหนึ่งของเทคนิคที่ใช้สำหรับวิเคราะห์ข้อมูลประเภทนี้เช่นกัน [26]

4. การวิเคราะห์แบบให้คำแนะนำ (Prescriptive Analytics) เป็นการวิเคราะห์ที่มุ่งเน้นถึงแนวทางที่จะดำเนินการต่อไปโดยอ้างอิงจากข้อมูลและปัจจัยต่าง ๆ ที่เกี่ยวข้อง รวมถึงผลลัพธ์ที่ได้

จากแนวทางนั้นซึ่งอาจมีได้หลายแนวทาง การวิเคราะห์นี้จะช่วยให้ทราบถึงผลที่จะได้รับจากแนวทางปฏิบัติที่หลากลาย [28]

เพื่อเพิ่มความได้เปรียบในการแข่งขันกับอุตสาหกรรมเดียวกัน ข้อมูลเชิงลึกทางธุรกิจจากในอดีตถูกนำมาใช้เพื่อประโยชน์สูงสุด เช่น การปรับกระบวนการให้เหมาะสม การลดต้นทุน การปรับปรุงประสิทธิภาพการดำเนินงาน หรือแม้กระทั่งการบำรุงรักษาเชิงพยากรณ์(Predictive Maintenance) [10, 18, 19] ข้อมูลขนาดใหญ่ที่ถูกจัดเก็บจากอุปกรณ์ต่าง ๆ อาทิเช่น เซนเซอร์หรือไมโครโปรเซสเซอร์ แม้กระทั่งข้อมูลภายนอกที่เกี่ยวข้องจะถูกนำขึ้นไปยังระบบคลาวด์ (Cloud System) ผ่านกระบวนการวิเคราะห์และประมวลผลเพื่อประกอบการตัดสินใจ [10] นอกจากนี้การลงทุนในข้อมูลและอุปกรณ์ที่มีคุณภาพเพียงพอจะมีส่วนในการขับเคลื่อนโอกาสทางธุรกิจ เพื่อให้การนำข้อมูลขนาดใหญ่มาใช้ประโยชน์ได้จริงจำเป็นต้องมีผู้เชี่ยวชาญหรือนักวิเคราะห์ที่มีทักษะความรู้ในด้านนี้ร่วมด้วยไม่เพียงแต่เฉพาะโครงสร้างพื้นฐานและเทคโนโลยีที่เกี่ยวข้องเท่านั้น [21] เทคนิคที่เลือกใช้ก็มีส่วนสำคัญที่จะทำให้การวิเคราะห์ข้อมูลขนาดใหญ่เป็นไปอย่างมีประสิทธิภาพในระยะเวลาที่จำกัด จากการศึกษางานวิจัยพบว่ามีหลากหลายเทคนิคที่ใช้สำหรับวิเคราะห์ข้อมูลขนาดใหญ่ดังที่จะกล่าวต่อไปนี้

1. การทำเหมืองข้อมูล (Data Mining) เป็นกระบวนการที่จะวิเคราะห์ค้นหารูปแบบและความเชื่อมโยงในข้อมูลที่มีปริมาณมาก เพื่อให้สามารถใช้ประโยชน์จากข้อมูลที่มีได้ ถือเป็นเทคนิคหลักที่ใช้สำหรับการวิเคราะห์เชิงพยากรณ์ [29]

2. การวิเคราะห์เครือข่ายทางสังคม (Social Network Analysis: SNA) เป็นการแสดงความสัมพันธ์ทางสังคมของระบบในทฤษฎีเครือข่ายทางสังคม ซึ่งเป็นวิธีที่ได้รับความนิยมทั้งในทางสังคมและทั้งการประมวลผลแบบคลาวด์(Cloud Computing) [30]

3. การเรียนรู้ของเครื่อง (Machine Learning) เป็นเทคนิคพื้นฐานของปัญญาประดิษฐ์ มีทั้งที่เป็นการเรียนรู้แบบมีผู้สอน (Supervised) ไม่มีผู้สอน (Unsupervised) และแบบเสริมกำลัง (Reinforcement) ซึ่งความฉลาดของเครื่องจะขึ้นอยู่กับอัลกอริทึมและข้อมูลที่ใช้สอน

4. แนวทางจินตทัศน์ (Visualization Approaches) ช่วยให้สามารถทำความเข้าใจข้อมูลได้ง่ายขึ้นโดยการสร้างเป็นรูปภาพ ตาราง หรือไดอะแกรม (Diagram) แต่การที่ใช้เทคนิคนี้สำหรับข้อมูลขนาดใหญ่นั้นยังมีความยากกว่าข้อมูลขนาดเล็กแบบดั้งเดิมอยู่ [30]

5. วิธีการหาค่าเหมาะสมที่สุด (Optimization Methods) ใช้สำหรับปัญหาในเชิงปริมาณช่วยเพิ่มประสิทธิภาพให้กับข้อมูลขนาดใหญ่ ให้ประสิทธิภาพสูงแต่ก็แลกมาด้วยความซับซ้อนและใช้

ระยะเวลาพอสมควร รวมถึงความต้องการในการทำงานแบบเรียลไทม์(Real-Time) เพื่อที่จะประมวลผลข้อมูลขนาดใหญ่ [30]

2.3 ปัญญาประดิษฐ์และการเรียนรู้ของเครื่อง

ในปี พ.ศ. 2499 คำว่าปัญญาประดิษฐ์ (Artificial Intelligence: AI) ได้ถูกนิยามขึ้นเป็นครั้งแรกโดย John McCarthy ซึ่งมีแนวคิดพื้นฐานมาจากมนุษย์กล่าวคือ ปัญญาประดิษฐ์เป็นแนวคิดที่ต้องการให้เครื่องจักรหรือคอมพิวเตอร์สามารถเรียนรู้ทักษะบางอย่างจนกระทั่งมีความฉลาดแบบสติปัญญาของมนุษย์ [31, 32] แน่นอนว่าคำกล่าวของ John McCarthy สร้างแรงกระตุ้นให้ผู้คนหันมาสนใจสิ่งนี้อย่างมหาศาล มีการศึกษาพัฒนาความรู้ในด้านนี้อย่างต่อเนื่องตั้งแต่อดีตถึงปัจจุบัน และได้มีการแบ่งระดับของปัญญาประดิษฐ์ออกเป็นสามระดับโดยอ้างอิงตามความฉลาดได้แก่ ปัญญาประดิษฐ์ระดับอ่อน ปัญญาประดิษฐ์ระดับเข้ม และปัญญาประดิษฐ์ระดับสุดยอด [33, 34, 35]

1. ปัญญาประดิษฐ์ระดับอ่อน (Weak AI) หมายถึงปัญญาประดิษฐ์ที่มีความสามารถเหนือมนุษย์ในเฉพาะด้าน กล่าวคือเครื่องจักรหรือคอมพิวเตอร์จะมีสติปัญญาและความเชี่ยวชาญที่มากกว่ามนุษย์แต่แค่เฉพาะสาขาเท่านั้น เช่น ระบบผู้เชี่ยวชาญ (Expert System) เป็นต้น [34] อย่างไรก็ตามหากนำระบบดังกล่าวไปใช้งานนอกเหนือจากที่ได้ทำการฝึกฝนไว้จะไม่สามารถทำงานได้

2. ปัญญาประดิษฐ์ระดับเข้ม (Strong AI) หมายถึงการที่เครื่องจักรหรือคอมพิวเตอร์มีสติปัญญาและความฉลาดในการใช้เหตุผลได้เหมือนกับมนุษย์รวมถึงการมีสภาวะอารมณ์และจิตใจที่เทียบเท่ามนุษย์ด้วยเช่นกัน [35, 34] อย่างไรก็ตามระบบดังกล่าวยังไม่ถูกพัฒนาให้สามารถทำงานได้ในระดับนี้

3. ปัญญาประดิษฐ์ระดับสุดยอด(Artificial Superintelligence: ASI) หมายถึงปัญญาประดิษฐ์ที่มีความฉลาดเหนือมนุษย์ในทุกด้าน ถึงแม้ว่าจะไม่มีอยู่จริงในขณะนี้แต่ก็เป็นที่คาดเดาว่าในอนาคตอาจมีปัญญาประดิษฐ์ระดับนี้เกิดขึ้น ซึ่งยังคงเป็นที่ถกเถียงกันถึงอันตรายและหายนะที่อาจตามมาอยู่ [33]

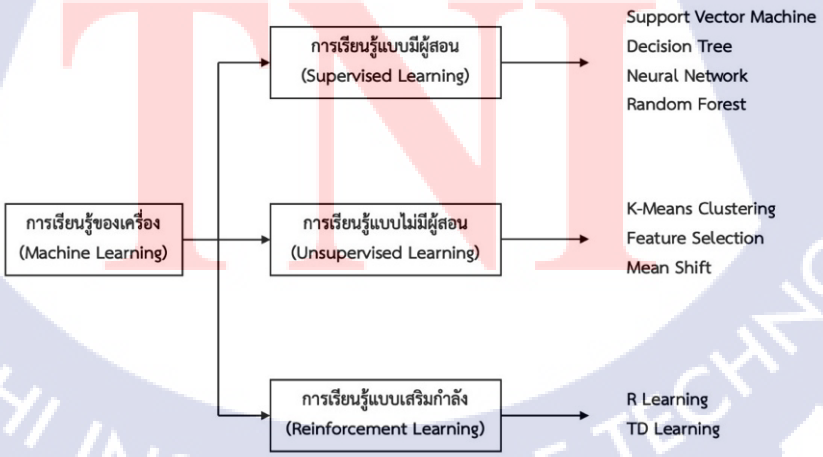
สำหรับการที่จะทำให้เครื่องจักรหรือคอมพิวเตอร์มีความฉลาดหรือกลายเป็นปัญญาประดิษฐ์ได้นั้นจะอาศัยสิ่งที่เรียกว่า การเรียนรู้ของเครื่อง (Machine Learning) เป็นหลัก ซึ่งการเรียนรู้ของเครื่องเป็นศาสตร์ที่มีความเกี่ยวข้องกับการพัฒนาอัลกอริทึม (Algorithm) และโมเดล (Model) หลักการพื้นฐานคือคอมพิวเตอร์จะเรียนรู้งานผ่านประสบการณ์ซึ่งก็คือการฝึกฝน (Training) ด้วยชุดข้อมูล (Datasets) เมื่อมีประสบการณ์มากขึ้นหรือผ่านการฝึกฝนอย่างเพียงพอ

คอมพิวเตอร์จะสามารถประมวลผลหรือตัดสินใจข้อมูลใหม่โดยอ้างอิงตามสิ่งที่เรียนรู้มาได้ [36] แนวทางการเรียนรู้ของเครื่องถูกแบ่งออกเป็นสามประเภทหลักตามรูปที่ 2.3 ได้แก่การเรียนรู้แบบมีผู้สอน การเรียนรู้แบบไม่มีผู้สอน และ การเรียนรู้แบบเสริมกำลัง

1. การเรียนรู้แบบมีผู้สอน (Supervised Learning) หมายถึงการเรียนรู้จากการหาความสัมพันธ์ระหว่างข้อมูลนำเข้า (Input) และค่าเป้าหมาย (Target) เพื่อพยากรณ์ผลลัพธ์ (Output) สำหรับข้อมูลที่ไม่เคยพบ (Unseen Data) สามารถจำแนกได้เป็นสองกลุ่มคือ การถดถอย (Regression) ซึ่งหมายถึงตัวแปรที่ได้มาเป็นค่าจริง (Real Value) หรือค่าต่อเนื่อง (Continuous Value) และการจำแนกประเภท (Classification) ที่ตัวแปรที่ได้มาจะอยู่ในกลุ่มใดกลุ่มหนึ่ง การเรียนรู้ประเภทนี้จำเป็นต้องใช้ชุดข้อมูลฝึกสอนมากเพียงพอที่จะให้คอมพิวเตอร์สามารถเรียนรู้กระทั่งมีความเชี่ยวชาญ [37, 38]

2. การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) จะเป็นการเรียนรู้แบบที่ไม่มีคำตอบหรือข้อเสนอแนะ (Feedback) ให้จะมีเพียงชุดข้อมูลที่ประกอบไปด้วยคุณลักษณะต่าง ๆ เท่านั้น ซึ่งอัลกอริทึมการเรียนรู้ประเภทนี้จะเน้นที่การจัดกลุ่มและการลดขนาดเป็นส่วนใหญ่ [39]

3. การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) หมายถึงการเรียนรู้ที่จะตัดสินใจหรือกระทำการบางอย่างกับสถานการณ์ที่พบด้วยการลองผิดลองถูก (Trial-And-Error) ซึ่งจะไม่มีการป้อนคำสั่งหรือการสอนใด ๆ กับเครื่องคอมพิวเตอร์ทั้งสิ้น และตัวชี้วัดถึงความสำเร็จในการเรียนรู้จะเป็นผลตอบแทนที่ได้รับ ความพยายามที่จะได้มาซึ่งรางวัลสูงสุดจะเป็นแรงกระตุ้นให้เกิดการเรียนรู้ประเภทนี้ [37]

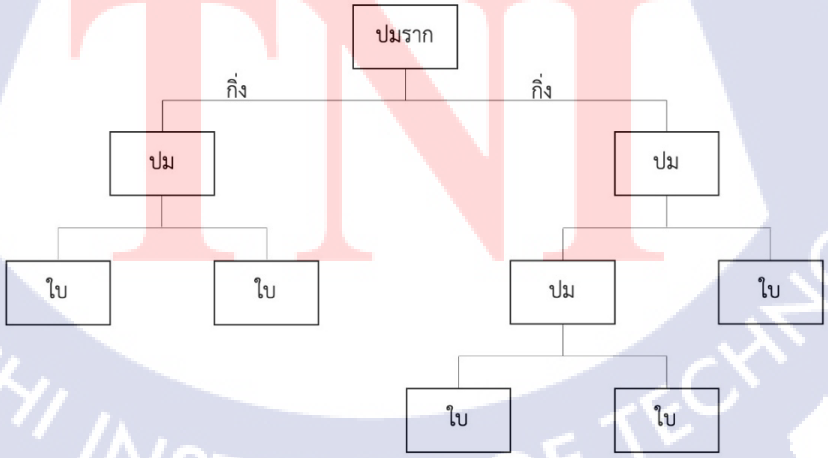


รูปที่ 2.3 ประเภทของการเรียนรู้ของเครื่องและตัวอย่างอัลกอริทึม [40]

โดยหลักแล้วจะมีการนำการเรียนรู้ของเครื่องไปใช้แก้ปัญหาอยู่สามประเภทได้แก่ การจำแนกประเภท การถดถอย และการจัดกลุ่ม (Clustering) ซึ่งจะใช้อัลกอริทึมที่แตกต่างกันไปตามประเภทของปัญหา แต่มีบางอัลกอริทึมที่สามารถแก้ปัญหาได้ทั้งสองแบบ เช่น ต้นไม้ตัดสินใจ (Decision Tree) เป็นต้น ในปัจจุบันการเรียนรู้ของเครื่องค่อนข้างเป็นที่นิยมมาก ด้วยเหตุผลที่ว่าสามารถแก้ปัญหาที่มีความซับซ้อนได้จึงมีการนำการเรียนรู้ของเครื่องไปประยุกต์ใช้ในหลายด้าน เช่น หุ่นยนต์(Robotics) การรู้จำรูปแบบ (Pattern Recognition) การทำเหมืองข้อมูล (Data Mining) การประมวลผลภาษาธรรมชาติ (Natural Language Processing) การแนะนำสินค้า (Product Recommendation) ผู้ช่วยส่วนตัวเสมือน (Virtual Personal Assistants) และอื่น ๆ [36]

2.4 ต้นไม้ตัดสินใจ

ต้นไม้ตัดสินใจ (Decision Tree) เป็นอัลกอริทึมที่เป็นที่รู้จักและมีความแพร่หลายเนื่องจากมีรูปแบบที่เข้าใจง่ายและมีประสิทธิภาพ มีลักษณะการเรียนรู้แบบมีผู้สอนสำหรับแก้ปัญหาได้ทั้งการจำแนกประเภท (Classification) ที่ผลลัพธ์จะเป็นกลุ่มหรือหมวดหมู่ (Categorical) และการถดถอย (Regression) ที่ตัวแปรจะเป็นแบบต่อเนื่อง (Continuous) แต่ส่วนใหญ่จะนิยมใช้สำหรับการจำแนกประเภท ต้นไม้ตัดสินใจมีองค์ประกอบหลักสามอย่างด้วยกันได้แก่ ปม (Node) กิ่ง (Branch) และใบ (Leaf) ในแต่ละปมจะแสดงถึงลักษณะประจำ (Attributes) ในหมวดหมู่และจะมีปมราก (Root Node) เป็นปมแรก การตัดสินใจหรือผลลัพธ์ที่ได้จะถูกแทนด้วยใบ ซึ่งจะมีกิ่งสำหรับแสดงค่าของแต่ละคุณลักษณะดังแสดงในรูปที่ 2.4



รูปที่ 2.4 ต้นไม้ตัดสินใจ (Decision Tree) [41]

ต้นไม้ตัดสินใจแบบ Iterative Dichotomiser 3 หรือ ID3 ได้ถูกนำเสนอขึ้นในปี พ.ศ. 2529 โดย J.R.Quinlan [42] ซึ่งมีหลักการพื้นฐานคือการทำซ้ำและมีการใช้ค่าความไม่แน่นอนหรือเอนโทรปี (Entropy) มาใช้ในการหาค่า Information Gain เพื่อเลือกลักษณะประจำในแต่ละปม โดยค่าเอนโทรปีและ Information Gain สามารถหาได้จากสมการ (2.1) และ (2.2) ตามลำดับ [43]

$$H = -\sum_{i=1}^{n} P_i log_2(P_i) \tag{2.1}$$

$$IG = H(T) - H(A) \tag{2.2}$$

เมื่อ P_i คือความน่าจะเป็นที่จะเกิดเหตุการณ์ในแต่ละเหตุการณ์ขึ้น

n คือเหตุการณ์ทั้งหมดที่เกิดขึ้น

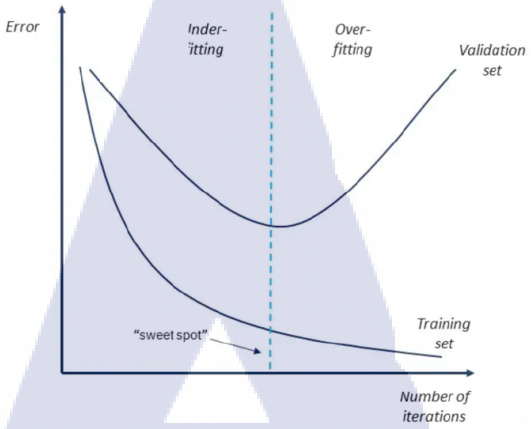
$H(\mathrm{T})$ คือค่าเอนโทรปีของทั้งหมด

$\mathrm{H(A)}$ คือค่าเอนโทรปีของลักษณะประจำที่เลือก

การเลือกลักษณะประจำของแต่ละปมจะทำการค่าหา Information Gain จากสมการที่ได้กล่าวไว้ข้างต้นด้วยการคำนวณซ้ำในทุกเงื่อนไขหรือทุกลักษณะประจำ เพื่อหาค่าที่มากที่สุดซึ่งหมายถึงลักษณะประจำที่เหมาะสมที่สุดสำหรับการคำนวณในรอบนั้น จากนั้นจึงทำการคำนวณแบบเดียวกันในเงื่อนไขหรือลักษณะประจำที่เหลือเพื่อหาลักษณะประจำที่เหมาะสมในแต่ละปมต่อไป

2.4.1 Generalization

ปัญหา Overfitting เป็นปัญหาที่พบได้บ่อยในการฝึกฝนโมเดล ซึ่งหมายถึงการที่วัดประสิทธิภาพจากข้อมูลทดสอบ (Test Data) ได้ต่ำกว่าข้อมูลฝึกฝน (Train Data) ที่ใช้ฝึกสอนโมเดล เมื่อโมเดลมีความซับซ้อนสูงขึ้น หรือกล่าวอย่างง่ายว่าโมเดลขณะฝึกสอนให้ผลลัพธ์ที่ดีกว่าความเป็นจริงดังรูปที่ 2.5 เมื่อนำไปใช้กับชุดข้อมูลที่โมเดลไม่เคยเจอจะส่งผลให้ประสิทธิภาพลดลง ดังนั้นวิธีการทำ Generalization จึงได้รับการเสนอเพื่อแก้ไขปัญหาที่พบดังกล่าว [44]



รูปที่ 2.5 Overfitting

การทำ Generalization เป็นการทำให้โมเดลนั้นสามารถใช้กับข้อมูลทั่วไปได้อย่างมีประสิทธิภาพ โดยทั่วไปแล้วการป้องกันการเกิด Overfitting สามารถทำได้โดยเตรียมชุดข้อมูลสำหรับการฝึกฝนที่ครอบคลุมและมีจำนวนมากเพียงพอที่จะฝึกฝนโมเดลให้มีความรู้มากขึ้น แต่ในความเป็นจริงยังคงมีอุปสรรคอยู่มากสำหรับการแก้ปัญหาด้วยวิธีนี้เนื่องจากการมีอยู่อย่างจำกัดของข้อมูลที่ใช้ฝึกสอน ดังนั้นการตัดทอน (Pruning) จึงเป็นวิธีที่ถูกใช้สำหรับการทำ Generalization ให้กับโมเดล โดยการตัดทอนสามารถทำได้สองลักษณะได้แก่ การตัดทอนก่อนที่ต้นไม้จะสมบูรณ์และการตัดทอนหลังจากที่ต้นไม้สมบูรณ์แล้ว

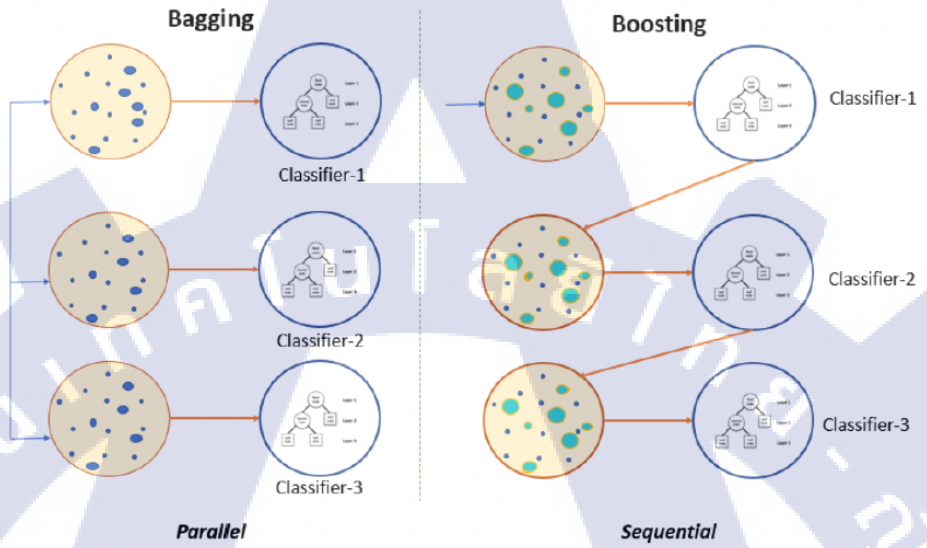
1. การตัดทอนก่อนที่ต้นไม้จะสมบูรณ์ (Pre-Pruning) หมายถึงการพยายามตัดกิ่งหรือคุณสมบัติที่พิจารณาแล้วว่าไม่มีความสำคัญหรือไม่น่าเชื่อถือก่อนที่โมเดลจะถูกสร้างเสร็จ

2. การตัดทอนหลังจากที่ต้นไม้สมบูรณ์ (Post-Pruning) หมายถึงการตัดกิ่งในลักษณะเดียวกัน แต่จะตัดหลังจากที่โมเดลถูกสร้างเสร็จสมบูรณ์แล้ว

2.5 ป่าแบบสุ่ม

ถึงแม้ว่าต้นไม้ตัดสินใจจะสามารถเข้าใจได้ง่ายและมีประสิทธิภาพ แต่ก็พบว่ามักจะไม่มีความเสถียร (Unstable) ควบคุมขนาดได้ยาก และอาจเจอปัญหาที่โมเดลขณะฝึกสอนมีประสิทธิภาพดีกว่าความเป็นจริง (Overfitting) จึงได้มีการพัฒนาอัลกอริทึมป่าแบบสุ่ม (Random Forest) มาเพื่อแก้ปัญหานี้ [41] ป่าแบบสุ่มจัดเป็นเทคนิคการเรียนรู้ของเครื่องที่พัฒนามาจากต้นไม้ตัดสินใจ ถือเป็นแนวทางการเรียนรู้แบบ Ensemble คือการรวมหลายโมเดลเข้าด้วยกันสำหรับแก้ปัญหาที่มีความ

ซับซ้อนสามารถ ใช้ได้ทั้งการจำแนกประเภทและการถดถอย แบ่งออกเป็นสองประเภทได้แก่ Bagging ซึ่งเป็นการเรียนรู้แบบแยกจากกันโดยอิสระแล้วจึงนำมาประมวลผลในตอนสุดท้ายและ Boosting ซึ่งเป็นการเรียนรู้แบบที่จะทำการเรียนรู้และปรับปรุงแบบจำลองไปทีละขั้นตอนดังรูปที่ 2.6



รูปที่ 2.6 Bagging และ Boosting [45]

2.5.1 Bagging

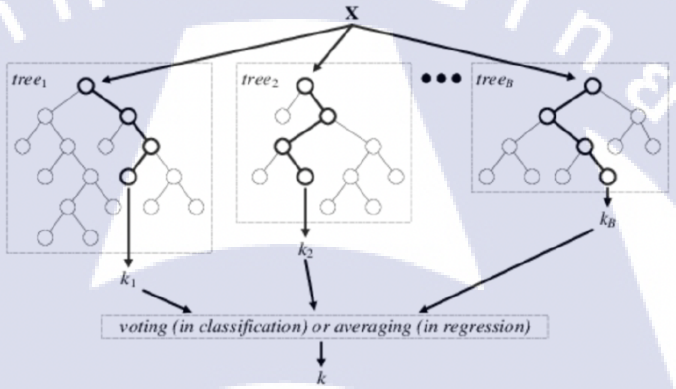
วิธี Bootstrap Aggregating หรือที่เรียกกันว่า Bagging นี้ถูกนำเสนอครั้งแรกในปี พ.ศ. 2539 โดย L. Breiman [46] วิธีการนี้จะทำการแบ่งชุดข้อมูลออกเป็นกลุ่มย่อย ด้วยการสุ่มตัวอย่างจากชุดข้อมูลเดิมพร้อมการแทนที่เข้าไป กลุ่มของข้อมูลใหม่จะมีขนาดเล็กกว่าชุดข้อมูลและตัวอย่างในแต่ละกลุ่มย่อยจะต้องไม่เหมือนกันทั้งหมด กล่าวคือสามารถสุ่มตัวอย่างซ้ำได้บางส่วนเท่านั้น วิธีการ Bagging มีขั้นตอนพื้นฐานอยู่สามขั้นตอนได้แก่การบูทแสตรป การฝึกแบบคู่ขนาน และการรวม

1. การบูทแสตรป (Bootstrapping) เป็นขั้นตอนการใช้เทคนิคบูทแสตรปเพื่อสุ่มตัวอย่างในแต่ละเซทย่อยให้มีความหลากหลาย

2. การฝึกแบบคู่ขนาน (Parallel Training) เป็นขั้นตอนการฝึกฝนของแต่ละเซทย่อยซึ่งจะไม่มีความเกี่ยวข้องกันโดยสิ้นเชิง

3. การรวม (Aggregation) เป็นขั้นตอนที่จะนำผลที่ได้จากการฝึกฝนในแต่ละเซทย่อยมาประมวลผลร่วมกัน โดยจะแบ่งออกเป็นกรณีของการถดถอยซึ่งจะใช้การเฉลี่ยผลที่ได้หรือที่เรียกว่า Soft Voting ในส่วนของกรณีของการจัดจำแนกจะใช้การลงคะแนนเสียงมากสุดที่เรียกว่า Hard Voting หรือ Majority Voting

ป่าแบบสุ่มถือเป็นอัลกอริทึมที่มีการเรียนรู้แบบ Bagging โดยป่าแบบสุ่มจะสร้างต้นไม้ตัดสินใจจำนวนมากขึ้นมา จากนั้นจะทำการสุ่มเลือกชุดข้อมูลให้กับต้นไม้แต่ละต้นแบบไม่ซ้ำกัน เพื่อทำการประมวลผลแยกแล้วจึงทำการโหวดเลือกค่าที่ดีที่สุดในกรณีที่เป็นการจำแนกประเภท และค่าเฉลี่ยสำหรับกรณีที่เป็นการถดถอยดังรูปที่ 2.7 [47]



รูปที่ 2.7 ป่าแบบสุ่ม [48]

2.6 Extreme Gradient Boosting

จากที่ได้กล่าวในหัวข้อป่าแบบสุ่ม การเรียนรู้แบบ Ensemble นั้นสามารถแบ่งได้เป็นสองประเภทได้แก่ Bagging และ Boosting การเรียนรู้ทั้งสองประเภทนี้ล้วนมีต้นแบบมาจากต้นไม้ตัดสินใจทั้งสิ้น แตกต่างเพียงแต่ขั้นตอนวิธีในการเรียนรู้ โดยป่าแบบสุ่มเป็นการเรียนรู้แบบ Bagging ซึ่งจะแตกต่างจากเทคนิค Extreme Gradient Boosting หรือ XGBoost ที่จะเป็นการเรียนรู้แบบ Boosting

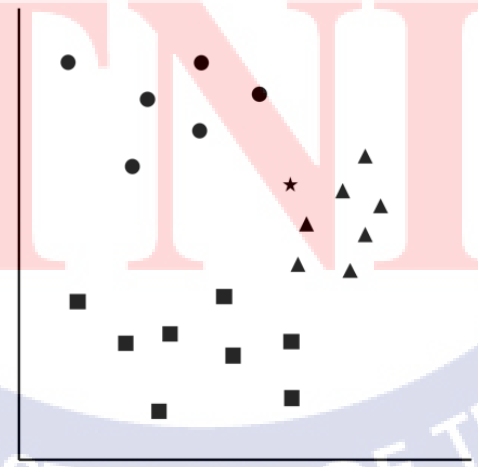
การเรียนรู้แบบ Boosting คือการเรียนรู้แบบเป็นลำดับ (Sequential) มีลักษณะการปรับปรุงแบบจำลองไปทีละขั้นตอนโดยอาศัยการเรียนรู้จากขั้นตอนก่อนหน้า ดังนั้นโมเดลที่พัฒนา

จากเทคนิคนี้จะมีการฝึกฝนจากการเรียนรู้ก่อนหน้าเพื่อปรับค่าความแม่นยำของโมเดล ในขณะที่เทคนิค Bagging นั้นจะแยกการเรียนรู้ออกจากกันอย่างชัดเจน

เทคนิค XGBoost เป็นเทคนิคที่ได้รับการพัฒนาเพิ่มเติมจากเทคนิค Gradient Boosting ให้สามารถทำงานได้อย่างรวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น ถึงแม้ว่าลักษณะการเรียนรู้แบบทำซ้ำนั้นจะส่งผลให้โมเดลมีความแม่นยำสูง แต่ก็อาจเกิดปัญหา Overfitting ได้ ดังนั้นเทคนิคนี้จึงลดปัญหาดังกล่าวด้วยการเพิ่ม Regularization [49]

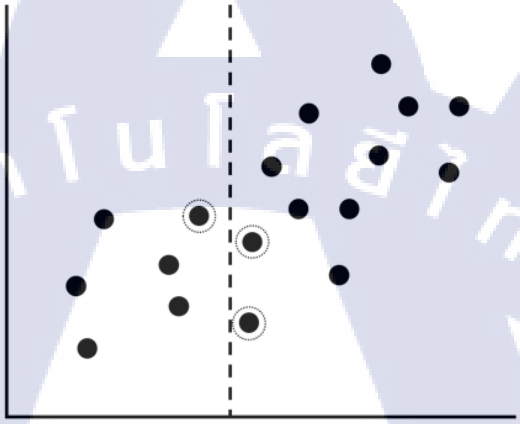
2.7 ขั้นตอนวิธีเพื่อนบ้านใกล้สุด

เทคนิคนี้ค่อนข้างเป็นที่นิยมสำหรับการจำแนกประเภทเนื่องด้วยหลักการพื้นฐานที่สามารถเข้าใจได้ง่ายคือ ให้จัดเข้ากลุ่มที่มีระยะห่างน้อยที่สุด วิธีการที่เทคนิคนี้ใช้คือการหาระยะห่างระหว่างข้อมูลใหม่กับข้อมูลที่มีอยู่เดิมจำนวน k ตัว จึงได้ชื่อว่าเทคนิค k-Nearest Neighbors เพื่อทำการจัดให้ข้อมูลใหม่เข้าไปอยู่ในกลุ่มของข้อมูลที่มีระยะห่างจากข้อมูลใหม่น้อยที่สุด เพื่อให้เห็นภาพได้อย่างชัดเจนยิ่งขึ้นจากรูปที่ 2.8 จะเห็นได้ว่ามีกลุ่มของข้อมูลอยู่ 3 กลุ่ม แต่ละกลุ่มจะมีการเกาะกลุ่มและระยะห่างที่แตกต่างกันออกไป เมื่อมีข้อมูลใหม่เพิ่มเข้ามา (ในที่นี้ให้เป็นรูปดาว) เทคนิคนี้จะทำการวัดระยะห่างจากจุดของข้อมูลที่เพิ่มเข้ามาใหม่กับจุดของข้อมูลที่กระจายตัวอยู่ในภาพ k ตัว หากกำหนดให้ k เท่ากับ 1 นั้นหมายความว่าข้อมูลใหม่จะถูกจัดเข้าไปอยู่ในกลุ่มเดียวกับชุดข้อมูลที่ใกล้ที่สุดทันทีโดยไม่ได้มีการเปรียบเทียบระยะห่างกับข้อมูลตัวถัดไป จากรูปภาพหากกำหนดให้ค่า k เท่ากับ 1 ข้อมูลใหม่จะถูกจัดให้อยู่ในกลุ่มของข้อมูลสามเหลี่ยมเป็นต้น [50]



รูปที่ 2.8 การจัดกลุ่มของ k-Nearest Neighbors

ถึงแม้ว่าเทคนิคที่กล่าวถึงนี้จะเป็นที่นิยมในการจัดจำแนกประเภทแต่ก็ยังสามารถใช้สำหรับการถดถอยได้เช่นกัน ในแง่ของการถดถอยเทคนิคนี้จะทำการหาระยะห่างตามสมการ Euclidean Distance จากรูปที่ 2.9 ให้เส้นประที่ตั้งฉากกับแกน x เป็นตำแหน่งอ้างอิงที่ต้องการคำนวณโดยมีค่า k = 3 หรือเลือกหาค่าเฉลี่ยระยะห่างจากข้อมูล 3 จุด ณ ตำแหน่งที่ใกล้กับเส้นอ้างอิงมากที่สุด และจะคำนวณซ้ำโดยเปลี่ยนตำแหน่งอ้างอิง ผลลัพธ์ที่ได้จะเป็นลักษณะของเส้นกราฟต่อเนื่องกัน



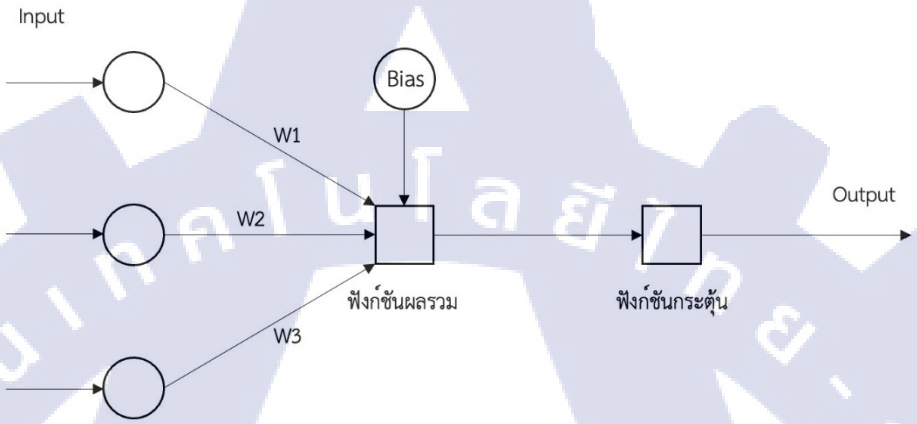
รูปที่ 2.9 การเลือกข้อมูลในการคำนวณการถดถอยของ k-Nearest Neighbors

2.8 โครงข่ายประสาทเทียม

การศึกษาในเรื่องของโครงข่ายประสาทเทียม (Artificial Neural Network) เริ่มต้นขึ้นในปี พ.ศ. 2486 เมื่อ W. S. McCulloch และ W. Pitts [51] ได้นำเสนอแบบจำลองทางคณิตศาสตร์ที่แสดงให้เห็นถึงแบบจำลองเซลล์ประสาทและพยายามทำความเข้าใจการทำงานของสมอง ร่วมกับกฎของ Hebb ซึ่งเสนอขึ้นในปี พ.ศ. 2492 ที่แสดงให้เห็นถึงกระบวนการเรียนรู้ที่เกิดขึ้นระหว่างเซลล์ที่เชื่อมต่อกัน [52] ถือเป็นจุดเริ่มต้นของการศึกษาในเรื่องนี้

จากที่ได้กล่าวไปข้างต้นจะเห็นได้ว่าโครงข่ายประสาทเทียมมีต้นแบบแนวคิดมาจากการทำงานของสมองมนุษย์ซึ่งมีส่วนย่อยคือเซลล์ประสาท (Neuron) เชื่อมต่อกันเป็นโครงข่าย บริเวณที่เชื่อมต่อกันเรียกว่าไซแนปส์ (Synapses) ซึ่งในปี พ.ศ. 2501 F. Rosenblatt [53] ได้นำเสนอถึงโครงข่ายประสาทเทียมแบบเพอร์เซปตรอน (Perceptron) ซึ่งถือเป็นรูปแบบพื้นฐานและง่ายที่สุดของโครงข่ายประสาทเทียม จากรูปที่ 2.10 จะแสดงให้เห็นถึงแบบจำลองโครงข่ายประสาทเทียมแบบเพอร์เซปตรอน ปมที่มีลักษณะวงกลมแสดงให้เห็นถึงเซลล์ประสาท แทนการเชื่อมต่อไซแนปส์ด้วย

เส้นเชื่อมซึ่งถูกกำกับไว้ด้วยค่าถ่วงน้ำหนัก (Weight) ในแต่ละเส้นมีฟังก์ชันผลรวมและค่าความเอนเอียง (Bias) สำหรับการเชื่อมต่อสื่อสารของเซลล์ประสาทจะเป็นลักษณะของการกระตุ้นด้วยศักย์ไฟฟ้าหรือที่เรียกว่าฟังก์ชันกระตุ้น (Activation Function) เมื่อการกระตุ้นถึงจุดที่กำหนดการส่งกระแสประสาทจะเกิดขึ้น [54] โดยการ เรียนรู้จะเป็นไปดังสมการที่ 2.3 และ 2.4



รูปที่ 2.10 แบบจำลองโครงข่ายประสาทเทียมแบบเพอร์เซปตรอน

$$\hat{y} = \sum_{i=1}^{n} x_i w_i + b \tag{2.3}$$

$$w_i \leftarrow w_i + \alpha(y - \hat{y})x_i \tag{2.4}$$

เมื่อ $\hat{y}$ แทนค่าที่ทำนายได้

x_i แทนข้อมูลนำเข้า

w_i แทนค่าถ่วงน้ำหนัก

b แทนค่าความเอนเอียง

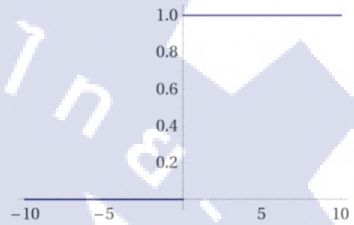
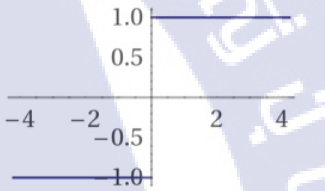
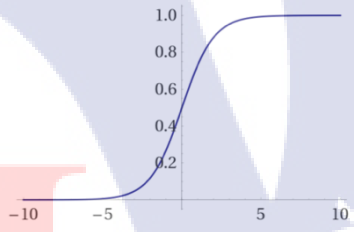
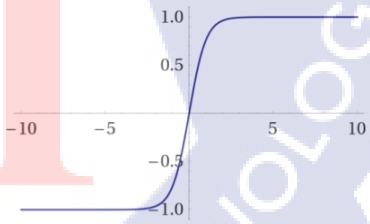
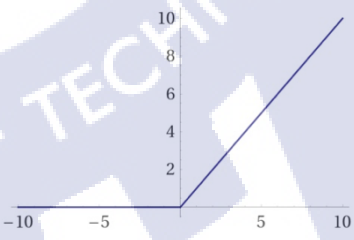
α แทนอัตราการเรียนรู้

y แทนค่าจริง

2.8.1 ฟังก์ชันกระตุ้น

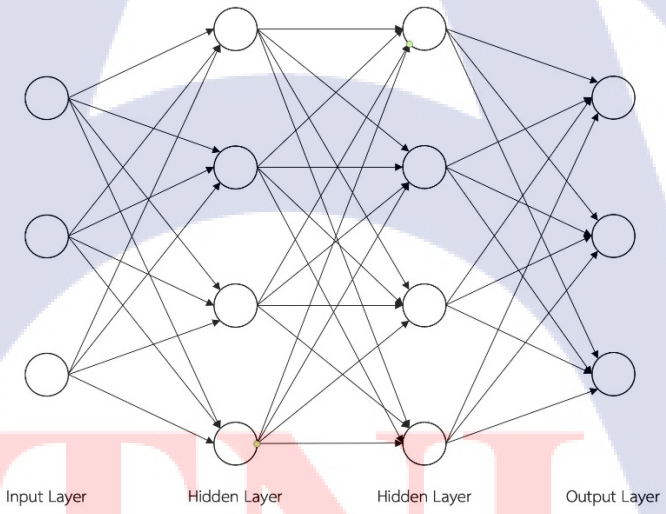
ฟังก์ชันกระตุ้นหรืออีกชื่อหนึ่งคือฟังก์ชันการถ่ายโอน (Transfer Function) ถือเป็นส่วนสำคัญของโครงข่ายประสาทเทียม เนื่องจากการเลือกฟังก์ชันในส่วนนี้จะส่งผลต่อข้อมูลส่งออกดังนั้น การเลือกฟังก์ชันกระตุ้นให้มีความเหมาะสมจะทำให้ผลลัพธ์ที่ได้มีความแม่นยำมากยิ่งขึ้น ตัวอย่างฟังก์ชันกระตุ้นที่พบแสดงดังตารางที่ 2.1 [55]

ตารางที่ 2.1 แสดงฟังก์ชันกระตุ้น สมการ และกราฟ

ฟังก์ชัน	สมการ	กราฟ
ฟังก์ชันขั้นบันได (Step Function)	$f(x) = \begin{cases} 1 \; if \; x \geq 0 \\ 0 \; if \; x < 0 \end{cases}$	
ฟังก์ชันขั้นบันไดแบบไบโพลาร์ (Bipolar Step Function)	$f(x) = \begin{cases} 1 \; if \; x > 0 \\ 0 \; if \; x = 0 \\ -1 \; if \; x < 0 \end{cases}$	
ฟังก์ชันซิกมอยด์ (Sigmoid Function)	$\sigma(x) = \frac{1}{1 + e^{-x}}$	
ฟังก์ชันแทนเจนต์ไฮเพอร์โบลิก (Hyperbolic Tangent Function: Tanh Function)	$tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$	
ฟังก์ชันเรคติไฟต์เชิงเส้น (Rectified Linear Unit Function: ReLu Function)	$f(x) = \begin{cases} 0 \; if \; x < 0 \\ x \; if \; x \geq 0 \end{cases}$	

2.9 การเรียนรู้เชิงลึก

จากที่ได้กล่าวมาข้างต้นจะสังเกตได้ว่าโครงข่ายประสาทเทียมแบบเพอร์เซปตรอนนั้นจะมีชั้นการคำนวณ (Computational Layer) เพียงชั้นเดียวเท่านั้น งานวิจัยของ M. Minsky และ S. Papert [56] ได้แสดงให้เห็นถึงข้อจำกัดในการแก้ปัญหาบางประเภทของโครงข่ายประสาทเทียมแบบนี้อยู่ จึงได้มีการพัฒนาโครงข่ายประสาทเทียมหลายชั้น (Multi-Layer Networks) ขึ้นมาในภายหลังเพื่อให้สามารถแก้ปัญหาที่โครงข่ายประสาทเทียมแบบเพอร์เซปตรอนไม่สามารถแก้ได้โดย โครงข่ายประสาทเทียมแบบหลายชั้นจะมีชั้นของการคำนวณเพิ่มขึ้นมา ซึ่งจะเรียกว่าชั้นซ่อน (Hidden Layer) การประมวลผลของสถาปัตยกรรมในรูปแบบนี้จะเป็นแบบป้อนไปข้างหน้า (Feed-Forward) กล่าวคือจะมีการทำงานในลักษณะของการรับข้อมูลเข้าที่ชั้นข้อมูลนำเข้า จากนั้นจึงส่งไปประมวลผลที่ชั้นซ่อนและส่งออกที่ชั้นเอาท์พุตเป็นลำดับ โดยลักษณะการเชื่อมต่อจะเป็นไปดังรูปที่ 2.11 [54] ซึ่งโครงสร้างที่กล่าวมานี้ถือเป็นต้นแบบของการเรียนรู้เชิงลึก (Deep Learning) นั้นเอง

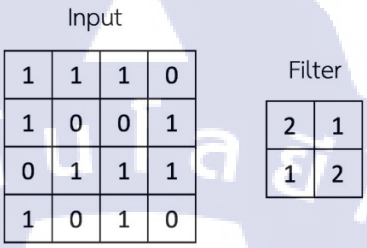


รูปที่ 2.11 โครงข่ายประสาทเทียมแบบหลายชั้น (Multi-Layer Networks)

2.9.1 <u>โครงข่ายประสาทเทียมคอนโวลูชัน</u>

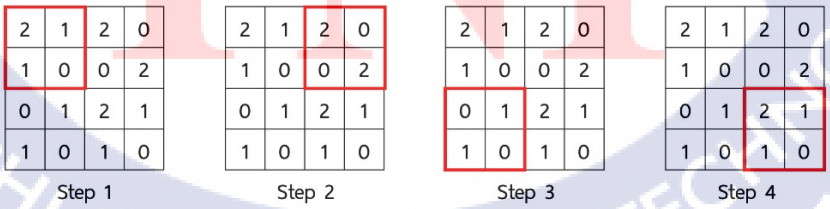
โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Networks) เป็นการเรียนรู้เชิงลึกประเภทหนึ่งที่มีแนวคิดหรือแรงบันดาลใจจากการแยกแยะคุณลักษณะและการรู้จำภาพ ดังนั้นโครงข่ายประสาทเทียมรูปแบบนี้จึงถูกนำมาใช้ในคอมพิวเตอร์วิทัศน์ (Computer Vision) สำหรับการจำแนกภาพและการตรวจจับวัตถุ (Object Detection) เนื่องจากมีการออกแบบมาสำหรับข้อมูล

นำเข้าที่มีโครงสร้างเป็นตาราง (Grid-Structured) ซึ่งรูปภาพแบบสองมิติก็เป็นตัวอย่างที่พบได้มากที่สุดของข้อมูลประเภทนี้นอกเหนือจากข้อมูลที่เป็นรูปภาพ โครงข่ายประสาทเทียมประเภทนี้ก็ยังสามารถใช้งานได้ดีกับข้อมูลที่เป็นข้อมูลตามลำดับ (Sequential Data) เช่นข้อความหรืออนุกรมเวลา (Time-Series) [54] โดยจะประกอบไปด้วยชั้นที่สำคัญ 3 ชั้นได้แก่ชั้นคอนโวลูชัน ชั้นพูลลิ่ง และชั้นเชื่อมโยงเต็มรูปแบบ

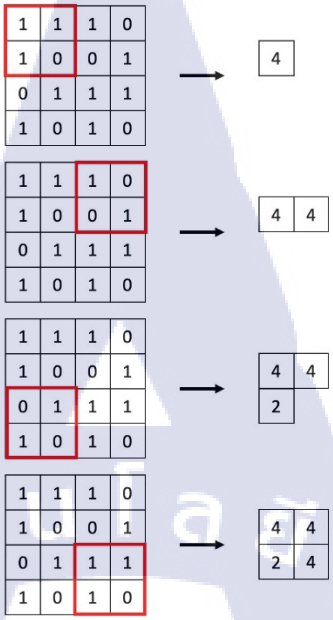


รูปที่ 2.12 ตัวอย่างข้อมูลนำเข้าและฟิลเตอร์

1. ชั้นคอนโวลูชัน (Convolutional Layer) จะมีข้อมูลนำเข้าและฟิลเตอร์ (Filter) หรือที่เรียกว่าเคอร์เนล (Kernel) ส่วนใหญ่จะมีลักษณะเป็นสี่เหลี่ยมจัตุรัสขนาดเล็กตัวอย่างแสดงดังรูปที่ 2.12 ซึ่งสามารถมีได้มากกว่าหนึ่งฟิลเตอร์ ในแต่ละตำแหน่งฟิลเตอร์จะมีการกำกับค่าน้ำหนักที่ไม่เท่ากัน ค่าน้ำหนักเริ่มต้นจะมาจากการสุ่มแล้วจึงปรับให้เหมาะสมจากการเรียนรู้ ในการคำนวณจะมีการสร้างฟีเจอร์แมพ (Feature Map) หรือการสกัดฟีเจอร์จากการคำนวณผลคูณแบบดอท (Dot Product) ระหว่างรูปภาพและฟิลเตอร์ซึ่งจะมีการเลื่อนฟิลเตอร์ไปจนกระทั่งครบทุกพื้นที่ของรูป โดยขนาดของการเลื่อนฟิลเตอร์ เรียกว่า การก้าวข้าม (Stride) ตัวอย่างการคำนวณคอนโวลูชันแสดงดังรูปที่ 2.13 และ 2.14

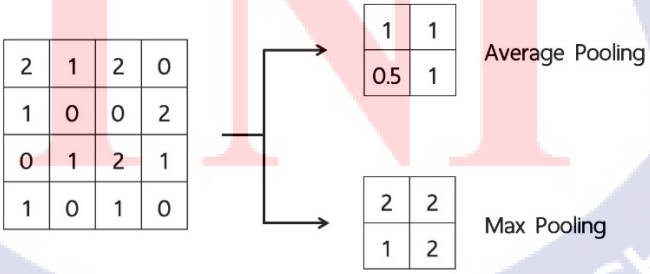


รูปที่ 2.13 การเปลี่ยนตำแหน่งของฟิลเตอร์เมื่อการก้าวข้ามเท่ากับสอง



รูปที่ 2.14 ตัวอย่างการคำนวณคอนโวลูชัน

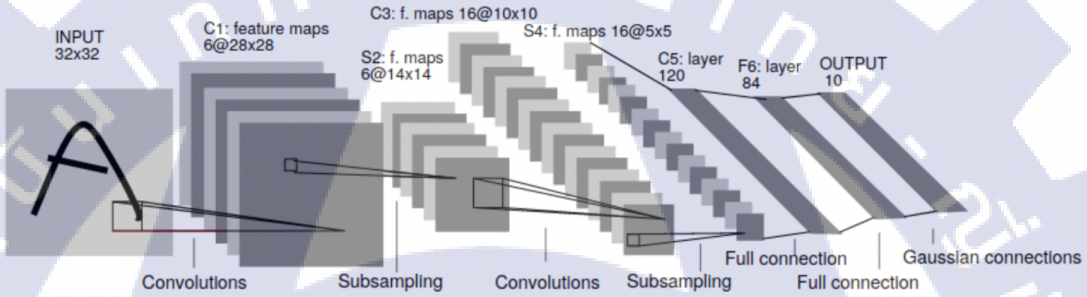
2. ชั้นพูลลิ่ง (Pooling Layer) เป็นชั้นที่จะทำการลดขนาดของฟีเจอร์แมพหรือฟีเจอร์ที่ทำการสกัดมา เพื่อให้สามารถทำการคำนวณหรือประมวลผลต่อได้อย่างสะดวกและมีความรวดเร็วมากยิ่งขึ้น โดยการทำพูลลิ่งจะเป็นการรวมฟีเจอร์ที่คล้ายกันไว้ด้วยกัน นิยมอยู่สองแบบได้แก่ พูลลิ่งแบบหาค่าเฉลี่ย (Average Pooling) จะทำการหาค่าเฉลี่ยของค่าน้ำหนักทั้งหมดในพื้นที่ย่อยที่กำหนด และพูลลิ่งแบบหาค่ามากสุด (Max Pooling) จะทำการเลือกค่าน้ำหนักที่มากที่สุดในพื้นที่ย่อยที่กำหนดดังรูปที่ 2.15



รูปที่ 2.15 การทำพูลลิ่งแบบค่าเฉลี่ยและค่ามากสุด (การก้าวข้ามเท่ากับสอง)

3. ชั้นเชื่อมโยงเต็มรูปแบบ (Fully-Connected Layer) เป็นชั้นที่จะทำการเชื่อมต่อระหว่างกระบวนการสกัดฟีเจอร์และการประมวลผลด้วยโครงข่ายประสาทเทียมเข้าด้วยกัน ซึ่งชั้นนี้จะทำการเปลี่ยนรูป (Reshape) เพื่อเตรียมสำหรับการคำนวณ ดังนั้นชั้นนี้จึงเปรียบเสมือนเป็นชั้นเตรียมส่งข้อมูลนำเข้าเพื่อนำไปประมวลผลแบบโครงข่ายประสาทเทียมแล้วจึงออกมาเป็นผลลัพธ์

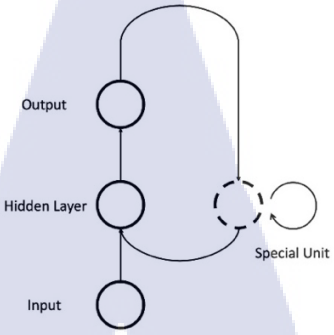
LeNet-5 เป็นสถาปัตยกรรมพื้นฐานชิ้นแรกที่มีลักษณะแบบโครงข่ายประสาทเทียมคอนโวลูชัน ถูกนำเสนอขึ้นโดย Y. Lecun และคณะ [57] ประกอบไปด้วยชั้นนำเข้า ชั้นคอนโวลูชันจำนวนสามชั้น ชั้นพูลลิ่งจำนวนสองชั้น ชั้นการเชื่อมโยงเต็มรูปแบบ และชั้นแสดงผลลัพธ์ รายละเอียดแสดงดังรูปที่ 2.16



รูปที่ 2.16 สถาปัตยกรรม LeNet-5 [57]

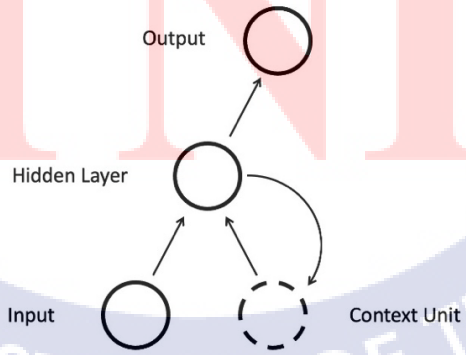
2.9.2 โครงข่ายประสาทเทียมวกกลับ

สถาปัตยกรรมของโครงข่ายประสาทเทียมอีกรูปแบบหนึ่งที่ได้รับความนิยมสำหรับชุดข้อมูลแบบอนุกรมเวลาได้แก่ โครงข่ายประสาทเทียมวกกลับ (Recurrent Neural Network) จุดเด่นของโครงข่ายประสาทเทียมรูปแบบนี้คือส่วนของการจดจำ (Memory) ซึ่งจะทำการเก็บข้อมูลของส่วนที่ได้คำนวณไปก่อนหน้า จากนั้นจึงส่งต่อไปยังขั้นตอนถัดไปเสมือนเป็นข้อมูลนำเข้าอีกข้อมูลหนึ่ง หากเปรียบเทียบกับโครงข่ายประสาทเทียมแบบดั้งเดิมจะพบว่าข้อมูลนำเข้าและเอาท์พุตทั้งหมดจะถูกแยกออกจากกันอย่างชัดเจนหรือเป็นอิสระต่อกัน จะมีเพียงค่าน้ำหนักเท่านั้นที่ถูกปรับสำหรับการเรียนรู้ในขั้นตอนถัดไป จุดนี้จึงถือเป็นส่วนสำคัญที่ทำให้โครงข่ายประสาทเทียมวกกลับนั้นมีความเหมาะสมที่จะใช้กับข้อมูลที่มีลักษณะของความต่อเนื่อง อาทิเช่น การประมวลผลภาษาธรรมชาติ (Natural Language Processing) รหัสพันธุกรรม (DNA) และข้อมูลอนุกรมเวลา (Times-Series) เป็นต้น



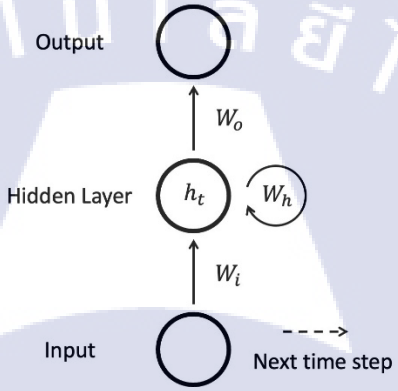
รูปที่ 2.17 แบบจำลองสถาปัตยกรรมของ Jordan

ในปี พ.ศ. 2525 แนวคิดเกี่ยวกับการจดจำของเซลล์ประสาทและการคำนวณที่เกี่ยวข้องได้ถูกเสนอขึ้นโดย J. Hopfield [58] ส่วนนี้เองถือเป็นจุดเริ่มต้นของการพัฒนาโครงข่ายประสาทเทียมวกกลับสำหรับสถาปัตยกรรมการเรียนรู้ของเครื่องแบบมีผู้สอนสำหรับข้อมูลแบบเป็นลำดับได้ถูกเสนอขึ้นในอีกสี่ปีถัดมาโดย M. I. Jordan [59] สถาปัตยกรรมนี้มีการเพิ่มโหนดพิเศษ (Special Unit) ขึ้นมานอกเหนือจากโหนดของข้อมูลนำเข้า โหนดในชั้นซ่อน และโหนดของเอาท์พุตดังรูปที่ 2.17 ซึ่งส่วนของโหนดพิเศษนี้จะทำหน้าที่ในการเก็บข้อมูลการคำนวณที่ได้จากขั้นตอนของเวลาในช่วงก่อนหน้าและส่งข้อมูลกลับเข้าโหนดดังกล่าวรวมถึงส่งไปยังโหนดในชั้นซ่อนสำหรับช่วงเวลาถัดไปด้วยเช่นกัน อีกหนึ่งสถาปัตยกรรมของโครงข่ายประสาทเทียมวกกลับที่มีความน่าสนใจได้แก่ สถาปัตยกรรมของ J. L. Elman [60] ดังรูปที่ 2.18 ปรากฏขึ้นในปี พ.ศ. 2533 เป็นสถาปัตยกรรมที่มีความคลายคลึงกับสถาปัตยกรรมของ Jordan แต่สามารถเข้าใจได้ง่ายและไม่มีความซับซ้อนมาก เนื่องจากมีการกำหนดให้ค่าน้ำหนักที่ถูกส่งไปยัง Context Unit มีค่าคงที่เท่ากับ 1

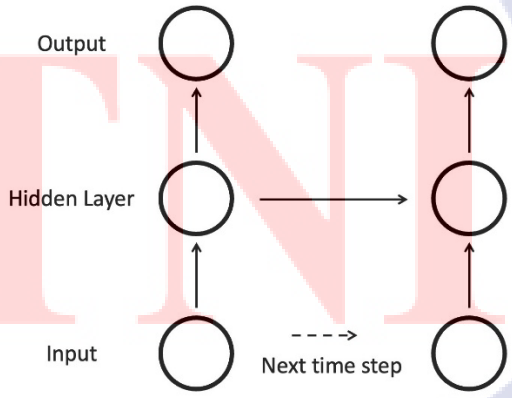


รูปที่ 2.18 แบบจำลองสถาปัตยกรรมของ Elman

เมื่อพิจารณาถึงสถาปัตยกรรมของ Elman จะพบว่าสถาปัตยกรรมนี้เป็นต้นแบบของโมเดลที่เรียกกันอย่างคุ้นเคยในปัจจุบันว่าโครงข่ายประสาทเทียมวกกลับแบบ Simple จากรูปที่ 2.19 จะแสดงให้เห็นถึงโครงสร้างของโครงข่ายประสาทเทียมวกกลับแบบ Simple ในหนึ่งช่วงเวลาที่ประกอบไปด้วยโหนดของข้อมูลนำเข้า ชั้นซ่อน และชั้นเอาท์พุตเพียงอย่างละหนึ่งเท่านั้น จะเห็นได้ว่าที่ชั้นซ่อนจะมีการส่งข้อมูลกลับเข้าโหนดอีกครั้ง หากพิจารณาจากรูปที่ 2.20 จะพบว่าการส่งข้อมูลกลับเข้าโหนดในชั้นซ่อนนั้นหมายความได้ว่า ชั้นซ่อนของช่วงเวลาหนึ่งได้ส่งผ่านข้อมูลไปยังโหนดในชั้นซ่อนของช่วงเวลาถัดไป

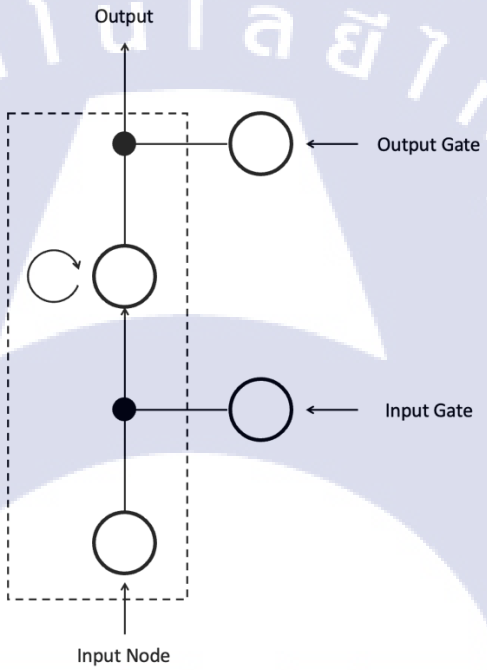


รูปที่ 2.19 โครงข่ายประสาทเทียมวกกลับแบบ Simple ในหนึ่งช่วงเวลา



รูปที่ 2.20 แบบจำลองการส่งข้อมูลของชั้นซ่อนไปยังช่วงเวลาถัดไป

ในอีกโมเดลหนึ่งซึ่งมีความคล้ายคลึงกับโครงข่ายประสาทเทียมวกกลับพื้นฐานและมีความน่าสนใจเป็นอย่างยิ่งนั้นคือโมเดลในรูปแบบ Long Short-Term Memory (LSTM) ความแตกต่างที่ทำให้โมเดลนี้มีความน่าสนใจคือ ในแต่ละโหนดของชั้นซ่อนจะถูกแทนที่ด้วยโหนดความจำ (Memory Cell) ที่มีลักษณะดังรูปที่ 2.21 การผสมผสานระหว่างคุณสมบัติแบบ Long Term Memory ซึ่งเป็นคุณสมบัติหลักที่พบได้ในโมเดลโครงข่ายประสาทเทียมวกกลับแบบ Simple และ Short term memory ส่งผลให้โมเดลนี้สามารถแก้ปัญหา Gradient ที่พบในโมเดลโครงข่ายประสาทเทียมวกกลับแบบ Simple ได้



รูปที่ 2.21 โหนดความจำ (Memory Cell) ของโครงข่ายประสาทเทียมวกกลับแบบ LSTM

กระบวนการภายในโหนดความจำของโครงข่ายประสาทเทียมวกกลับแบบ LSTM จะเริ่มต้นจากส่วนที่ใช้รับข้อมูลเข้าได้แก่ Input Node และ Input Gate จากนั้นจะเข้าไปสู่ Internal State ซึ่งจะมีลักษณะเป็นการเชื่อมต่อวกกลับภายในตัว (Self-Recurrent Connection) ที่มีการปรับปรุงค่าในแต่ละช่วงเวลา จากนั้นจึงจะถูกส่งออกไปโดยค่าของเอาท์พุตจะเป็นค่าที่ได้จาก Internal State และ Output Gate

2.10 การบำรุงรักษาเชิงคาดการณ์

งานด้านการบำรุงรักษา (Maintenance) ถือเป็นงานที่มีความสำคัญต่ออุตสาหกรรมเป็นอย่างมาก เนื่องจากในอุตสาหกรรมมีการใช้เครื่องจักรเป็นส่วนหนึ่งของการผลิต ดังนั้นการที่จะบำรุงรักษาให้เครื่องจักรอยู่ในสภาพที่สามารถใช้งานได้เสมอจึงมีความจำเป็น เห็นได้ว่ามนุษย์รู้จักการบำรุงรักษามาตั้งแต่เริ่มมีการใช้เครื่องจักรมาเป็นส่วนหนึ่งของอุตสาหกรรมและมีการพัฒนาเทคนิคในการบำรุงรักษาอย่างต่อเนื่อง วิธีที่หลากหลายถูกนำมาใช้เป็นส่วนหนึ่ง จึงได้มีการจัดกลุ่มแนวทางการซ่อมบำรุงไว้ทั้งหมดสามแนวทางหลัก ได้แก่การบำรุงรักษาเชิงรับ การบำรุงรักษาเชิงป้องกัน และการบำรุงรักษาเชิงคาดการณ์ [61]

1. การบำรุงรักษาเชิงรับ (Corrective Maintenance) หรือที่คุ้นเคยกันว่าซ่อมก็ต่อเมื่อเสีย (Breakdown Maintenance) เป็นการบำรุงรักษาแบบแรกที่เกิดขึ้นในโรงงานหรืออุตสาหกรรม ซึ่งหมายถึงจะดำเนินการซ่อมก็ต่อเมื่อพังหรือเสียหายเท่านั้น ถึงแม้จะไม่ต้องดำเนินการบำรุงรักษาก่อนที่เครื่องจักรจะเสียหาย แต่กลับเป็นแนวทางที่มีค่าใช้จ่ายสูงซึ่งจะเป็นค่าใช้จ่ายทางอ้อมและค่าอะไหล่เสียส่วนใหญ

2. การบำรุงรักษาเชิงป้องกัน (Preventive Maintenance: PvM) เป็นการบำรุงรักษาก่อนที่จะเกิดความเสียหายขึ้นกับเครื่องจักร โดยส่วนใหญ่จะอาศัยเวลาเป็นตัวแปรสำคัญ ซึ่งจะมีการประมาณอายุการใช้งานหรือโอกาสที่จะเกิดความเสียหายขึ้นในแต่ละช่วงเวลาของเครื่องจักรเพื่อดำเนินการวางแผนการซ่อมบำรุง

3. การบำรุงรักษาเชิงคาดการณ์ (Predictive Maintenance: PdM) เป็นการตรวจสอบสภาพหรือตรวจจับสัญญาณที่สามารถบ่งชี้ได้ว่าอาจเกิดความเสียหายขึ้นกับเครื่องจักรเพื่อทำการซ่อมบำรุงก่อนเกิดความเสียหาย มีความแตกต่างจากการซ่อมบำรุงเชิงป้องกันในส่วนที่จะไม่มีการกำหนดเวลาที่แน่นอน จึงกล่าวได้ว่าเป็นการบำรุงรักษาที่อ้างอิงจากสภาพจริงของเครื่องจักร รวมถึงสภาพแวดล้อมในการทำงานของเครื่องจักรร่วมด้วย

เพื่อหลีกเลี่ยงการบำรุงรักษาเชิงรับที่ไม่ได้คาดการณ์ไว้และการบำรุงรักษาเชิงป้องกันที่อาจก่อให้เกิดค่าใช้จ่ายที่ไม่จำเป็นจำนวนมาก การบำรุงรักษาเชิงคาดการณ์จึงเป็นอีกหนึ่งประเภทของการบำรุงรักษาที่กำลังได้รับความสนใจเป็นอย่างมาก [36] ด้วยข้อมูลอ้างอิงตามเวลา (Time-Based Information) ที่ได้จากอุปกรณ์ตรวจวัดต่าง ๆ เช่น เซนเซอร์และความรู้ที่เกี่ยวข้องถูกนำมาใช้ร่วมกับการวิเคราะห์ขั้นสูงในการพิจารณาถึงรูปแบบรวมถึงแนวโน้มเพื่อพยากรณ์ความเสียหายที่อาจเกิดขึ้น [6, 36] เมื่อจุดประสงค์หลักของการการบำรุงรักษาเชิงคาดการณ์คือการคาดการณ์ความเสียหายและ

ดำเนินการบำรุงรักษาเมื่อจำเป็นเพื่อป้องกันไม่ให้เกิดความเสียหาย [3] สิ่งสำคัญที่จะช่วยให้ประสบความสำเร็จคือปริมาณข้อมูลที่เพียงพอและมีความครอบคลุมในทุกส่วนของกระบวนการ [6] ด้วยสิ่งที่กล่าวมานี้การที่เครื่องจักรหยุดชะงักหรือความสิ้นเปลืองจากการเปลี่ยนอุปกรณ์ที่ไม่จำเป็นนั้นแทบจะไม่เกิดขึ้น อีกทั้งยังสามารถปรับปรุงความปลอดภัยในการทำงาน ความพร้อมใช้งานของอุปกรณ์หรือเครื่องจักร ลดต้นทุนการบำรุงรักษา และเพิ่มประสิทธิภาพในกระบวนการผลิต [6, 36, 62] ในงานการบำรุงรักษาเชิงคาดการณ์ ได้มีการจัดจำแนกแนวทางหลักที่ใช้ไว้ทั้งหมดสามแนวทาง ได้แก่ แนวทางอ้างอิงโมเดลทางกายภาพ แนวทางอ้างอิงความรู้และ แนวทางขับเคลื่อนด้วยข้อมูล [6]

1. แนวทางอ้างอิงโมเดลทางกายภาพ (Physical Model-Based) เป็นการใช้วิธีการทางสถิติและโมเดลทางคณิตศาสตร์ในการสร้างแบบจำลองการเสื่อมสภาพและอธิบายลักษณะทางกายภาพของความเสียหายที่เกิดขึ้น วิธีนี้จำเป็นต้องอาศัยความรู้ความเข้าใจเกี่ยวกับความเสียหายที่เกิดขึ้นโดยธรรมชาติร่วมด้วย [37]

2. แนวทางอ้างอิงความรู้(Knowledge-Based) ช่วยลดความซับซ้อนของโมเดลทางกายภาพด้วยการใช้ความรู้เข้ามาเพิ่มความเข้าใจในลักษณะทางกายภาพของอุปกรณ์หรือเครื่องจักร จึงนิยมใช้แนวทางนี้ร่วมกับแนวทางอ้างอิงโมเดลทางกายภาพเป็นแบบผสมผสาน (Hybrid) ตัวอย่างเช่น ระบบผู้เชี่ยวชาญ (Expert System)

3. แนวทางขับเคลื่อนด้วยข้อมูล (Data-Driven) มีการใช้ข้อมูลจำนวนมากเพื่อค้นหาความเชื่อมโยงหรือความเกี่ยวข้องของการเสื่อมสภาพและความเสียหายที่อาจเกิดขึ้นผ่านวิธีการต่าง ๆ เช่น วิธีทางสถิติปัญญาประดิษฐ์ (Artificial Intelligence: AI) หรือโมเดลที่อ้างอิงจากอัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning Algorithms) เป็นต้น [37]

ในปัจจุบันแนวทางขับเคลื่อนด้วยข้อมูลกำลังเป็นที่นิยมมากเมื่อเปรียบเทียบกับการใช้แบบจำลอง เนื่องจากอุตสาหกรรม 4.0 และข้อมูลขนาดใหญ่เข้ามามีบทบาทสำคัญกับโลกอุตสาหกรรม จึงมีการนำเทคโนโลยีที่เกี่ยวข้องตัวอย่างเช่น การนำเซนเซอร์มาใช้ในการเก็บค่าพารามิเตอร์ (Parameters) จากเครื่องจักรเป็นจำนวนมาก ด้วยข้อมูลในส่วนนี้จะสามารถออกแบบระบบที่ใช้แนวทางนี้ได้ง่ายยิ่งขึ้น [37] สำหรับเทคนิคที่ใช้ในการรวบรวมข้อมูลเพื่อนำมาวิเคราะห์การบำรุงรักษาเชิงคาดการณ์มีอยู่สามเทคนิคด้วยกันเมื่อจำแนกตามแหล่งที่มา ได้แก่ เซนเซอร์ที่มีอยู่ในกระบวนการ เซนเซอร์ทดสอบ และสัญญาณทดสอบ [62]

1. เซนเซอร์ที่มีอยู่ในกระบวนการ (Process Sensors) เป็นอุปกรณ์วัดที่ถูกติดตั้งอยู่ในระบบมาแต่เดิม อาจมีวัตถุประสงค์เพื่อการวัดค่าทั่วไปด้านอุตสาหกรรม เช่น อุณหภูมิความร้อน

ความดัน หรืออัตราการไหล ซึ่งค่าเหล่านี้สามารถนำมาเป็นข้อมูลเพื่อทำการวิเคราะห์การบำรุงรักษาได้

2. เซนเซอร์ทดสอบ (Test Sensors) เป็นอุปกรณ์ที่ถูกติดตั้งเพิ่มเติมบนชิ้นส่วนเครื่องจักรเพื่อวัดค่าพารามิเตอร์บางอย่างเช่น ขนาดการสั่นสะเทือน (Vibration Amplitude) ที่สามารถบ่งบอกความเสียหายที่เกิดขึ้นกับชิ้นส่วนเครื่องจักรได้หรืออาจหมายถึงการวัดค่าที่ไม่เกี่ยวข้องกับเครื่องจักรแต่คาดว่าจะมีผลกระทบต่อการบำรุงรักษาอย่างปัจจัยแวดล้อม เช่น ความชื้น เป็นต้น

3. สัญญาณทดสอบ (Test Signal) เป็นการนำสัญญาณทดสอบฉีดหรือใส่เข้าไปในชิ้นส่วนที่ต้องการทดสอบเช่น การฉีดสัญญาณทดสอบเข้าไปในสายเคเบิลเพื่อค้นหาตำแหน่งความเสียหายของสายเคเบิล เทคนิคนี้สามารถทดสอบการแตก การกัดกร่อน หรือการสึกหรอรวมถึงสามารถวัดประสิทธิภาพของชิ้นส่วนได้ด้วยเช่นกัน

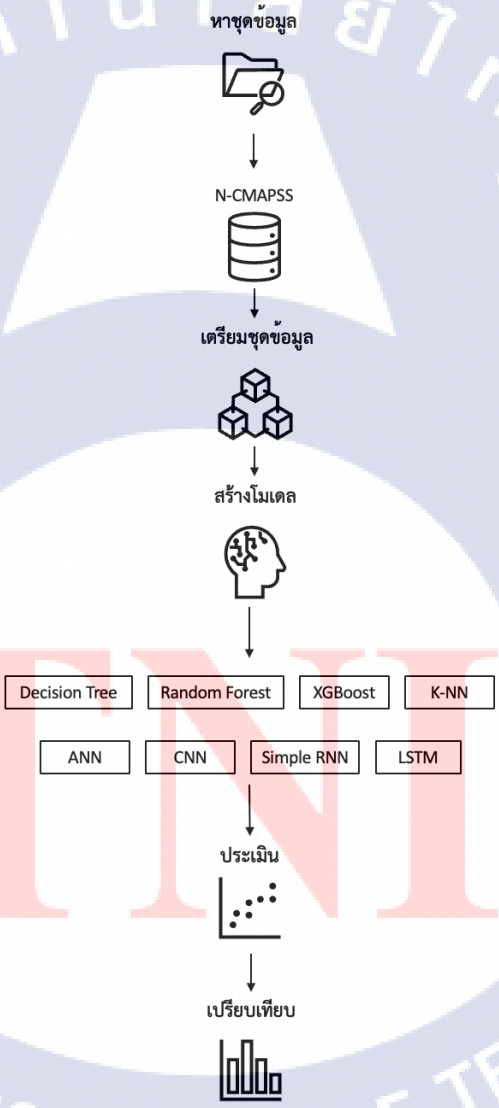
ความนิยมที่เพิ่มสูงขึ้นในแนวทางขับเคลื่อนด้วยข้อมูล (Data-Driven Approach) จากจุดเด่นที่สามารถปรับแต่งอัลกอริทึมเพียงเล็กน้อยเพื่อประยุกต์กับระบบอื่นได้และต้องการความรู้ความเข้าใจพื้นฐานทางด้านกายภาพที่ต่ำ [63] ร่วมกับเทคโนโลยีด้านอุตสาหกรรมที่พัฒนาอย่างก้าวกระโดด ส่งผลให้ข้อมูลในทุกรูปแบบที่มีอย่างมหาศาลถูกนำมาใช้ให้เกิดประโยชน์สูงสุดและมีความพยายามที่จะนำเสนอหลักการหรือวิธีการที่ใช้สำหรับคาดการณ์ความเสียหายของเครื่องจักรที่หลากหลายมากขึ้นจากสิ่งเหล่านี้ หนึ่งในวิธีที่ได้รับความนิยมอย่างสูงมากคือการคาดการณ์อายุการใช้งานคงเหลือ (Remaining Useful Life: RUL) ซึ่งหมายถึงระยะเวลาที่อุปกรณ์จะสามารถให้ประโยชน์หรือใช้งานได้นับตั้งแต่ปัจจุบันจนกระทั่งสิ้นสุดอายุการใช้งาน วิธีการนี้จะทำการประมาณอายุการใช้งานคงเหลือของชิ้นส่วนหรืออุปกรณ์นั้นที่ช่วงเวลาหนึ่งของการดำเนินงานและจะสามารถคาดการณ์ช่วงเวลาที่จะเกิดความเสียหายกับชิ้นส่วนหรืออุปกรณ์นั้นได้ ถือได้ว่ามีความซับซ้อนสูงแต่ก็เป็นประโยชน์มากเช่นกัน [6, 64, 65]

อัลกอริทึมที่หลากหลายถูกนำมาใช้เพื่อคาดการณ์อายุการใช้งานคงเหลือ อาทิเช่น ตัวเข้ารหัสอัตโนมัติ (Auto-Encoder) โครงข่ายประสาทเทียมความเชื่อเชิงลึก (Deep Belief Network: DBN) โครงข่ายประสาทคอนโวลูชัน (Convolutional Neural Network: CNN) หรือโครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network: RNN) ซึ่งถูกจัดอยู่ในกลุ่มของการเรียนรู้เชิงลึก (Deep Leaning) [63] รวมถึงการถดถอยเชิงเส้น (Linear Regression) ต้นไม้ตัดสินใจ (Decision Tree) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) ป่าแบบสุ่ม (Random Forest) เทคนิคเพื่อนบ้านใกล้สุด (K-Nearest Neighbors) และอื่น ๆ ที่ถูกจัดอยู่ในกลุ่ม

ของการเรียนรู้ของเครื่องเช่นกัน [66] และถึงแม้ว่าการใช้แนวทางขับเคลื่อนด้วยข้อมูลจะมีประโยชน์และข้อได้เปรียบอยู่มากแต่ก็พบว่ายังคงมีข้อจำกัดอยู่ เช่น ต้องการข้อมูลที่มากเพียงพอในการฝึกสอนโมเดล ความแม่นยำของโมเดลในการคาดการณ์ซึ่งอาจเกิด Overfitting ขึ้นได้หรือแม้กระทั่งความไม่เข้าใจในผลลัทธ์ที่ได้เนื่องจากขัดกับสัญชาตญาณมนุษย์ [63]

บทที่ 3
วิธีการดำเนินงานวิจัย

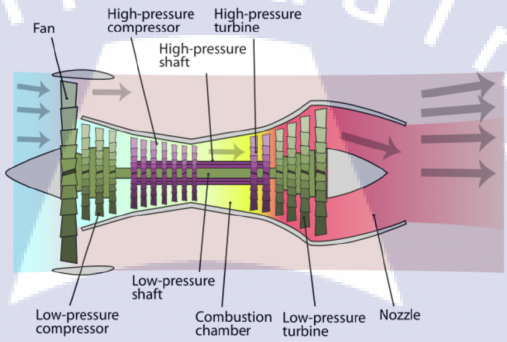
ในส่วนของบทนี้จะกล่าวเกี่ยวกับวิธีการดำเนินงานวิจัยซึ่งจะแสดงให้เห็นถึงรายละเอียดของขั้นตอนการทำวิจัยทั้งหมด 5 ขั้นตอนได้แก่ การจัดหาชุดข้อมูล การเตรียมชุดข้อมูล การสร้างโมเดล การประเมินโมเดล และการเปรียบเทียบประสิทธิภาพ เพื่อให้เห็นภาพได้อย่างชัดเจนยิ่งขึ้น ขั้นตอนทั้งหมดได้แสดงไว้ดังรูปที่ 3.1



รูปที่ 3.1 ขั้นตอนการดำเนินงาน

3.1 ชุดข้อมูล

ชุดข้อมูลที่นำมาใช้ในงานวิจัยฉบับนี้เป็นชุดข้อมูลที่เก็บรวบรวมค่าพารามิเตอร์ต่าง ๆ ของเครื่องยนต์อากาศยานตั้งแต่เริ่มต้นจนกระทั่งเกิดความเสียหายจากระบบฐานข้อมูล Standard Datasets ของ NASA [67] ถูกสร้างขึ้นจากโปรแกรม Commercial Modular Aero-Propulsion System Simulation หรือ C-MAPSS ซึ่งเป็นโปรแกรมสำหรับจำลองข้อมูลเครื่องยนต์อากาศยานเชิงพาณิชย์ขนาดใหญ่ ที่มีลักษณะเครื่องยนต์แบบ Turbofan ดังรูปที่ 3.2 งานวิจัยฉบับนี้จะใช้ข้อมูลจากโมเดลที่ปรับปรุงใหม่ในชื่อ N-CMAPSS มาทำการเปรียบเทียบการคาดการณ์ความเสียหายของเครื่องยนต์อากาศยานด้วยโมเดลการเรียนรู้ของเครื่องแบบมีผู้สอน



รูปที่ 3.2 เครื่องยนต์แบบ Turbofan [68]

C-MAPSS เป็นโปรแกรมจำลองที่มีความสมจริง โดยชุดข้อมูลที่ได้จากโมเดลปรับปรุงใหม่ (N-CMAPSS) จะถูกจำลองขึ้นภายใต้ข้อกำหนดทางสภาพแวดล้อมเดียวกันกับโมเดลเดิมซึ่งมีรายละเอียดดังต่อไปนี้

(1) ระดับความสูง (Altitudes) ตั้งแต่ระดับน้ำทะเลไปถึงที่ระดับความสูง 40,000 ฟุต

(2) เลขมัค (Mach Numbers) ตั้งแต่ 0 ถึง 0.90

(3) อุณหภูมิที่ระดับน้ำทะเล (Sea-Level Temperatures) ตั้งแต่ -60 ถึง 103 ฟาเรนไฮต์

โมเดลที่ปรับปรุงใหม่จะมีลักษณะของสมการแบบไม่เป็นเชิงเส้น (Non-Linear) โดยข้อมูลนำเข้าจะถูกแบ่งออกเป็นสองส่วนได้แก่ ส่วนของสถานการณ์หรือข้อบ่งชี้และพารามิเตอร์แสดงสถานะการเสื่อมสภาพ ในส่วนของเอาท์พุตจะถูกแบ่งออกเป็นสองส่วนเช่นกันได้แก่ คุณสมบัติทางกายภาพที่ได้จากการวัดและคุณสมบัติจากเซนเซอร์เสมือนจริง รายละเอียดแสดงดังตารางที่ 3.1 ถึงตารางที่ 3.4

ตารางที่ 3.1 สถานการณ์หรือข้อบ่งชี้

สัญลักษณ์	คำอธิบาย	หน่วย
alt	ระดับความสูง	ft
Mach	เลขมัคของเที่ยวบิน	-
TRA	Throttle-resolver angle	%
T2	อุณหภูมิรวมที่ทางเข้าใบพัด	°R

ตารางที่ 3.2 คุณสมบัติทางกายภาพจากการวัด

สัญลักษณ์	คำอธิบาย	หน่วย
Wf	อัตราการไหลของเชื้อเพลิง	pps
Nf	ความเร็วใบพัด	rpm
T24	ความเร็วแกน	rpm
T30	อุณหภูมิรวมที่ทางออก LPC	°R
T48	อุณหภูมิรวมที่ทางออก HPC	°R
T50	อุณหภูมิรวมที่ทางออก HPT	°R
P15	อุณหภูมิรวมที่ทางออก LPT	°R
P2	ความดันรวมในท่อ Bypass	psia
P21	ความดันรวมที่ทางออกใบพัด	psia
P24	ความดันรวมที่ทางออก LPC	psia
Ps30	ความดันสถิตที่ทางออก HPC	psia
P40	ความดันรวมที่ทางออก Burner	psia
P50	ความดันรวมที่ทางออก LPT	psia

ตารางที่ 3.3 คุณสมบัติจากเซนเซอร์เสมือนจริง

สัญลักษณ์	คำอธิบาย	หน่วย
T40	อุณหภูมิรวมที่ทางออก Burner	°R
P30	ความดันรวมที่ทางออก HPC	psia
P45	ความดันรวมที่ทางออก HPT	psia
W21	Fan Flow	pps
W22	อัตราไหลออกของ LPC	lbm/s

ตารางที่ 3.3 คุณสมบัติจากเซนเซอร์เสมือนจริง (ต่อ)

สัญลักษณ์	คำอธิบาย	หน่วย
W25	อัตราไหลเข้า HPC	lbm/s
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed	lbm/s
W48	อัตราไหลออกของ HPT	lbm/s
W50	อัตราไหลออกของ LPT	lbm/s
SmFan	Fan stall margin	-
SmLPC	LPC stall margin	-
SmHPC	HPC stall margin	-
phi	อัตราส่วนอัตราการไหลเชื้อเพลิง Ps30	pps/psi

ตารางที่ 3.4 พารามิเตอร์แสดงสถานะการเสื่อมสภาพ

สัญลักษณ์	คำอธิบาย	หน่วย
fan_eff_mod	Fan efficiency modifier	-
fan_flow_mod	Fan flow modifier	-
LPC_eff_mod	LPC efficiency modifier	-
LPC_flow_mod	LPC flow modifier	-
HPC_eff_mod	HPC efficiency modifier	-
HPC_flow_mod	HPC flow modifier	-
HPT_eff_mod	HPT efficiency modifier	-
HPT_flow_mod	HPT flow modifier	-
LPT_eff_mod	LPT efficiency modifier	-
LPT_flow_mod	LPT flow modifier	-

ชุดข้อมูลที่ได้ทำการดาวน์โหลดมาจากเว็บไซต์ [67] เป็นไฟล์ประเภท Hierarchical Data Format 5 (.h5) ทั้งหมด 9 ไฟล์ ซึ่งไฟล์ประเภทนี้ถูกใช้สำหรับเก็บข้อมูลขนาดใหญ่โดยเฉพาะ ในแต่ละไฟล์จะมีข้อมูลของพารามิเตอร์แสดงสถานะการเสื่อมสภาพแตกต่างกันออกไปตามที่แสดงในตารางที่ 3.5

ตารางที่ 3.5 ชื่อไฟล์และข้อมูลพารามิเตอร์แสดงสถานะการเสื่อมสภาพของแต่ละไฟล์

ชื่อไฟล์	พารามิเตอร์ที่เก็บ
N-CMAPSS_DS01-005.h5	HPT_eff_mod
N-CMAPSS_DS02-006.h5	HPT_eff_mod LPT_eff_mod LPT_flow_mod
N-CMAPSS_DS03-012.h5	HPT_eff_mod LPT_eff_mod LPT_flow_mod
N-CMAPSS_DS04.h5	fan_eff_mod fan_flow_mod
N-CMAPSS_DS05.h5	HPC_eff_mod HPC_flow_mod
N-CMAPSS_DS06.h5	LPC_eff_mod LPC_flow_mod HPC_eff_mod HPC_flow_mod
N-CMAPSS_DS07.h5	LPT_eff_mod LPT_flow_mod
N-CMAPSS_DS08a-009.h5	เก็บทุกพารามิเตอร์
N-CMAPSS_DS08c-008.h5	เก็บทุกพารามิเตอร์

นอกเหนือจากข้อมูลในตารางที่ได้กล่าวไปในตารางที่ 3.1 ถึง 3.4 ชุดข้อมูลเกี่ยวกับอายุการใช้งานคงเหลือ (Remaining Useful Life: RUL) ก็ได้ถูกเก็บไว้ในแต่ละไฟล์ด้วยเช่นกัน และถึงแม้ว่าข้อมูลที่ได้จากโปรแกรมจะมีทั้งหมดสี่ส่วนดังที่อธิบายไว้ข้างต้น แต่คุณสมบัติจากเซนเซอร์เสมือนจริงนั้นไม่ได้มีความเกี่ยวข้องกับการคาดการณ์ความเสียหายตามที่ได้อธิบายไว้ในงานวิจัยของ A. C. Manuel และคณะ [9] ดังนั้นงานวิจัยฉบับนี้จะใช้ข้อมูลในส่วนของอายุการใช้งานคงเหลือมาใช้ในการฝึกฝนและทดสอบโมเดลสำหรับพยากรณ์ความเสียหายของชิ้นส่วนเครื่องยนต์อากาศยานและจะไม่นำข้อมูลส่วนของคุณสมบัติจากเซนเซอร์เสมือนจริงมาใช้ ทั้งนี้รายละเอียดเกี่ยวกับการจัดการข้อมูลและการเตรียมชุดข้อมูลจะอธิบายในหัวข้อถัดไป

3.2 การจัดการชุดข้อมูล

ในขั้นตอนของการจัดการกับชุดข้อมูลจะเป็นการเตรียมชุดข้อมูลให้มีความพร้อมสำหรับใช้ในการฝึกฝนและทดสอบโมเดล ในขั้นตอนนี้ค่อนข้างมีความสำคัญอย่างมากเนื่องจากความไม่เข้าใจในลักษณะของข้อมูลจะส่งผลให้การวิเคราะห์ผลลัพธ์ที่ได้นั้นผิดพลาดไป และข้อมูลที่ไม่มีความสมบูรณ์จะส่งผลให้การฝึกฝนโมเดลนั้นเกิดปัญหาต่าง ๆ ขึ้นได้ ดังนั้นในขั้นตอนนี้จะอธิบายถึงรายละเอียดเกี่ยวกับชุดข้อมูลและขั้นตอนที่ใช้ในการเตรียมข้อมูลเพื่อเข้าสู่กระบวนการถัดไป [69]

3.2.1 การอ่านชุดข้อมูล

ตามที่ได้กล่าวไว้ในส่วนของรายละเอียดข้อมูลว่าไฟล์ที่ดาวน์โหลดมาได้นั้นเป็นไฟล์นามสกุล .h5 หรือ Hierarchical Data Format 5 ซึ่งเป็นไฟล์ที่นิยมใช้สำหรับเก็บข้อมูลขนาดใหญ่ สามารถแบ่งออกได้เป็นสองประเภทหลักได้แก่ ประเภทกลุ่ม (Groups) ที่จัดเก็บข้อมูลในรูปแบบ Dictionaries และประเภทชุดข้อมูล (Datasets) ที่จะจัดเก็บข้อมูลในรูปแบบ Numpy Array ในส่วนของชุดข้อมูลนี้จะถูกเก็บในประเภทของชุดข้อมูล

สำหรับการอ่านข้อมูลที่ถูกจัดเก็บในรูปแบบนี้จำเป็นต้องใช้โมดูลที่มีชื่อว่า h5py ข้อมูลที่อ่านได้จากทั้ง 9 ไฟล์ มีจำนวน 69,900,301 ชุดข้อมูล เนื่องด้วยข้อจำกัดทางฮาร์ดแวร์ของคอมพิวเตอร์ ทางผู้วิจัยจึงได้ทำการเลือกชุดข้อมูลสำหรับใช้ในงานวิจัยนี้ โดยจะใช้ชุดข้อมูลจากไฟล์ N-CMAPSS_DS01-005.h5 ทั้งหมด 7,641,868 แถว

3.2.2 การวิเคราะห์เชิงสถิติและค่าว่าง

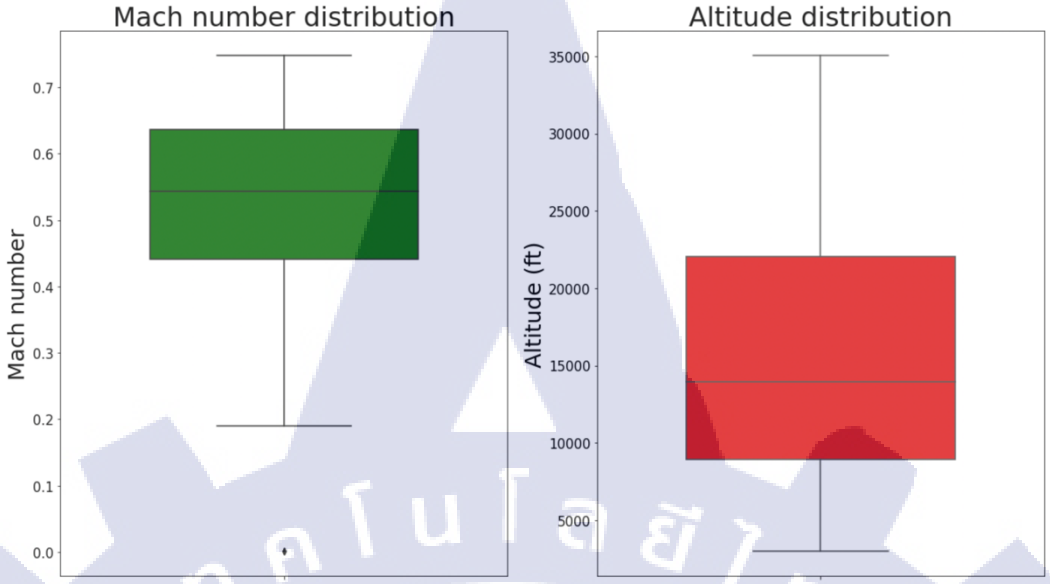
ในการวิเคราะห์ทางสถิติของชุดข้อมูลจะใช้คำสั่ง .describe() ผลลัพธ์ของการวิเคราะห์จากทั้งชุดข้อมูลพารามิเตอร์ต่าง ๆ และชุดข้อมูลแสดงสถานะการเสื่อมสภาพแสดงดังตารางที่ 3.6 สำหรับการตรวจสอบค่าว่าง (Null) จะใช้คำสั่ง .isna().sum() จากการตรวจสอบไม่พบค่าว่างในชุดข้อมูลที่นำมาใช้

ตารางที่ 3.6 การวิเคราะห์ทางสถิติของชุดข้อมูลพารามิเตอร์ต่าง ๆ

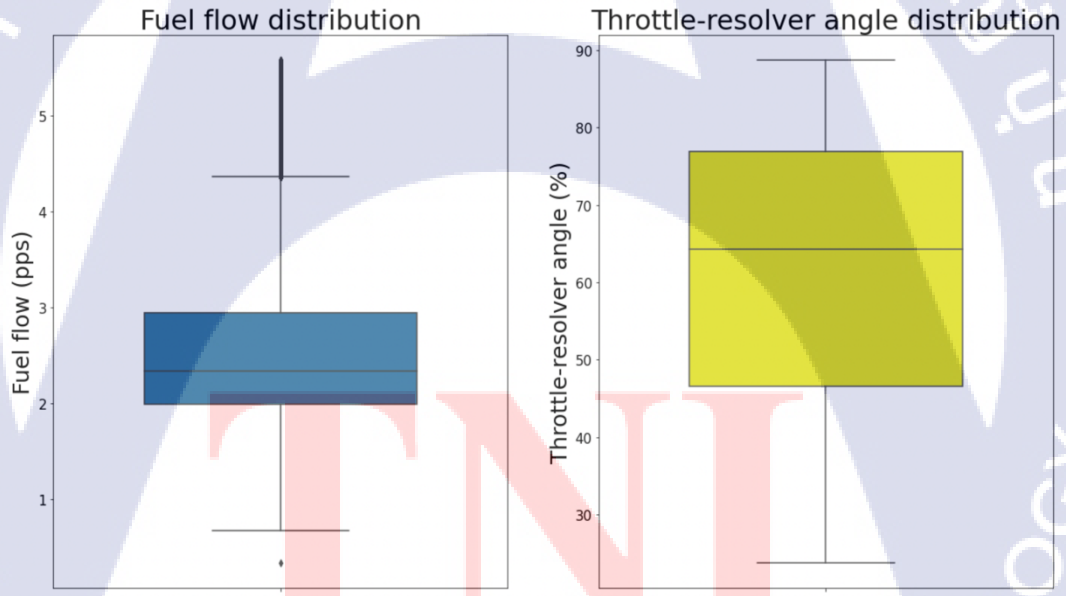
	mean	std	min	25%	50%	75%	max
alt	15447.74	8002.19	3001.00	8953.00	13968.00	22037.00	35032.00
Mach	0.54	0.12	0.00	0.44	0.54	0.64	0.75
TRA	60.29	18.43	23.73	46.58	64.25	76.90	88.77
T2	490.79	19.61	421.88	474.82	495.23	506.80	534.38
T24	570.08	20.90	484.21	555.28	567.79	583.59	634.00
T30	1331.13	68.32	1069.98	1284.78	1326.67	1370.42	1525.87
T48	1641.04	124.19	955.92	1555.21	1649.61	1719.52	1985.97
T50	1130.82	62.35	698.96	1087.10	1121.03	1165.61	1344.46
P15	13.02	2.85	5.92	10.56	13.27	15.25	20.45
P2	10.16	2.40	4.40	8.04	10.46	12.05	15.68
P21	13.21	2.90	6.01	10.72	13.47	15.48	20.76
P24	16.04	3.42	6.92	13.28	16.06	18.43	26.42
Ps30	237.67	58.78	80.35	193.93	225.20	271.14	448.50
P40	241.90	59.60	82.11	197.55	229.52	275.91	455.39
P50	10.16	2.73	4.13	7.71	10.34	12.25	16.72
Nf	1956.39	188.14	1486.24	1837.30	2002.11	2117.22	2285.51
Nc	8239.15	226.91	7371.68	8085.30	8231.17	8373.47	8865.98
Wf	2.55	0.78	0.34	1.99	2.34	2.94	5.59
RUL	43.88	25.92	0.00	22.00	44.00	65.00	99.00

3.2.3 การวิเคราะห์การกระจายและค่า Outliers

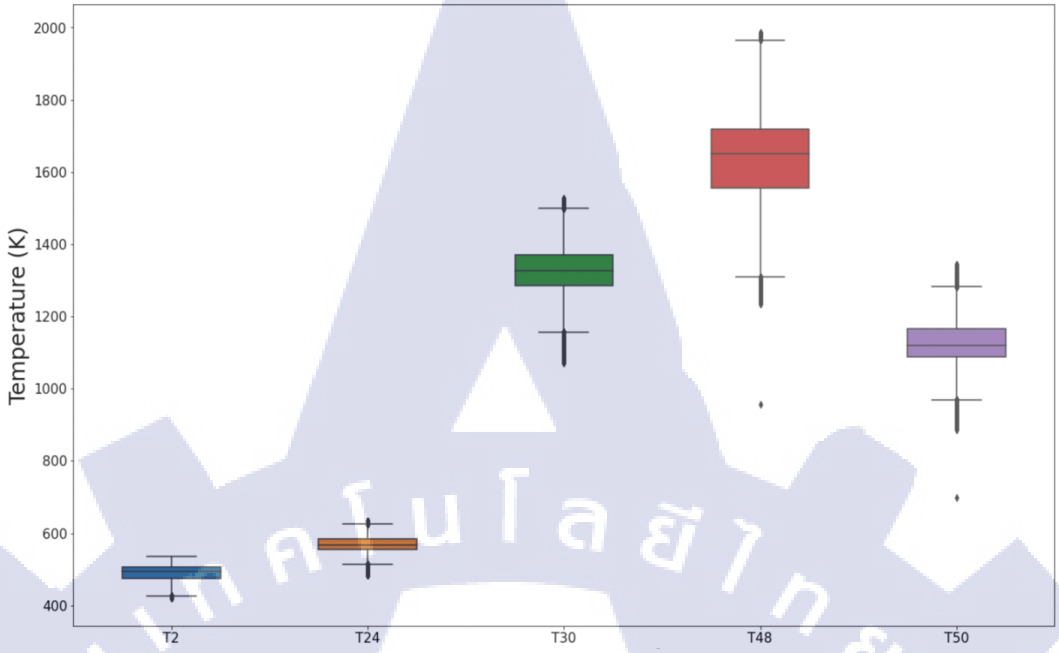
ในการวิเคราะห์ส่วนนี้จะใช้ Boxplot หรือ Box and whisker plot ในการแสดงผล ซึ่งจะมีลักษณะเป็นกล่องสี่เหลี่ยมแสดงค่าควอไทล์ที่ 1 ค่ามัธยฐาน ค่าควอไทล์ที่ 3 และจะมีเส้นที่ต่อเนื่องมาจากตัวกล่องเพื่อแสดงค่าสูงสุดและต่ำสุด สำหรับค่า Outliers จะแสดงเป็นจุดที่อยู่นอกเหนือจากเส้นดังกล่าว Boxplot ของชุดข้อมูลนี้แสดงดังรูปที่ 3.3 ถึงรูปที่ 3.8



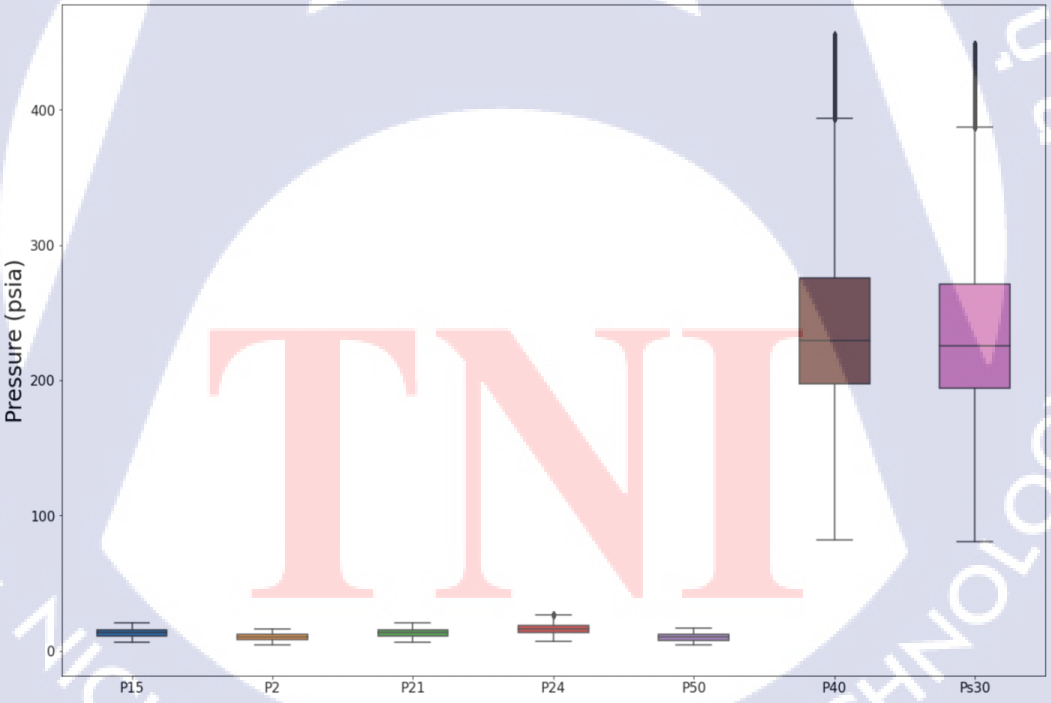
รูปที่ 3.3 การกระจายของเลขมัคและความสูง



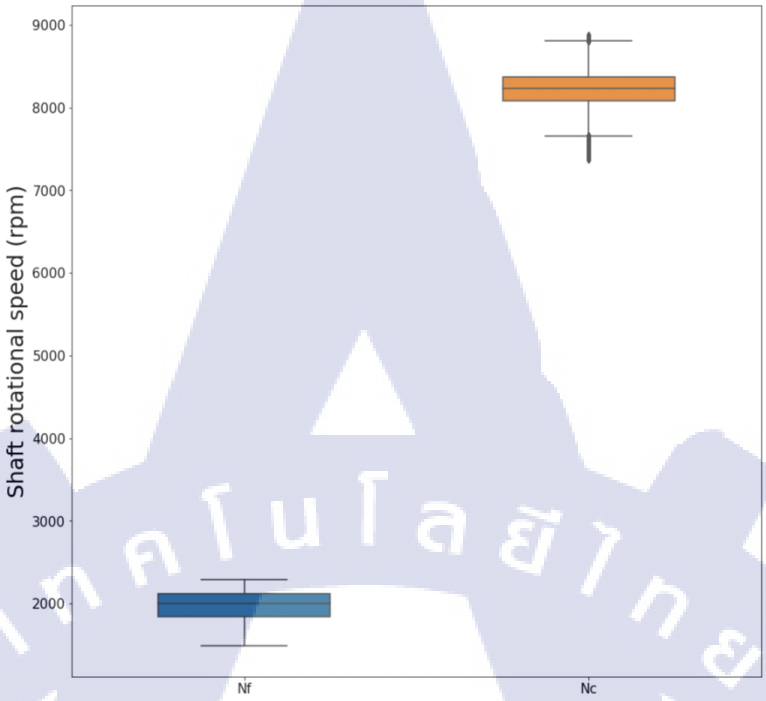
รูปที่ 3.4 การกระจายของอัตราการไหลเชื้อเพลิงและมุม Throttle-resolver



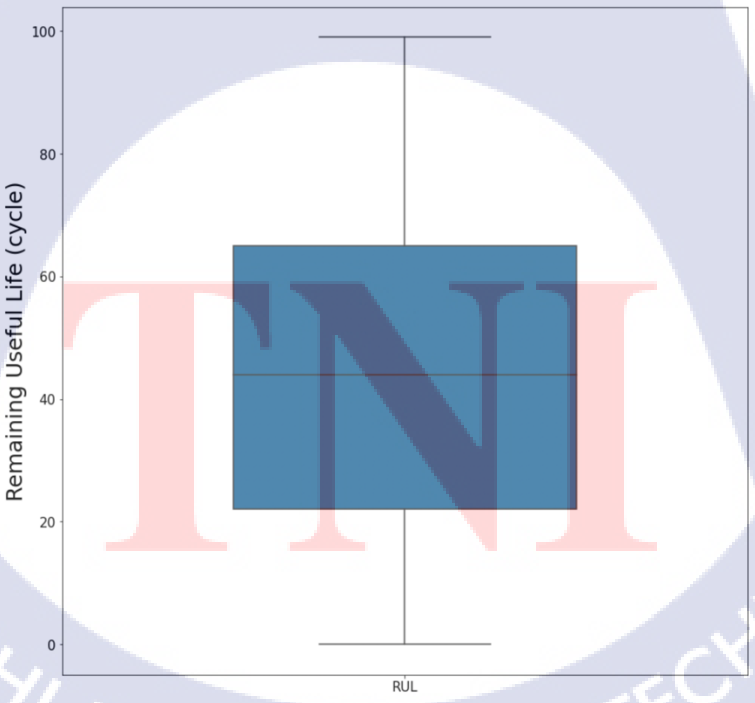
รูปที่ 3.5 การกระจายของอุณหภูมิที่ได้จากการวัด



รูปที่ 3.6 การกระจายของความดันที่ได้จากการวัด



รูปที่ 3.7 การกระจายของความเร็วการหมุนของเพลา

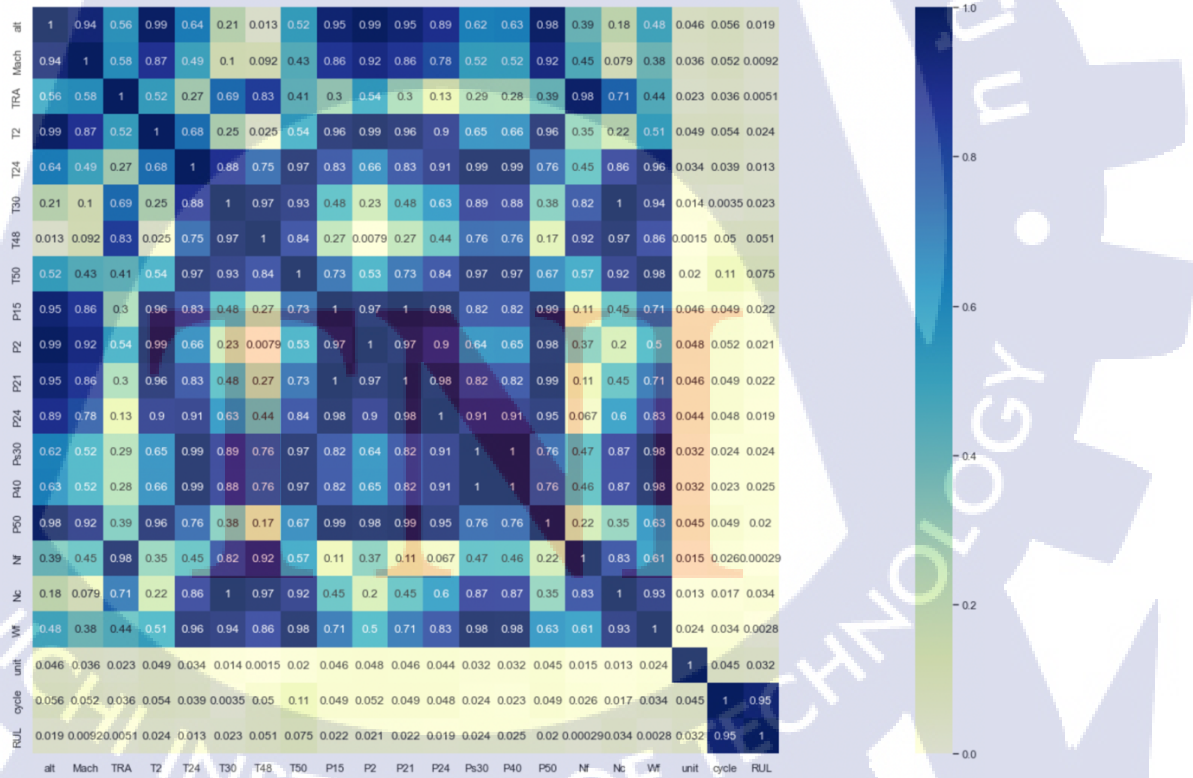


รูปที่ 3.8 การกระจายของอายุการใช้งานคงเหลือ

3.2.4 การวิเคราะห์สหสัมพันธ์ (Correlation)

การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร 2 ตัวจะทำให้ทราบถึงความเกี่ยวข้องหรือความสัมพันธ์ของตัวแปรที่ทำการวิเคราะห์อยู่ โดยปกติแล้วค่าสัมประสิทธิ์สหสัมพันธ์ซึ่งถูกใช้เป็นตัวชี้วัดความสัมพันธ์นี้จะมีค่าอยู่ระหว่าง -1 ถึง 1 ในกรณีที่ค่าติดลบจะหมายถึงตัวแปรทั้งสองมีความสัมพันธ์ในเชิงตรงกันข้าม ในกรณีที่มีค่าบวกหมายถึงมีความสัมพันธ์กัน หากเท่ากับ 0 จะสามารถตีความได้ว่า ไม่มีความสัมพันธ์ต่อกันเลย

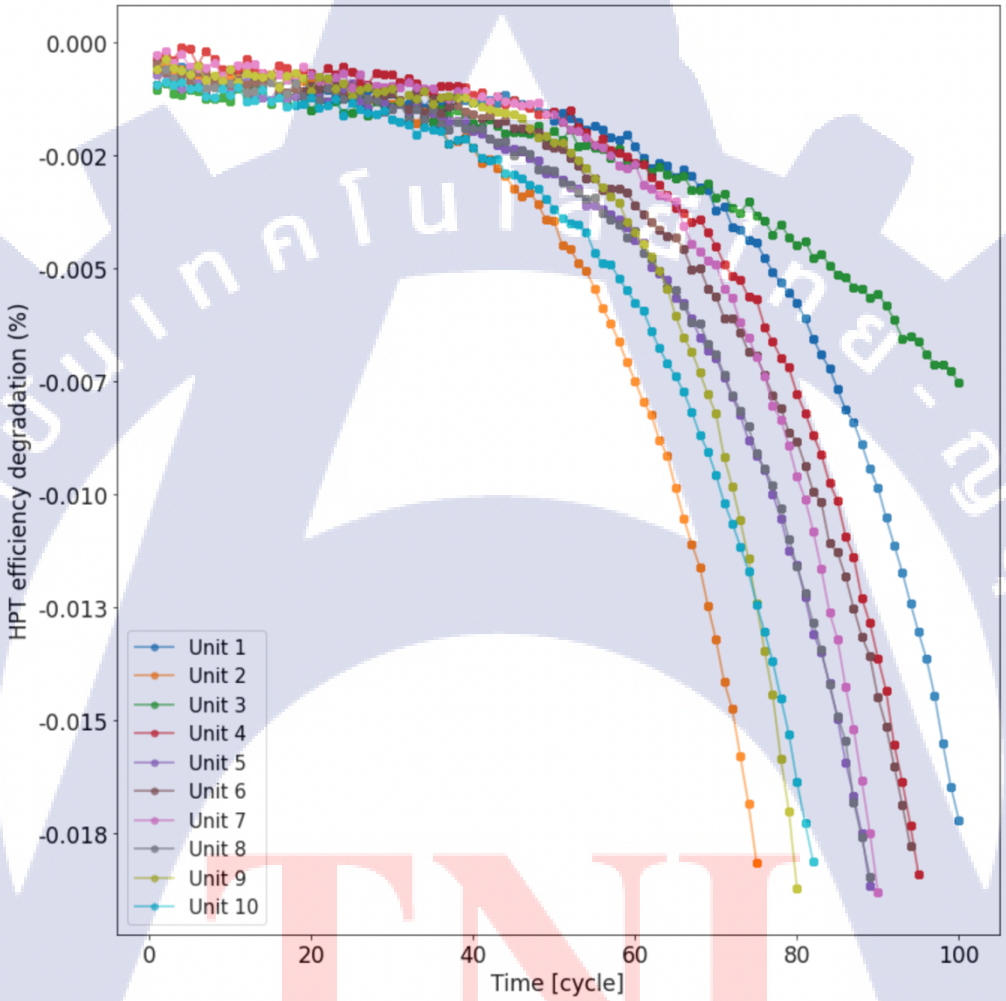
ในส่วนของงานวิจัยฉบับนี้ต้องการจะพิจารณาความเกี่ยวข้องกันระหว่างตัวแปรโดยไม่คำนึงถึงค่าติดลบ ซึ่งหมายความว่าการวิเคราะห์นี้ต้องการพิจารณาว่ามีความเกี่ยวข้องกันหรือไม่ แต่ไม่ได้พิจารณาถึงทิศทางของความสัมพันธ์ระหว่างตัวแปร ดังนั้นในรูปที่ 3.9 จะแสดงให้เห็นถึงค่าสัมประสิทธิ์สหสัมพันธ์ที่มีขอบเขตตั้งแต่ 0 ถึง 1 โดยจะแสดงในรูปแบบของ Heatmap เพื่อให้สามารถเข้าใจได้ง่ายยิ่งขึ้น เห็นได้ว่าพื้นที่ที่มีสีน้ำเงินเข้มแสดงถึงตัวแปร 2 ตัวที่มีความสัมพันธ์กันสูงมากและจะลดหลั่นกันลงมาตามแถบสีที่แสดงไว้ทางด้านขวามือของรูปภาพ เว้นแนวเส้นทแยงมุมเนื่องจากเป็นค่าของตัวแปรนั้น ๆ



รูปที่ 3.9 การวิเคราะห์สหสัมพันธ์ของชุดข้อมูล

3.2.5 การวิเคราะห์แนวโน้มการเสื่อมสภาพ (Degradation)

พารามิเตอร์แสดงสถานะการเสื่อมสภาพสำหรับไฟล์ที่เลือกมานั้นจะมีการเก็บเฉพาะค่าที่แสดงสถานะการเสื่อมสภาพของประสิทธิภาพกังหันความดันสูงเท่านั้นและจะเก็บข้อมูลเครื่องยนต์ตั้งแต่หมายเลข 1 ถึงหมายเลข 10 ดังแสดงในรูปที่ 3.10



รูปที่ 3.10 แนวโน้มการเสื่อมสภาพวัดจากประสิทธิภาพกังหันความดันสูง

3.3 การสร้างโมเดล

งานวิจัยฉบับนี้ได้เลือกเทคนิคการเรียนรู้ของเครื่องแบบมีผู้สอนมาใช้ในการเปรียบเทียบประสิทธิภาพในการคาดการณ์ความเสียหายทั้งหมด 8 เทคนิค พารามิเตอร์ที่ใช้ในแต่ละโมเดลแสดงดังตารางที่ 3.7 และรายละเอียดของ Library ที่ใช้มีดังต่อไปนี้

1. scikit-learn เป็น Library ที่บรรจุคำสั่ง DecisionTreeRegressor RandomForest-Regressor และ KNeighborsRegressor สำหรับการเรียกใช้เพื่อสร้างโมเดลด้วยเทคนิคต้นไม้ตัดสินใจ ป่าแบบสุ่ม และเพื่อนบ้านใกล้เคียงตามลำดับ

2. XGBoost เป็น Library ที่บรรจุคำสั่ง XGBRegressor สำหรับการเรียกใช้เพื่อสร้างโมเดลด้วยเทคนิค Extreme Gradient Boosting

3. TensorFlow เป็น Library ที่บรรจุคำสั่งสำหรับสร้างโครงข่ายประสาทเทียมทั้ง 4 เทคนิคของงานวิจัยฉบับนี้

ตารางที่ 3.7 รายละเอียดพารามิเตอร์ในแต่ละเทคนิค

เทคนิค	พารามิเตอร์	
Decision Tree	criterion='squared_error'	splitter='best'
	max_depth=None	min_samples_split=2
	min_samples_leaf=1	min_weight_fraction_leaf=0.0
	max_features=None	random_state=None
	max_leaf_nodes=None	min_impurity_decrease=0.0
	ccp_alpha=0.0	
Random Forest	n_estimators=100	criterion='squared_error'
	max_depth=None	min_samples_split=2
	min_samples_leaf=1	min_weight_fraction_leaf=0.0
	max_features=1.0	max_leaf_nodes=None
	min_impurity_decrease=0.0	bootstrap=True
	oob_score=False	n_jobs=None
	random_state=None	verbose=0
	warm_start=False	ccp_alpha=0.0
	max_samples=None	
XGBoost	booster=gbtree	verbosity=1
	nthread=maximun	disable_default_eval_metric=false
	eta=3	gamma=0
	max_depth=6	min_child_weight=1
	max_delta_step=0	subsample=1

ตารางที่ 3.7 รายละเอียดพารามิเตอร์ในแต่ละเทคนิค (ต่อ)

เทคนิค	พารามิเตอร์	
XGBoost	sampling_method=uniform	colsample_bytree,
	lambda=1	colsample_bylevel,
	alpha=0	colsample_bynode=1
	tree_method=auto	sketch_eps=0.03
	scale_pos_weight=1	refresh_leaf=1
	process_type=default	grow_policy=depthwise
	max_leaves=0	max_bin=256
	predictor=auto	num_parallel_tree=1
k-Nearest Neighbors	n_neighbors=5	weights='uniform'
	algorithm='auto'	leaf_size=30
	p=2	metric='minkowski'
	metric_params=None	n_jobs=None
ANN	input node=18	kernel_initializer='normal'
	hidden layer=3	activation='relu'
	hidden node=32	loss='mean_squared_error'
	output node=1	optimizer='adam'
	Dropout=0.12	epochs=2000 *Earlystopping
	monitor='loss'	patience=64
CNN	Conv1D filters=64	kernel_size=3
	activation='relu'	input_shape=(18,1)
	padding='same'	MaxPool1D(pool_size=2)
	Convolution layer=2	Pooling layer =2
	hidden node=64	activation='relu'
	optimizer='adam'	loss='mean_squared_error'
	patience=64	epochs=2000 *Earlystopping
	monitor='loss'	
RNN	input node=18	activation='tanh'
	hidden layer=3	return_sequences=True

ตารางที่ 3.7 รายละเอียดพารามิเตอร์ในแต่ละเทคนิค (ต่อ)

เทคนิค	พารามิเตอร์	
RNN	output node=1	hidden node=32
	optimizer='adam'	loss='mean_squared_error'
	patience=64	epochs=2000 *Earlystopping
	monitor='loss'	
LSTM	input node=18	activation='tanh'
	hidden layer=2	return_sequences=True
	output node=1	hidden node=32
	optimizer='adam'	loss='mean_squared_error'
	patience=64	epochs=2000 *Earlystopping
	monitor='loss'	

3.4 การประเมินโมเดล

ถัดมาหลังจากที่ได้ทำการสร้างและทดสอบโมเดลทั้งหมดแล้วคือขั้นตอนของการประเมินโมเดล ซึ่งก็คือการประเมินความสามารถและความแม่นยำในการทำงานของโมเดล เพื่อทำการตัดสินใจว่าโมเดลที่ได้ทำการสร้างขึ้นนั้นมีความเหมาะสมเพียงพอต่อการนำไปใช้งานหรือไม่ เนื่องจากในงานวิจัยฉบับนี้ต้องการการวิเคราะห์แบบถดถอย ดังนั้นเทคนิคการเรียนรู้ของเครื่องแบบมีผู้สอนที่ผู้วิจัยเลือกใช้ต้องสามารถใช้กับการวิเคราะห์เช่นนี้ได้ ซึ่งจะส่งผลให้การประเมินโมเดลนั้นมีความแตกต่างออกไปเช่นกัน

ในงานวิจัยฉบับนี้ต้องการคาดการณ์ความเสียหายของชิ้นส่วนเครื่องจักรซึ่งเป็นลักษณะข้อมูลที่มีการอ้างอิงตามเวลา จึงเลือกใช้เทคนิคแบบถดถอยและจะให้ค่าผลลัพธ์ในลักษณะของค่าตัวเลข ดังนั้นการประเมินโมเดลในงานวิจัยฉบับนี้จึงจะเลือกใช้การวัดระยะคลาดเคลื่อนของการคาดการณ์ (Errors) ด้วยวิธีดังต่อไปนี้

1. สัมประสิทธิ์การกำหนด (Coefficient of Determination: R^2)

$$R^2 = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y - \bar{y})^2} \tag{3.1}$$

2. ความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error: MSE) และรากที่สองของความผิดพลาดกำลังสองเฉลี่ย (Root Mean Square Error: RMSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y - \hat{y})^2 \tag{3.2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y - \hat{y})^2} \tag{3.3}$$

เมื่อ y คือค่าจริง (Actual Value)

$\hat{y}$ คือค่าทำนาย (Predicted Value)

$\bar{y}$ คือค่าเฉลี่ยของค่าจริง

n จำนวนข้อมูล

3.5 การเปรียบเทียบประสิทธิภาพ

โมเดลที่ผ่านการประเมินด้วยค่าสัมประสิทธิ์การกำหนดและค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ยจะถูกนำมาเปรียบเทียบผลลัพธ์ที่ได้ โดยจะแบ่งออกเป็นการเปรียบเทียบในภาพรวมจากทั้งหมด 8 โมเดลและการเปรียบเทียบในแต่ละพารามิเตอร์แสดงสถานะการเสื่อมสภาพของแต่ละโมเดล ซึ่งจะแสดงผลในรูปแบบของตารางและกราฟในบทถัดไป

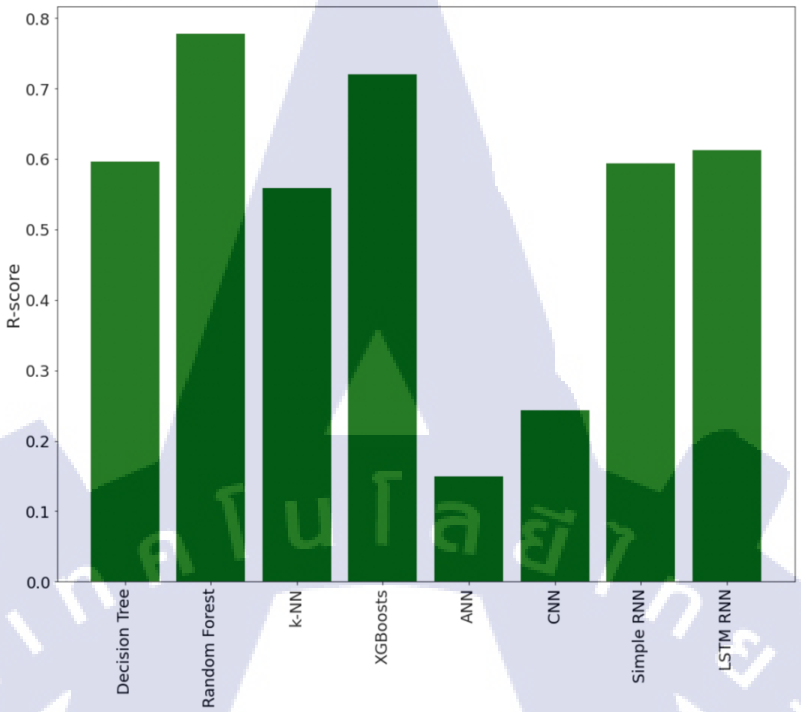
บทที่ 4
ผลการวิจัย

จากการฝึกฝนและทดสอบโมเดลทั้งหมด 8 เทคนิคด้วยชุดข้อมูล N-CMAPSS ประสิทธิภาพของโมเดลวัดจากค่าสัมประสิทธิ์การกำหนด (R^2) และค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ย (RMSE) แสดงดังตารางที่ 4.1 โดยสัมประสิทธิ์การกำหนดจะมีค่าอยู่ระหว่าง 0-1 หากมีค่าเข้าใกล้ 0 นั้นหมายความว่าโมเดลนั้นมีความผิดพลาดสูง แต่หากเข้าใกล้ 1 ก็สามารถแปลความได้ว่าโมเดลนั้นค่อนข้างมีประสิทธิภาพสูง ในส่วนของการวัดประสิทธิภาพจากค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ยนั้นจะเป็นการวัดค่าความคลาดเคลื่อน ดังนั้นค่าที่ต่ำก็แสดงถึงความคลาดเคลื่อนที่ต่ำและตีความได้ว่าโมเดลนั้นมีประสิทธิภาพสูงสำหรับชุดข้อมูลที่นำมาใช้

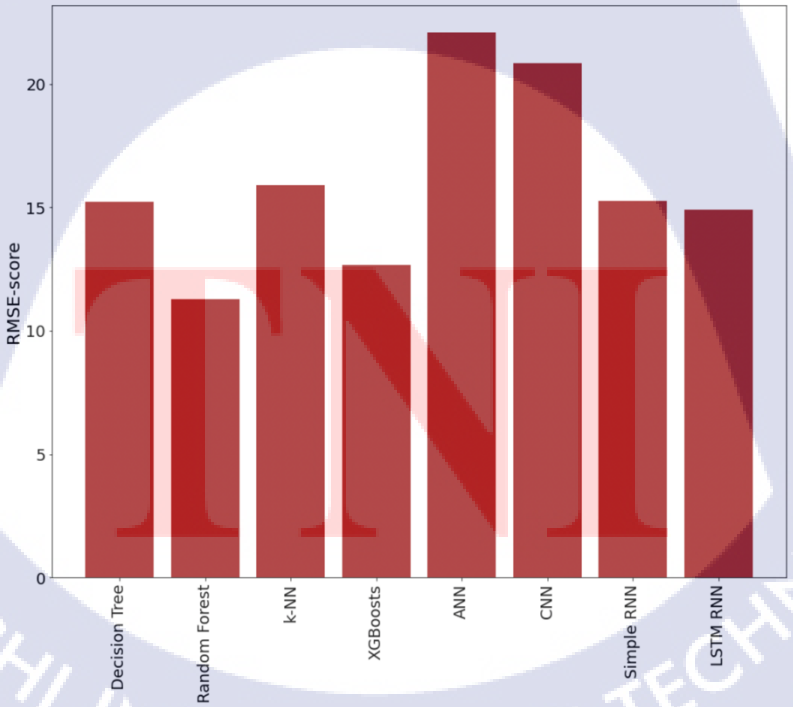
ตารางที่ 4.1 ผลการประเมินโมเดลในแต่ละเทคนิค

เทคนิค	R^2	RMSE
Decision Tree	0.5959	15.2267
Random Forest	0.7780	11.2844
k-Nearest Neighbors	0.5584	15.9170
XGBoosts	0.7203	12.6674
ANN	0.1495	22.0890
CNN	0.2431	20.8389
Simple RNN	0.5942	15.2574
LSTM RNN	0.6129	14.9031

จากผลลัพธ์ที่ได้แสดงให้เห็นว่าเทคนิคป่าแบบสุ่มนั้นให้ค่าสัมประสิทธิ์การกำหนดที่เข้าใกล้ 1 มากที่สุด รองลงมาได้แก่เทคนิค Extreme Gradient Boosting และโครงข่ายประสาทเทียมวกกลับแบบ LSTM ตามลำดับ ในส่วนของค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ยที่ต่ำที่สุดได้แก่เทคนิคป่าแบบสุ่ม เพิ่มขึ้นมาเล็กน้อยได้แก่เทคนิค Extreme Gradient Boosting และโครงข่ายประสาทเทียมวกกลับแบบ LSTM ตามลำดับ เพื่อให้เห็นภาพที่ชัดเจนยิ่งขึ้นการเปรียบเทียบในรูปแบบของกราฟแสดงดังรูปที่ 4.1 และ 4.2



รูปที่ 4.1 กราฟแสดงการเปรียบเทียบประสิทธิภาพจากค่า R^2



รูปที่ 4.2 กราฟแสดงการเปรียบเทียบประสิทธิภาพจากค่า RMSE

บทที่ 5
สรุปและอภิปรายผล

5.1 สรุปและอภิปรายผล

หากพิจารณาถึงชุดข้อมูลที่นำมาใช้ในงานวิจัยฉบับนี้จะพบว่าเป็นข้อมูลที่มีลักษณะเป็นแบบอนุกรมเวลาและเป็นการวิเคราะห์แบบถดถอย โดยค่าที่ต้องกาคาดการณ์ได้แก่ อายุการใช้งานคงเหลือของชิ้นส่วนเครื่องยนต์อากาศยาน ดังนั้นรูปแบบของโมเดลจึงมีลักษณะเป็นแบบเชิงเส้น จากผลการวิจัยที่ได้แสดงไว้ในบทที่ 4 สามารถสรุปได้ว่าโมเดลที่มีประสิทธิภาพสูงที่สุดได้แก่โมเดลป่าแบบสุ่ม รองลงมาได้แก่ Extreme Gradient Boosting และโครงข่ายประสาทเทียมวกกลับแบบ LSTM ตามลำดับ โดยวัดจากค่าสัมประสิทธิ์การกำหนดและค่ารากที่สองของความผิดพลาดกำลังสองเฉลี่ย ซึ่งการประเมินประสิทธิภาพด้วยค่าทั้งสองค่านี้ให้ผลลัพธ์ที่มีความสอดคล้องกัน

เทคนิคป่าแบบสุ่มเป็นเทคนิคที่ได้รับการปรับปรุงมาจากเทคนิคต้นไม้ตัดสินใจโดยใช้แนวทางการเรียนรู้แบบ Ensemble คือมีการสร้างต้นไม้ตัดสินใจหลายต้นเช่นเดียวกับเทคนิค Extreme Gradient Boosting แตกต่างเพียงเทคนิคป่าแบบสุ่มจะใช้วิธี Bagging หรือการเรียนรู้แบบคู่ขนานแล้วจึงทำการหาค่าเฉลี่ยของต้นไม้แต่ละต้นที่สร้างขึ้นมา ในขณะที่เทคนิค Extreme Gradient Boosting จะใช้วิธี Boosting หรือการเรียนรู้เป็นลำดับมีการปรับปรุงโมเดลในทุกรอบของการเรียนรู้ ดังนั้นเทคนิคป่าแบบสุ่มและเทคนิค Extreme Gradient Boosting จึงให้ผลลัพธ์ที่ดีกว่าต้นไม้ตัดสินใจ

จากเทคนิคทั้งหมดที่ผู้วิจัยเลือกใช้ในงานวิจัยฉบับนี้สามารถแบ่งออกได้เป็น 2 กลุ่มหลัก ได้แก่ กลุ่มของเทคนิคการเรียนรู้ของเครื่องและกลุ่มของเทคนิคโครงข่ายประสาทเทียม โดยกลุ่มเทคนิคการเรียนรู้ของเครื่องจะประกอบไปด้วยเทคนิคต้นไม้ตัดสินใจ ป่าแบบสุ่ม ขั้นตอนวิธีเพื่อนบ้านใกล้สุด และ Extreme Gradient Boosting ในขณะที่กลุ่มของเทคนิคโครงข่ายประสาทเทียมจะประกอบไปด้วยเทคนิคโครงข่ายประสาทเทียม โครงข่ายประสาทเทียมคอนโวลูชัน โครงข่ายประสาทเทียมวกกลับแบบ Simple และโครงข่ายประสาทเทียมวกกลับแบบ LSTM

การแบ่งกลุ่มในที่นี้ผู้วิจัยต้องการชี้ให้เห็นว่าเมื่อพิจารณาเฉพาะกลุ่มของโครงข่ายประสาทเทียมจะพบว่าเทคนิคโครงข่ายประสาทเทียมวกกลับแบบ LSTM นั้นให้ผลลัพธ์ที่ดีที่สุด รองลงมาคือโครงข่ายประสาทเทียมวกกลับแบบ Simple เห็นได้ชัดว่าโครงข่ายประสาทเทียมวกกลับนั้นมีประสิทธิภาพสูงกว่าโครงข่ายประสาทเทียมอื่นสำหรับชุดข้อมูลแบบอนุกรมเวลา

อย่างไรก็ตามเทคนิคโครงข่ายประสาทเทียมนั้นประกอบด้วยองค์ประกอบหลายประการและสามารถมีรูปแบบของสถาปัตยกรรมที่มีความหลากหลายได้ แต่ด้วยข้อจำกัดที่จะขอกล่าวในหัวข้อถัดไปนั้นส่งผลให้รูปแบบของโครงข่ายประสาทเทียมที่ใช้ในงานวิจัยฉบับนี้เป็นเพียงรูปแบบอย่างง่ายหรือรูปแบบพื้นฐานเพียงเท่านั้น จึงอาจสรุปได้เพียงว่ากลุ่มของโครงข่ายประสาทเทียมอย่างง่ายนั้นมีประสิทธิภาพที่ด้อยกว่าเทคนิคป่าแบบสุ่มสำหรับชุดข้อมูลนี้ แต่ไม่สามารถสรุปได้ว่ากลุ่มโครงข่ายประสาทเทียมไม่มีประสิทธิภาพเพียงพอหรือไม่เหมาะสมที่จะนำมาใช้กับข้อมูลชุดนี้ได้

5.2 ข้อจำกัดในการวิจัย

เนื่องจากชุดข้อมูลที่เลือกใช้นั้นมีขนาดที่ค่อนข้างใหญ่ มีการเก็บตัวแปรและข้อมูลจำนวนมาก การประมวลผลทั้งหมดจำเป็นที่จะต้องใช้คอมพิวเตอร์ประสิทธิภาพสูง แต่ด้วยข้อจำกัดด้านฮาร์ดแวร์ของคอมพิวเตอร์ที่ใช้ในการวิจัย จึงจำเป็นต้องเลือกข้อมูลบางส่วนมาทำการวิจัยเท่านั้น นอกเหนือจากนี้ยังส่งผลต่อการสร้างโมเดลในกลุ่มของโครงข่ายประสาทเทียม ซึ่งต้องการหน่วยประมวลผลที่มีประสิทธิภาพและหน่วยความจำที่มากเพียงพอสำหรับประมวลผลในรูปแบบสถาปัตยกรรมที่มีความซับซ้อนสูง

5.3 ข้อเสนอแนะ

จากการดำเนินงานวิจัยฉบับนี้ทางผู้วิจัยได้ประสบปัญหาและข้อเสนอแนะต่าง ๆ ที่เล็งเห็นว่าควรนำเสนอไว้ในหัวข้อนี้ เพื่อเป็นประโยชน์ให้กับผู้ที่ต้องการศึกษางานวิจัยฉบับนี้ต่อหรือต้องการที่จะนำไปประยุกต์ใช้งานจริง โดยมีรายละเอียดดังต่อไปนี้

1. ด้วยข้อจำกัดในการวิจัยฉบับนี้ที่ได้กล่าวไปในหัวข้อก่อนหน้า ทางผู้วิจัยเห็นว่ามีความน่าสนใจที่จะศึกษาโครงข่ายประสาทเทียมในเทคนิคต่าง ๆ ที่มีรูปแบบของสถาปัตยกรรมที่มีความซับซ้อนสูงขึ้นสำหรับชุดข้อมูลนี้ ซึ่งอาจต้องการหน่วยความจำ (RAM) จำนวนมากเพื่อทำการฝึกฝนโมเดลที่มีความซับซ้อนให้มีประสิทธิภาพเพียงพอต่อการนำไปใช้งานจริง

2. โมเดลการคาดการณ์ความเสียหายจากเทคนิคการเรียนรู้ของเครื่องที่ถูกพัฒนาขึ้นในงานวิจัยฉบับนี้ใช้ข้อมูลที่มีความเกี่ยวข้องกับการไหลและอุณหพลศาสตร์เป็นชุดข้อมูลสำหรับฝึกฝนและทดสอบเป็นหลัก จึงอาจตั้งสมมติฐานได้ว่าผลการวิจัยฉบับนี้สามารถใช้สำหรับความเสียหายในเชิงของการไหลและอุณหพลศาสตร์เท่านั้น หากเป็นความเสียหายในรูปแบบอื่นอาทิเช่น ความเสียหายที่เกิดขึ้นจากความล้า (Fatigue Failure) ซึ่งหมายถึงการที่ชิ้นส่วนนั้น ๆ ต้องรับแรงกระทำ

อย่างต่อเนื่องเป็นระยะเวลาหนึ่งจนกระทั่งเกิดความเสียหายขึ้น ก็จำเป็นที่จะต้องทำการศึกษาและนำชุดข้อมูลที่เกี่ยวข้องมาทำการฝึกฝนและทดสอบโมเดลใหม่ทั้งหมด

3. งานการซ่อมบำรุงเชิงคาดการณ์นั้นควรที่จะต้องคาดการณ์ถึงระยะเวลาที่จะเกิดความเสียหายให้ได้อย่างแม่นยำเพื่อให้การซ่อมบำรุงนั้นเกิดประโยชน์สูงสุด แต่ในทางปฏิบัติจริงความปลอดภัยเป็นสิ่งที่สำคัญยิ่ง ทางผู้วิจัยจึงแนะนำให้ศึกษาเพิ่มเติมในส่วนของค่าความปลอดภัย (Safety factor) สำหรับการนำไปประยุกต์ใช้งานจริง

ภาคผนวก

ภาคผนวก ก.

รายละเอียดโมเดล

Decision Tree

```
from sklearn.tree import DecisionTreeRegressor
dt = DecisionTreeRegressor(random_state=12345)
dt.fit(X_train, y_train)
```

Random Forest

```
from sklearn.ensemble import RandomForestRegressor
rf = RandomForestRegressor()
rf.fit(X_train, y_train)
```

k-Nearest Neighbors

```
from sklearn.neighbors import KNeighborsRegressor
knn = KNeighborsRegressor()
knn.fit(X_train, y_train)
```

XGBoost

```
from xgboost import XGBRegressor
my_xgb = XGBRegressor()
my_xgb.fit(X_train_ar, y_train_ar)
```

Artificial Neural Network

```
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import Dropout
model = Sequential()
model.add(Dense(32, input_dim=18, kernel_initializer='normal', activation='relu'))
```

```
model.add(Dropout(0.12))
model.add(Dense(32, kernel_initializer='normal', activation='relu'))
model.add(Dense(10, kernel_initializer='normal'))
model.compile(loss='mean_squared_error', optimizer='adam')

model.summary()
```

Convolutional Neural Network

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv1D, MaxPool1D, Flatten, Dense

model_cnn = Sequential()
model_cnn.add(Conv1D(filters=64, kernel_size=3, activation='relu',
input_shape=(18,1), padding='same'))
model_cnn.add(MaxPool1D(pool_size=2))
model_cnn.add(Conv1D(filters=128, kernel_size=3, activation='relu', padding='same'))
model_cnn.add(MaxPool1D(pool_size=2))
model_cnn.add(Flatten())
model_cnn.add(Dense(32, activation='relu'))
model_cnn.add(Dense(64, activation='relu'))
model_cnn.add(Dense(1))
model_cnn.compile(loss='mean_squared_error', optimizer='adam')

model_cnn.summary()
```

Recurrent Neural Network

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, SimpleRNN
```

```
simple_model = Sequential([
  SimpleRNN(32, activation='tanh',input_shape=(18,1),
  return_sequences=True),
  SimpleRNN(32, activation='tanh', return_sequences = True),
  SimpleRNN(32, activation='tanh'),
  Dense(y_train_ar.shape[1]),
])
simple_model.compile(loss='mean_squared_error', optimizer='adam')

simple_model.summary()
```

LSTM Recurrent Neural Network

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, LSTM

LSTM_model = Sequential([
   LSTM(32, activation='tanh',input_shape=(18, 1),
   return_sequences=True),
   LSTM(32, activation='tanh'),
  Dense(y_train.shape[1]),
])
LSTM_model.summary()
```

บรรณานุกรม

บรรณานุกรม

[1] L. Junjie et al., “Distributed kernel extreme learning machines for aircraft engine failure diagnostics,” *Applied Sciences,* vol. 9, no. 8, pp. 1-19, April 2019.

[2] L. N. Efstratios and B. Pantelis, “Diagnostic methods for an aircraft engine performance,” *Journal of Engineering Science and Technology Review,* vol. 8, no. 4, pp. 64-72, November 2015.

[3] C. Ana et al., “Maintenance 4.0: Intelligent and predictive maintenance system architecture,” in *IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, Turin, Italy, September 4-7, 2018, pp. 139-146.

[4] T. Khanh and M. Kamal, “A new dynamic predictive maintenance framework using deep learning for failure prognostics,” *Reliability Engineering & System Safety,* vol. 188, pp. 251-262, August 2019.

[5] A. S. Gian et al., “Machine learning for predictive maintenance: A multiple classifier approach,” *IEEE Transactions on Industrial Informatics,* vol. 11, no. 3, pp. 812-820, June 2015.

[6] Z. Tiago et al., “Predictive maintenance in the Industry 4.0: A systematic literature review,” *Computers & Industrial Engineering,* vol. 150, pp. 1-17, December 2020.

[7] P. C. Thyago et al., “A systematic literature review of machine learning methods applied to predictive maintenance,” *Computers & Industrial Engineering,* vol. 137, pp. 1-10, November 2019.

[8] S. Abhinav et al., “Damage propagation modeling for aircraft engine run-to-failure simulation,” in *International Conference on Prognostics and Health Management*, Denver, CO, USA, October 6-9, 2008, pp. 1-9.

[9] A. C. Manuel et al., “Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics,” *Data,* vol. 6, no. 1, pp. 1-14, January 2021.

[10] K. Zhou et al., “Industry 4.0: Towards future industrial opportunities and challenges,” in *12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, Zhangjiajie, August 15-17, 2015, pp. 2147-2152.

[11] L. Heiner et al., “Industry 4.0,” *Business & Information Systems Engineering,* vol. 6, no. 4, pp. 239-242, June 2014.

[12] E. Oztemel and S. Gursev, “Literature review of Industry 4.0 and related technologie,” *Journal of Intelligent Manufacturing,* vol. 31, no. 1, pp. 127-182, July 2018.

[13] M. A. K. Bahrin et al., “Industry 4.0: A review on industrial automation and robotic,” *Advances on Intelligent Control and Robotic,* vol. 78, no. 6-13, pp. 137-143, June 2016.

[14] Z. Li et al., “Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario,” *Advances in Manufacturing,* vol. 5, no. 3, pp. 377-387, December 2017.

[15] A. Frank et al., “Industry 4.0 technologies: Implementation patterns in manufacturing companies,” *International Journal of Production Economics,* vol. 210, no. 3, pp. 15-26, April 2019.

[16] L. Doug, *Application Delivery Strategies*, Stamford: META Group, 2001.

[17] M. Sam, “From databases to big data,” *IEEE Internet Computing* , vol. 16, no. 3, pp. 4-6, April 2012.

[18] K. Venkatram and A. G. Mary, “Review on big data & analytics – concepts, philosophy, process and applications,” *Cybernetics and Information Technologies,* vol. 17, no. 2, pp. 3-27, June 2017.

[19] J. Yan et al., “Industrial big data in an industry 4.0 environment: Challenges, schemes and applications for predictive maintenance,” *IEEE Access,* vol. 5, pp. 23484-23491, October 2017.

[20] G. Maryam, “Understanding the impact of big data on firm performance: The necessity of conceptually differentiating among big data characteristics,” *International Journal of Information Management,* vol. 57, no. 2, pp. 1-13, April 2021.

[21] V. Grover et al., “Creating strategic business value from big data analytics: A research framework,” *Journal of Management Information Systems,* vol. 35, no. 2, pp. 388-423, May 2018.

[22] R. M. Losee, “Browsing mixed structured and unstructured data,” *Information Processing & Management,* vol. 42, no. 2, pp. 440-452, March 2006.

[23] S. Sugam, “Expanded cloud plumes hiding big data ecosystem,” *Future Generation Computer Systems,* vol. 59, pp. 63-92, June 2016.

[24] A. Gandomi and M. Haider, “Beyond the hype: Big data concepts, methods, and analytics,” *International Journal of information management,* vol. 35, no. 2, pp. 137-144, April 2015.

[25] T. A. Runkler, *Data Analytics*, Germany: Springer Vieweg Wiesbaden, 2020.

[26] W. Graham, *Data Mining with Rattle and R*, New York: Springer New York, 2011.

[27] A. Banerjee et al., “Data analytics: Hyped up aspirations or true potential?,” *Vikalpa,* vol. 38, no. 4, pp. 1-11, October-December 2013.

[28] R. Soltanpoor and T. Sellis, “Prescriptive analytics for big data,” *Databases Theory and Applications,* vol. 9877, pp. 245-256, September 2016.

[29] J. Han et al., *Data Mining: Concepts and Techniques*, Netherlands: Elsevier, 2011.

[30] Y. Ibrar et al., “Big data: From beginning to future,” *International Journal of Information Management,* vol. 36, no. 6, pp. 1231-1247, December 2016.

[31] B. Attila and S. L. Cecília, *Encyclopedia of Information Science and Technology*, Hungary: IGI Global, 2009.

[32] J. Karthikeyan et al., *Learning Outcomes of Classroom Research*, India: L Ordine Nuovo Publication, 2021.

[33] A. M. Barrett and S. D. Baum, "A model of pathways to artificial superintelligence catastrophe for risk and decision analysis," *Journal of Experimental & Theoretical Artificial Intelligence,* vol. 29, no. 2, pp. 397-414, May 2016.

[34] V. Z. Henk, "Difference between artificial intelligence and traditional methods," *Artificial Intelligence Applications to Critical Transportation Issues,* vol. 6, no. 3, pp. 360-375, November 2012.

[35] S. R. John, "Minds, brains, and programs," *Behavioral and Brain Sciences ,* vol. 3, no. 3, pp. 417-424, September 1980.

[36] K. Richter and J. Walther, *Supply Chain Integration Challenges in Commercial Aerospace*, Switzerland: Springer Cham, 2017.

[37] J. Deutsch and D. He, "Using deep learning-based approach to predict remaining useful life of rotating components," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 48, no. 1, pp. 2018, May 2017.

[38] S. B. Kotsiantis, *Emerging Artificial Intelligence Applications in Computer Engineering*, Netherlands: IOS Press, 2007.

[39] G. Zoubin, "Unsupervised learning," *Advanced Lectures on Machine Learning,* vol. 3176, pp. 72-112, September 2004.

[40] Harkiran78, "Top 10 Algorithms every Machine Learning Engineer should know," [Online]. Available: https://www.geeksforgeeks.org/top-10-algorithms-every-machine-learning-engineer-should-know/. [Accessed : February 23, 2022].

[41] T. J. Bahzad and M. A. Adnan, "Classification based on decision tree algorithm for machine learning," *Journal of Applied Science and Technology Trends,* vol. 2, no. 1, pp. 20-28, March 2021.

[42] J. R. Quinlan, “Induction of decision trees,” *Machine Learning,* vol. 1, no. 1, pp. 81-106, March 1986.

[43] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal,* vol. 27, no. 3, pp. 379-423, July 1948.

[44] J. R. Quinlan, *C4.5: Programs for Machine Learning*, California: Morgan Kaufman, 1933.

[45] S. Hong, “2021.01.17(pm): Ensemble – Bagging and Boosting,” [Online]. Available: https://seongjuhong.com/2021-01-17pm-ensemble-bagging-and-boosting/. [Accessed : February 23, 2022].

[46] L. Breiman, “Bagging predictors,” *Machine Learning ,* vol. 24, no. 2, pp. 123-140, August 1996.

[47] B. Leo, “Random forests,” *Machine Learning ,* vol. 45, no. 1, pp. 5-32, October 2001.

[48] M. L., “Random Forest ทำนายว่าจักรยานจะขายได้ก็ต่อเมื่อ...,” [Online]. Available: https://medium.com/mmp-li/random-forest-ทำนายว่าจักรยานจะ ขายได้ก็ต่อเมื่อ-c60080354380. [Accessed : 23 กุมภาพันธ์ 2565].

[49] G. Ridgeway, “The state of boosting,” *Computing Science and Statistics,* vol. 31, pp. 172-181, 1999.

[50] M. Francisco et al., “A methodology for applying k-nearest neighbor to time series forecasting,” *Artificial Intelligence Review,* vol. 3, no. 52, pp. 2019-2037, December 2017.

[51] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The Bulletin of Mathematical Biophysics,* vol. 5, no. 4, pp. 115-133, December 1943.

[52] D. Hebb, *The Organization of Behavior*, New York: Psychology Press, 2002.

[53] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review,* vol. 65, no. 6, pp. 386-408, November 1958.

[54] A. C. Charu, *Neural Networks and Deep Learning*, Switzerland: Springer Cham, 2018.

[55] J. Singh and R. Banerjee, “A study on single and multi-layer perceptron neural network,” in *International Conference on Computing Methodologies and Communication*, India, March 27-29, 2019, pp. 35-40.

[56] M. Minsky and S. Papert, *Perceptrons*, Cambridge: MIT Press, 1969.

[57] Y. Lecun et al., “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE,* vol. 86, no. 11, pp. 2278-2324, November 1998.

[58] J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” *Proceedings of the National Academy of Sciences,* vol. 79, no. 8, pp. 2554-2558, April 1982.

[59] M. I. Jordan, “Serial order: A parallel distributed processing approach,” *Advances in Psychology*, vol. 121, pp. 471-495, 1997.

[60] J. L. Elman, “Finding structure in time,” *Cognitive Science,* vol. 14, no. 2, pp. 179-211, March 1990.

[61] R. K. Mobley, “Impact of maintenance,” in *An Introduction to Predictive Maintenance*, Oxford, UK, Butterworth-Heinemann, September 26, 2002, pp. 1-22.

[62] H. Hashemian, “State-of-the-Art predictive maintenance techniques,” *IEEE Transactions on Instrumentation and Measurement,* vol. 60, no. 1, pp. 226-236, December 2010.

[63] Y. Wang et al., “Remaining useful life prediction using deep learning approaches: A review,” *Procedia Manufacturing,* vol. 49, no. 6, pp. 81-88, July 2020.

[64] S. Vollert and A. Theissler, “Challenges of machine learning-based RUL prognosis: A review on NASA's C-MAPSS data set,” in *26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, Vasteras, Sweden, September 7-10, 2021, pp. 1-8.

[65] X.-S. Si et al., “Remaining useful life estimation – A review on the statistical data driven approache,” *European Journal of Operational Research,* vol. 213, no. 1, pp. 1-14, August 2011.

[66] V. Mathew et al., “Prediction of remaining useful lifetime (RUL) of turbofan engine using machine learning,” in *IEEE International Conference on Circuits and Systems (ICCS)*, Thiruvananthapuram, India, December 20-21, 2017, pp. 306-311.

[67] A. C. Manuel et al., “NASA Ames Prognostics Data Repository,” NASA Ames Research Center, [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository. [Accessed : January 26, 2022].

[68] R. Duivis, “Jet Engine Propulsion: The Comparison of Power between a Car and an Aircraft?,” KLM Blog, [Online]. Available: https://blog.klm.com/jet-engine-propulsion-the-comparison-of-power-between-a-car-and-an-aircraft/. [Accessed : January 26, 2022].

[69] Luis-Conti, “Udacity-Data-Scientist/Turbofan-Predictive-Maintenance/Turbofan Health Monitoring and Predictive Maintenance.ipynb,” [Online]. Available: https://github.com/Luis-Conti/Udacity-Data-Scientist/blob/main/Turbofan-Predictive-Maintenance/Turbofan%20health%20monitoring%20and%20predictive%20maintenance.ipynb. [Accessed : May 13, 2022].