

การทำนายอายุการใช้งานคงเหลือของเครื่องจักรด้วยเน็ตเวิร์กคอนโวลูชันเชิงลึกที่เพิ่มประสิทธิภาพ



นางสาวมนัสกานต์ เสน่หา

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2559

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Remaining Useful Life Prediction of Machine Using Enhanced Deep Convolutional
Neural Network

Miss Manassakan Sanayha



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Computer Science

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2016

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การทำนายอายุการใช้งานคงเหลือของเครื่องจักรด้วยเน็ต
เวิร์กคอนโวลูชันเชิงลึกที่เพิ่มประสิทธิภาพ

โดย

นางสาวมนัสกานต์ เสน่หา

สาขาวิชา

วิทยาศาสตร์คอมพิวเตอร์

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็นส่วน
หนึ่งของการศึกษาตามหลักสูตรปริญญาโทบริหารธุรกิจ

..... คณบดีคณะวิศวกรรมศาสตร์

(รองศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ

(ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(อาจารย์ ดร.พีรพล เวทีกุล)

..... กรรมการ

(รองศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัทธนะ)

..... กรรมการภายนอกมหาวิทยาลัย

(ดร.เดโช สุรางค์ศรีรัฐ)

มนัสกานต์ เสน่หา : การทำนายอายุการใช้งานคงเหลือของเครื่องจักรด้วยเน็ตเวิร์กคอนโวลูชันเชิงลึกที่เพิ่มประสิทธิภาพ (Remaining Useful Life Prediction of Machine Using Enhanced Deep Convolutional Neural Network) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: อ.ดร.พีรพล เวทีกุล, 73 หน้า.

ในการดำเนินการของโรงไฟฟ้าหรือโรงงาน เครื่องจักรและอุปกรณ์ต้องมีความน่าเชื่อถือพร้อมใช้งาน และทำงานได้อย่างต่อเนื่องตลอดเวลาเพื่อการผลิตได้อย่างมีประสิทธิภาพ การทำนายอายุการใช้งานคงเหลือ (Remaining Useful Life Prediction) ของอุปกรณ์ เป็นการดำเนินการในงานวางแผนบำรุงรักษาแบบทำนายล่วงหน้า (Predictive Maintenance) ซึ่งมีข้อดีคือ สามารถลดค่าใช้จ่ายของการบำรุงรักษาเชิงป้องกันและแก้ไข โดยวิธีการทำนายส่วนใหญ่ที่มีอยู่ในปัจจุบันจะเลือกวิเคราะห์ข้อมูลจุดวัดแต่ละจุดของอุปกรณ์แยกกัน ซึ่งไม่มีการพิจารณาถึงความสัมพันธ์ระหว่างจุดวัดกับจุดวัดและความสัมพันธ์ระหว่างจุดวัดกับช่วงเวลา ในวิทยานิพนธ์นี้ได้มีการประยุกต์ใช้โครงข่ายประสาทเทียมคอนโวลูชัน (Convolution Neural Network; CNN) ซึ่งเป็นการวิเคราะห์อนุกรมเวลาแบบหลายตัวแปร ทั้งในแง่ของความสัมพันธ์ระหว่างจุดวัดกับจุดวัดและความสัมพันธ์ของจุดวัดกับช่วงเวลา นอกจากนี้ยังได้นำเทคนิคต่าง ๆ มาใช้ในการปรับปรุงการเรียนรู้ของแบบจำลองที่นำเสนอให้มีประสิทธิภาพดีขึ้น ได้แก่ การป้องกันการอิงกับข้อมูลเรียนรู้มากเกินไป (Overfitting) ด้วยดรอปเอาต์ การทำเรกูลาร์ไรเซชันแบบแอลสอง (L^2 Regularization) เพื่อลดความซับซ้อนของแบบจำลอง และการปรับค่าน้ำหนักและอัตราการเรียนรู้ด้วยเกรเดียนต์ที่ปรับตัวได้ (Adaptive Gradient Method; AdaGrad)

สำหรับการทดลองได้แสดงให้เห็นถึงประสิทธิภาพโดยใช้ค่าความคลาดเคลื่อนของการทำนาย ได้แก่ ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error; RMSE) บนชุดข้อมูลที่มีคุณสมบัติแบบอนุกรมเวลาหลายตัวแปรที่มีการเผยแพร่โดยทั่วไปและชุดข้อมูลจากการดำเนินการจริงของอุปกรณ์ตัวอย่างในโรงไฟฟ้า

ภาควิชา วิศวกรรมคอมพิวเตอร์

ลายมือชื่อนิสิต

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ลายมือชื่อ อ.ที่ปรึกษาหลัก

ปีการศึกษา 2559

5870956621 : MAJOR COMPUTER SCIENCE

KEYWORDS: DEEP LEARNING / CONVOLUTIONAL NEURAL NETWORK / REMAINING USEFUL LIFE / MULTIVARIATE TIME SERIES ANALYSIS

MANASSAKAN SANAYHA: Remaining Useful Life Prediction of Machine Using Enhanced Deep Convolutional Neural Network. ADVISOR: PEERAPON VATEEKUL, Ph.D., 73 pp.

In the operation of power plants or factories, machinery and equipment must be reliable, ready to use and work continuously for efficient production. The Remaining Useful Life (RUL) prediction of equipment is performed in Predictive Maintenance (PdM) process which can reduce the cost of corrective and preventive maintenance. Most of existing models analyse sensor data separately; univariate analysis and rarely considers relationship between sensors and time. In this research, we applied a Convolutional Neural Network (CNN) which considers both dimensions, including time and sensors; multivariate time series analysis to predict the RUL of machine. In addition, many techniques have been applied to improve the learning of the proposed model in order to increase the performance, including dropout which prevent overfitting problem of model, model's complexity reduction with L^2 Regularization and the Adaptive Gradient Descent (AdaGrad) to adjust the weight and learning rate of model while training.

The performance was demonstrated in the experiments using the error of the prediction in term of the Root Mean Square Error (RMSE) on the datasets with the multivariate time series property on a standard benchmark datasets and dataset from the actual operation of sample equipment in power plants.

Department: Computer Engineering Student's Signature

Field of Study: Computer Science Advisor's Signature

Academic Year: 2016

กิตติกรรมประกาศ

ในการดำเนินการทำงานวิจัยและจัดทำวิทยานิพนธ์ฉบับนี้ ผู้วิจัยได้เผชิญกับอุปสรรคนานัปการ ทำให้ผู้จัดทำได้รับประสบการณ์ที่มีค่าเป็นอย่างยิ่งในการทำความเข้าใจและเรียนรู้ถึงการแก้ไขปัญหาและการฝึกฝนทักษะต่าง ๆ ในการทำงานวิจัย ซึ่งการดำเนินงานวิจัยให้แล้วเสร็จในครั้งนี้จะไม่สามารถทำให้สำเร็จได้หากขาดการสนับสนุนจากบุคคลสำคัญหลายฝ่าย ผู้วิจัยรู้สึกซาบซึ้งในความกรุณาเหล่านี้เป็นอย่างมาก จึงใคร่ขอใช้เนื้อหาในส่วนกิตติกรรมประกาศของวิทยานิพนธ์ฉบับนี้แสดงความขอบพระคุณเป็นอย่างสูงมา ณ ที่นี้

ขอบพระคุณอาจารย์ที่ปรึกษาวิทยานิพนธ์ ดร.พิรพล เวทีกุล ผู้อบรมสั่งสอน เสียสละและชี้แนะแนวทางในการทำวิทยานิพนธ์อย่างดีกับผู้วิจัย ทำให้งานวิจัยและวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้

ขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ อันประกอบไปด้วย ศาสตราจารย์ ดร.บุญเสริม กิจสิริกุล ผู้ช่วยศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัทธนะ และดร.เดโช สุรางค์ศรีรัฐ ที่ได้กรุณาให้เกียรติเป็นคณะกรรมการ ให้คำปรึกษา และข้อเสนอแนะอันเป็นประโยชน์อย่างมากต่อการทำวิจัยและการจัดทำวิทยานิพนธ์ฉบับนี้

ขอบคุณเพื่อน ๆ พี่ ๆ และน้อง ๆ สมาชิก “Data Mining Group, MIND Lab” ทุกท่านสำหรับคำปรึกษาและการแนะนำแนวทางในรายละเอียดต่าง ๆ เกี่ยวกับการทำนายด้วยเน็ตเวิร์กคอนโวลูชันเชิงลึกสำหรับข้อมูลอนุกรมเวลา รวมถึงการช่วยอำนวยความสะดวกในทุก ๆ ด้านระหว่างการจัดทำวิทยานิพนธ์ ซึ่งถือเป็นอีกปัจจัยหนึ่งที่ทำให้การทำงานสำเร็จลุล่วงไปได้ด้วยดี

ขอบพระคุณเจ้าหน้าที่ในภาควิชาวิศวกรรมคอมพิวเตอร์ทุกท่าน ที่ได้ช่วยประสานงานให้คำแนะนำในเรื่องเอกสารและการดำเนินการในขั้นตอนต่าง ๆ แก่ผู้วิจัยในระหว่างการทำวิจัยเป็นอย่างดี

สุดท้ายนี้ขอขอบพระคุณครอบครัวและบุคคลอันเป็นที่รักของผู้วิจัยที่ให้การสนับสนุนในทุก ๆ ด้าน และให้กำลังใจตลอดระยะเวลาในการดำเนินการทำงานวิจัยนี้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
บทที่ 1 ที่มาและความสำคัญ.....	4
1.1 ที่มาและความสำคัญ.....	4
1.2 วัตถุประสงค์	6
1.3 ขอบเขตการดำเนินงาน	6
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	6
1.5 ขั้นตอนการดำเนินงานวิจัย.....	7
1.6 สรุปผลภาพรวมของการทำงานวิจัย	7
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	9
2.1 ทฤษฎีที่เกี่ยวข้อง	9
2.1.1 โครงข่ายประสาทเทียม (Artificial Neural Network).....	9
2.1.2 ฟังก์ชันกระตุ้น (Activation Function).....	12
2.1.3 ฟังก์ชันต้นทุน (Cost Function หรือ Loss Function หรือ Objective Function)	13
2.1.4 การหาค่าเหมาะที่สุด (Optimization).....	13
2.1.5 การดรอปเอาต์ (Dropout).....	14
2.1.6 โครงข่ายประสาทเทียมเชิงลึก (Deep Neural Network).....	15
2.1.7 การทำเรกูลาร์ไรเซชัน (Regularization).....	20
2.1.8 การพยากรณ์อนุกรมเวลาด้วยวิธีอาร์มา (ARIMA).....	21

2.1.9 การพยากรณ์อนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โพเนนเชียล (Exponential smoothing).....	23
2.1.10 การวัดประสิทธิภาพ (Performance Evaluation).....	23
2.2 งานวิจัยที่เกี่ยวข้อง.....	24
2.2.1 งานวิจัยเกี่ยวกับการบำรุงรักษาแบบพยากรณ์ล่วงหน้า.....	24
2.2.2 งานวิจัยเกี่ยวกับการวิเคราะห์ข้อมูลแบบอนุกรมเวลา.....	25
2.2.3 งานวิจัยเกี่ยวกับการทำนายอายุการใช้งานคงเหลือของเครื่องจักร.....	25
บทที่ 3 แบบจำลองการทำนายอายุการใช้งานคงเหลือของเครื่องจักรที่เพิ่มประสิทธิภาพ.....	27
3.1 แบบจำลองการทำนายอายุการใช้งานคงเหลือ.....	27
3.2 การเพิ่มประสิทธิภาพให้กับแบบจำลอง.....	28
3.3 การเตรียมข้อมูลด้วยการประมวลผลก่อน.....	36
3.4 การทำนายอายุการใช้งานคงเหลือ.....	37
3.4.1 การทำนายแบบตัวแปรเดียว (Univariate Forecasting).....	37
3.4.2 การทำนายแบบหลายตัวแปร (Multivariate Forecasting) ด้วยการใช้โครงข่ายประสาทเทียมคอนโวลูชันเชิงลึก.....	38
บทที่ 4 การทดลองและผลการทดลอง.....	40
4.1 ชุดข้อมูลที่ใช้ในการทดลอง.....	40
4.2 การดำเนินการทดลอง.....	48
4.2.1 การสอน (Training).....	48
4.2.2 การทดสอบ (Testing).....	48
4.2.3 การกำหนดค่าต่าง ๆ ของแบบจำลอง (Model Configuration).....	49
4.3 ผลการทดลอง.....	51
4.4 การวิเคราะห์ผลการทดลอง.....	53
บทที่ 5 สรุปการวิจัยและแนวทางการวิจัยในขั้นถัดไป.....	54

5.1 สรุปการวิจัย.....	54
5.2 ข้อจำกัดของงานวิจัย.....	55
5.3 แนวทางการวิจัยในอนาคต.....	56
5.4 ผลงานตีพิมพ์จากงานวิจัย.....	56
ดัชนีศัพท์.....	57
รายการอ้างอิง.....	61
ภาคผนวก ก รายละเอียดของข้อมูลเทอร์โบแฟนชุด C-MAPSS.....	64
ภาคผนวก ข ผลการทดลองค่า RMSE ในรูปแบบกราฟแท่งของชุดข้อมูลทั้ง 14 ชุดเมื่อทดลอง กับแบบจำลองทั้งเก้าแบบ.....	70
ภาคผนวก ค ผลการทดสอบการเปรียบเทียบประสิทธิภาพของแบบจำลองทั้งเก้าแบบด้วยวิธี Paired T Test.....	72
ประวัติผู้เขียนวิทยานิพนธ์.....	73

สารบัญรูปภาพ

	หน้า
รูปที่ 1 แสดงโครงสร้างของเพอร์เซ็ปตรอน โดยมีข้อมูลนำเข้า ค่าน้ำหนัก (Weights).....	9
รูปที่ 2 แสดงโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า.....	10
รูปที่ 3 เปรียบเทียบการดรอปเอาต์ในโครงข่ายประสาทเทียม.....	15
รูปที่ 4 โครงข่ายประสาทเทียมคอนโวลูชันแบบ LeNet-5.....	16
รูปที่ 5 แสดงขั้นตอนการทำคอนโวลูชันระหว่างภาพนำเข้า.....	16
รูปที่ 6 เปรียบเทียบการทำคอนโวลูชันแบบแคบและแบบกว้าง.....	17
รูปที่ 7 การทำคอนโวลูชันโดยมีข้อมูลรับเข้าขนาด 4x4 ตัวกรองขนาด 2x2	18
รูปที่ 8 การทำคอนโวลูชันโดยมีจำนวนตัวกรองเท่ากับ 3	18
รูปที่ 9 แสดงการพูลลิงแบบหาค่าเฉลี่ยและหาค่ามากที่สุด.....	19
รูปที่ 10 แสดงโครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก [10]	26
รูปที่ 11 แสดงโครงสร้างการทำนายอายุการใช้งานคงเหลือด้วยโครงข่ายประสาทเทียมคอนโวลูชันเชิงลึกที่มีข้อมูลเข้า 24 จุดวัด ทีละ 15 ช่วงเวลา.....	27
รูปที่ 12 ความแตกต่างของโครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึกแบบต่าง ๆ.....	28
รูปที่ 13 เปรียบเทียบประสิทธิภาพจากการทำนายด้วยแบบจำลองที่ใช้ฟังก์ชันกระตุ้นซิกมอยด์ เรคตีไฟเชิงเส้น และเอกโพเนนเชียลเชิงเส้น 100 รอบการเรียนรู้แรกบนชุดข้อมูล C-MAPSS3.....	30
รูปที่ 14 เปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองด้วยการใช้เทคนิคการดรอปเอาต์ 100 รอบการเรียนรู้แรกบนชุดข้อมูล C-MAPSS3	31
รูปที่ 15 เปรียบเทียบประสิทธิภาพการทำนายโดยการใช้เทคนิคการทำเรกูลาร์ไรเซชันแบบแอลสอง 100 รอบการเรียนรู้แรกบนชุดข้อมูล C-MAPSS3	32
รูปที่ 16 เปรียบเทียบประสิทธิภาพของแบบจำลองเมื่อใช้สโตแคสติกเกรเดียนต์เดสเซนท์และเกรเดียนต์ที่ปรับตัวในการทำนายอายุการใช้งานคงเหลือของชุดข้อมูล C-MAPSS1.....	33
รูปที่ 17 แสดงโครงสร้างการทำนายภาพรวมประกอบด้วยเน็ตเวิร์กคอนโวลูชันเชิงลึกแบบ VGG-16 และเทคนิคในการเพิ่มประสิทธิภาพ.....	34

รูปที่ 18 กระบวนการทำนายอายุการใช้งานคงเหลือแบบตัวแปรเดียว	38
รูปที่ 19 กระบวนการทำงานของแบบจำลองการทำนายอายุการใช้งานคงเหลือของเครื่องจักร ด้วยโครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึก.....	39
รูปที่ 20 ส่วนประกอบของเทอร์โบแฟน	40
รูปที่ 21 ตัวอย่างการขยายตัวของรอยแตกในสภาพการรับน้ำหนักเดียวกัน	43
รูปที่ 22 โครงสร้างส่วนประกอบ CWP	43
รูปที่ 23 ค่า RMSE ในรูปแบบกราฟแท่งจากการทดลองด้วยแบบจำลองเก้าแบบกับชุดข้อมูลด้าน การทำนายอายุการใช้งานคงเหลือ.....	70
รูปที่ 24 ค่า RMSE ในรูปแบบกราฟแท่งจากการทดลองด้วยแบบจำลองเก้าแบบกับชุดข้อมูล อนุกรมเวลาหลายตัวแปร	71



สารบัญตาราง

	หน้า
ตารางที่ 1 แผนการดำเนินงาน	8
ตารางที่ 2 แสดงประเภทของฟังก์ชันกระตุ้น ค่าผลลัพธ์ และสมการ	12
ตารางที่ 3 ความแตกต่างระหว่างเรกูลาร์ไรเซชันแบบแอลหนึ่งและแอลสอง	21
ตารางที่ 4 รหัสเทียมหรือ Pseudocode ของโครงสร้างเน็ตเวิร์กคอนโวลูชันที่เพิ่มประสิทธิภาพ ...	35
ตารางที่ 5 แสดงชื่อแอตทริบิวต์และคำอธิบายของชุดข้อมูลเทอร์โบแฟน	41
ตารางที่ 6 แสดงชื่อตัวแปรและคำอธิบายจุดวัดเพิ่มเติมของชุดข้อมูลเทอร์โบแฟน [11].....	41
ตารางที่ 7 แสดงจำนวนเครื่องเทอร์โบแฟนที่ใช้ในการจำลองการทำงานของแต่ละชุดข้อมูลย่อย	42
ตารางที่ 8 ข้อมูลประวัติเหตุการณ์ ช่วงเวลา และรายละเอียดที่อุปกรณ์ถูกตรวจสอบ ซ่อมแซม หรือมีการทำงานผิดปกติในปี 2558.....	44
ตารางที่ 9 แสดงแอตทริบิวต์ หน่วย และช่วงของข้อมูล CCPP	45
ตารางที่ 10 แสดงชื่อแอตทริบิวต์คำอธิบาย และหน่วยการวัดของชุดข้อมูลการใช้พลังงาน	45
ตารางที่ 11 แสดงแอตทริบิวต์ คำอธิบาย และหน่วยการวัดของชุดข้อมูล Beijing.....	47
ตารางที่ 12 แสดงจำนวนแอตทริบิวต์ ข้อมูลสอน และข้อมูลทดสอบของแต่ละชุดข้อมูล	48
ตารางที่ 13 แสดงขนาดและมิติที่ใช้ในโครงสร้างของแบบจำลองแบบต่าง ๆ	50
ตารางที่ 14 เทคนิคที่ใช้ในแต่ละการทดลองการทำนายด้วยโครงข่ายประสาทเทียม.....	51
ตารางที่ 15 ค่า RMSE จากการทดลองแบบจำลองกับชุดข้อมูลด้านการทำนายอายุการใช้งาน คังเหลื่อ.....	52
ตารางที่ 16 ค่า RMSE จากการทดลองแบบจำลองกับชุดข้อมูลอนุกรมเวลาหลายตัวแปร.....	52
ตารางที่ 17 แสดงค่า T Value จากการทำการทดสอบ Pair T Test	72
ตารางที่ 18 แสดงค่า P Value จากการทำการทดสอบ Pair T Test	72

บทที่ 1

ที่มาและความสำคัญ

1.1 ที่มาและความสำคัญ

การบำรุงรักษาเป็นกิจกรรมที่สำคัญที่สุดอย่างหนึ่งในการปฏิบัติงานทางด้านการผลิตไฟฟ้า เนื่องจากอุปกรณ์ต้องอยู่ในสภาพพร้อมใช้งานจึงสามารถดำเนินการผลิตไฟฟ้าได้อย่างมีประสิทธิภาพ ประเภทของการบำรุงรักษา มี 3 ประเภท ได้แก่ การบำรุงรักษาเชิงแก้ไข (Corrective Maintenance) จะดำเนินการซ่อมแซมเมื่อพบว่าอุปกรณ์เสียหาย การบำรุงรักษาเชิงป้องกัน (Preventive Maintenance) เป็นกิจกรรมการบำรุงรักษาที่ทำตามรอบ เพื่อลดความถี่หรือจำนวนของอุปกรณ์ที่เสียหาย นอกจากนี้ยังมีการบำรุงรักษาแบบพยากรณ์ล่วงหน้า (Predictive Maintenance) ที่จะคาดคะเน ทำนาย หรือพยากรณ์อาการเสียของอุปกรณ์ล่วงหน้า เพื่อสามารถวางแผนการบำรุงรักษาในอนาคต แตกต่างจากการบำรุงรักษาเชิงป้องกันตรงที่การบำรุงรักษาอาจมาจากประสบการณ์ หรือตามคู่มือการใช้งานอุปกรณ์ ในขณะที่การบำรุงรักษาแบบพยากรณ์ล่วงหน้าอาศัยผลการวัดค่าสถานะต่าง ๆ ของอุปกรณ์ ไม่สามารถใช้ระบบประสาทสัมผัสของบุคคลได้ เนื่องจากประสาทสัมผัสแต่ละคนนั้นมีความแตกต่างกัน ส่งผลให้เกิดความคลาดเคลื่อนได้

สิ่งที่นิยมตรวจวัดเพื่อนำมาทำนายอาการเสียของอุปกรณ์ ได้แก่ สัญญาณความสั่นสะเทือน (Vibration Signal) การวิเคราะห์สารหล่อลื่น (Oil Wear Particle Analysis) การวิเคราะห์ค่าต่าง ๆ เช่น อุณหภูมิ กระแสไฟฟ้า แรงดัน จากจุดวัด (Measuring Point) ที่ตัวอุปกรณ์ ซึ่งการที่จะเลือกวิธีใดวิธีหนึ่ง หรือหลายวิธีมาใช้งาน ต้องขึ้นอยู่กับชนิดและความเสี่ยงของเครื่องจักรที่จะเกิด ความเสียหาย

เนื่องจากกิจการที่มีเครื่องจักรซับซ้อน เช่น โรงไฟฟ้า ต้องการความน่าเชื่อถือในกระบวนการผลิตสูง และเครื่องจักรมีการทำงานอย่างต่อเนื่อง การบำรุงรักษาแบบพยากรณ์ล่วงหน้าจึงเริ่มเป็นที่นิยม เนื่องจากความต้องการในการทำนายอาการเสีย หรือความผิดปกติล่วงหน้า เพื่อจะได้ดำเนินการวางแผนแรงงาน ชิ้นส่วนที่ต้องใช้ และแผนการบำรุงรักษาล่วงหน้า อีกทั้งยังเป็นการลดค่าใช้จ่ายที่เกิดขึ้นจากการบำรุงรักษาเชิงแก้ไข ซึ่งเครื่องจักรเกิดการเสียหายแบบฉับพลัน และค่าใช้จ่ายที่ไม่จำเป็นในการบำรุงรักษาเชิงป้องกัน ซึ่งอาจดำเนินการทั้ง ๆ ที่เครื่องจักรยังไม่เกิดความเสียหาย ทำให้เสียค่าใช้จ่ายโดยเปล่าประโยชน์

ในการทำนายอายุการใช้งานคงเหลือ (Remaining Useful Life; RUL) ของอุปกรณ์ ไม่ว่าจะ เป็นล่วงหน้า 1 ชั่วโมง 12 ชั่วโมง 1 วัน 2 วันก็ตาม ต้องอาศัยข้อมูลจากผู้เชี่ยวชาญในอุปกรณ์นั้น ๆ

ข้อมูลประวัติ (Historical Data) และข้อมูลนำเข้า (Input Data) อันได้แก่ ตัวเลขค่าต่าง ๆ จากจุดวัดที่ติดอยู่กับตัวอุปกรณ์ มาทำการวิเคราะห์โดยใช้วิธีการทางสถิติ เช่น Principle Component Analysis (PCA), Auto Regression Integrated Moving Average (ARIMA) เป็นต้น ร่วมกับการทำเหมืองข้อมูล (Data Mining) ซึ่งสามารถทำได้หลายวิธี เช่น การเรียนรู้ของเครื่อง (Machine Learning) โครงข่ายประสาทเทียม (Neural Network) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine; SVM) วิธีการเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor; KNN) การแบ่งกลุ่มข้อมูลแบบเคมีน (K-Means Clustering) การถดถอยเชิงเส้น (Linear Regression) กฎความสัมพันธ์ (Association Rules) เป็นต้น โดยในอดีตนิยมทำการทำนายข้อมูลอนุกรมเวลาด้วยการวิเคราะห์แยกกันแบบตัวแปรเดียว (Univariate Time-Series Analysis) ที่ไม่มีการวิเคราะห์ถึงความสัมพันธ์ระหว่างตัวแปร ซึ่งเป็นปัจจัยสำคัญที่ส่งผลต่อข้อมูลส่งออกโดยตรง ทำให้ค่าที่ทำนายได้ไม่แม่นยำเท่าที่ควร

การทำนายด้วยโครงข่ายประสาทเทียมคอนโวลูชันเชิงลึก (Deep Convolution Neural Network) เป็นวิธีการที่นิยมกันอย่างแพร่หลายในการวิเคราะห์เพื่อจัดกลุ่มข้อมูลรูปภาพ (Image Classification) เนื่องจากสามารถดึงคุณลักษณะที่สำคัญ (Feature Extraction) และความสัมพันธ์ระหว่างตัวแปร หรือสามารถเรียกได้ว่าการวิเคราะห์ข้อมูลอนุกรมเวลาแบบหลายตัวแปร (Multivariate Time-Series Analysis) ได้ ทำให้ค่าที่ต้องการทำนายมีความแม่นยำมากกว่าวิธีการวิเคราะห์แบบเดิมที่ไม่มีการปรับเปลี่ยน หรือนำเทคนิคอื่น ๆ มาประยุกต์ใช้ อีกทั้งไม่มีการเปรียบเทียบประสิทธิภาพกับโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันเชิงลึกแบบเดิมที่มีอยู่อีกด้วย

โดยในวิทยานิพนธ์นี้นำเสนอโครงสร้างที่เหมาะสมที่สุดสำหรับการทำนายอายุการใช้งานคงเหลือของเครื่องจักร ซึ่งได้มีการนำเทคนิคต่าง ๆ มาประยุกต์ใช้และปรับเปลี่ยนโครงสร้างโครงข่ายประสาทเทียมให้เหมาะสมกับข้อมูลนำเข้าจากเดิมที่นิยมใช้กับรูปภาพมาเป็นใช้กับชุดข้อมูลที่เป็นค่าการทำงานต่าง ๆ ของเครื่องจักรซึ่งมีลักษณะเป็นข้อมูลต่อเนื่อง (Time-Series Data) ที่เป็นที่ยอมรับและน่าเชื่อถือ เพื่อให้ได้ประสิทธิภาพสูงและค่าความคลาดเคลื่อนของการทำนายลดลงเมื่อเทียบกับวิธีการและโครงสร้างแบบเดิม เพื่อการนำไปศึกษาและพัฒนาต่อไป

1.2 วัตถุประสงค์

เพื่อนำเสนอวิธีการในการทำนายอายุการใช้งานคงเหลือของเครื่องจักรโดยใช้โครงข่ายประสาทเทียมคอนโวลูชันเชิงลึกที่เพิ่มประสิทธิภาพให้มีความแม่นยำมากขึ้นกว่าวิธีการเดิมที่ยังไม่มีการปรับโครงสร้างให้เหมาะสมกับการทำนายข้อมูลอนุกรมเวลาแบบหลายตัวแปรและค่าที่ต้องการทำนาย

เพื่อสนับสนุนงานทางการบำรุงรักษาเชิงป้องกันในการเตรียมความพร้อมสำหรับการดำเนินการตรวจสอบอุปกรณ์ หรือเครื่องจักรที่อาจมีแนวโน้มที่จะเกิดอาการเสีย หรือทำงานผิดปกติได้ในอนาคต

1.3 ขอบเขตการดำเนินงาน

1. ทำนายอายุการใช้งานคงเหลือของเครื่องจักรที่มีข้อมูลนำเข้าเป็นอนุกรมเวลาแบบหลายตัวแปรด้วยเน็ตเวิร์กคอนโวลูชันเชิงลึก
2. เปรียบเทียบประสิทธิภาพเมื่อมีการปรับโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันเชิงลึกใหม่กับโครงสร้างแบบเดิม ได้แก่ LeNet-5 VGG-16 และ AlexNet
3. เปรียบเทียบประสิทธิภาพด้านโครงสร้างระหว่างวิธีการแบบเดิมกับวิธีการใหม่ที่มีการใช้เทคนิคต่าง ๆ เช่น การดรอปเอาต์ การทำเรกูลาร์ไรเซชันแบบแอลสอง วิธีการหาค่าเหมาะที่สุดในการปรับปรุงน้ำหนักของเส้นเชื่อม เป็นต้น

1.4 ประโยชน์ที่คาดว่าจะได้รับ

การทำนายอายุการใช้งานคงเหลือของเครื่องจักรสามารถช่วยผู้ปฏิบัติงานในการบำรุงรักษาอุปกรณ์ เนื่องจากทราบอายุการใช้งานคงเหลือล่วงหน้าจากแบบจำลอง ทำให้ดำเนินการตรวจสอบหรือซ่อมบำรุงได้ก่อนที่อุปกรณ์จะเสียหาย ซึ่งมีผลกระทบต่อระบบผลิตกระแสไฟฟ้าทั้งในด้านความปลอดภัย การดำเนินการผลิต และค่าใช้จ่าย

1.5 ขั้นตอนการดำเนินงานวิจัย

ในการดำเนินงานวิจัยนี้มีขั้นตอนและระยะเวลาในแต่ละขั้นตอน ดังตารางที่ 1

1. ศึกษางานวิจัยและทฤษฎีที่เกี่ยวข้อง
 - 1.1 ศึกษางานวิจัยการบำรุงรักษาแบบพยากรณ์ล่วงหน้า
 - 1.2 ศึกษางานวิจัยการวิเคราะห์ข้อมูลอนุกรมเวลา
 - 1.3 ศึกษางานวิจัยการทำนายอายุการใช้งานคงเหลือของเครื่องจักร
 - 1.4 ศึกษาขั้นตอนวิธีและเทคนิคโครงข่ายประสาทคอนโวลูชัน
 - 1.5 ศึกษาเครื่องมือที่ใช้ในงานวิจัยได้แก่ R Python
2. ออกแบบการทดลอง
3. สร้างแบบจำลอง
4. ทำการทดลอง
5. วิเคราะห์ผลการทดลอง
6. สรุปผลและเรียบเรียงวิทยานิพนธ์
7. สอบวิทยานิพนธ์

1.6 สรุปผลภาพรวมของการทำงานวิจัย

จากวัตถุประสงค์และขอบเขตการดำเนินงานทั้งหมด ผู้วิจัยได้ทำการทดลองและดำเนินการสำเร็จเป็นที่เรียบร้อยแล้ว โดยได้สรุปรายการทั้งหมดไว้ในบทที่ 5

ตารางที่ 1 แผนการดำเนินงาน

ลำดับ	ขั้นตอนงานวิจัย	2559							2560								
		ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.			
1	ศึกษางานวิจัยและทฤษฎีที่เกี่ยวข้อง																
1.1	ศึกษางานวิจัยการบำรุงรักษาเชิงป้องกัน																
1.2	ศึกษางานวิจัยการทำเหมืองข้อมูลกับข้อมูลอนุกรมเวลา																
1.3	ศึกษาขั้นตอนวิธีและเทคโนโลยีโครงข่ายประสาทคอนโวลูชัน																
1.4	ศึกษาเครื่องมือที่ใช้ในงานวิจัยได้แก่ R Python																
2	ออกแบบการทดลอง																
3	สร้างแบบจำลอง																
4	ทำการทดลอง																
5	วิเคราะห์ผลการทดลอง																
6	สรุปผลและเรียบเรียงวิทยานิพนธ์																
7	สอบวิทยานิพนธ์																

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

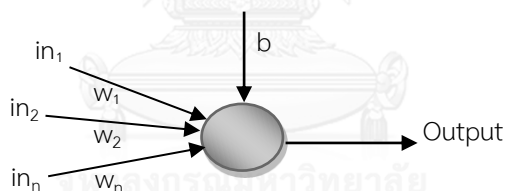
2.1 ทฤษฎีที่เกี่ยวข้อง

ในงานวิจัยนี้มีทฤษฎีที่เกี่ยวข้องได้แก่ โครงข่ายประสาทเทียม โครงข่ายประสาทเทียมแบบคอนโวลูชัน และการคำนวณความแม่นยำ

2.1.1 โครงข่ายประสาทเทียม (Artificial Neural Network)

เป็นระบบที่พัฒนาขึ้นโดยใช้แบบจำลองทางคณิตศาสตร์เลียนแบบการทำงานของสมอง จำลองให้เหมือนกระบวนการเรียนรู้ของมนุษย์ เพื่อให้สามารถเรียนรู้การจดจำรูปแบบและสร้างความรู้ใหม่ได้เช่นเดียวกับสมองมนุษย์ หลักการทำงานคือ เมื่อข้อมูลนำเข้ามายังโครงข่ายประสาทเทียมจะถูกส่งไปที่ชั้นซ่อน (Hidden Layer) เพื่อประมวลผลออกมาเป็นเอาต์พุต

โครงข่ายประสาทเทียมมีหน่วยพื้นฐานที่เล็กที่สุดเรียกว่า เพอร์เซ็ปตรอน (Perceptron) มีโครงสร้างดังรูปที่ 1 ทำหน้าที่แยกเอาต์พุตแต่ละกลุ่มออกจากกัน และมีฟังก์ชันการทำงาน ดังแสดงในสมการที่ 1



รูปที่ 1 แสดงโครงสร้างของเพอร์เซ็ปตรอน โดยมีข้อมูลนำเข้า ค่าน้ำหนัก (Weights)

$$f(x) = \sum_{i=1}^n w_i x_i + b \quad (1)$$

โดย w แทนค่าน้ำหนัก

b แทนค่าความเอนเอียง

n แทนจำนวนของข้อมูลนำเข้า

นอกจากนี้เพอร์เซ็ปตรอนยังมีกระบวนการเรียนรู้ ดังสมการที่ 2 และ 3 เป็นตัวกำหนด การเปลี่ยนแปลงน้ำหนักในแต่ละรอบของการเรียนรู้ เทียบกับอัตราส่วนของผลต่างของข้อมูลส่งออก

$$w_i \leftarrow w_i + \Delta w_i \quad (2)$$

$$\Delta w_i = \alpha(\hat{y} - y)x_i \quad (3)$$

โดย x แทนชุดข้อมูลตัวอย่าง

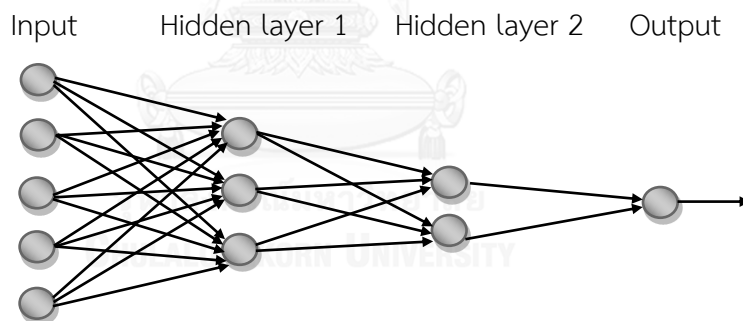
\hat{y} แทนค่าข้อมูลส่งออกที่ทำนายได้

y แทนค่าข้อมูลส่งออกจริง แทนด้วย

α แทนอัตราการเรียนรู้ (learning rate)

1. โครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า (Feedforward Neural Network)

เป็นโครงข่ายประสาทเทียมที่มีการคำนวณและส่งผ่านของข้อมูลไปในทิศทางเดียวและเป็นลำดับ มีโครงสร้างเป็นลำดับชั้น และในแต่ละลำดับชั้น มีเพอร์เซ็ปตรอนจำนวนหนึ่งซึ่งเพอร์เซ็ปตรอนในชั้นเดียวกันจะไม่เชื่อมถึงกัน แต่จะเชื่อมกับเพอร์เซ็ปตรอนที่อยู่ในลำดับชั้นติดกันทั้งหมด และข้อมูลส่งออกของเพอร์เซ็ปตรอนชั้นก่อนหน้า จะเป็นข้อมูลนำเข้าของเพอร์เซ็ปตรอนในชั้นปัจจุบัน โดยมีโครงสร้างดังรูปที่ 2



รูปที่ 2 แสดงโครงข่ายประสาทเทียมแบบป้อนไปข้างหน้า

นอกจากนี้สำหรับการคำนวณข้อมูลนำเข้าของเพอร์เซ็ปตรอนในชั้นถัดไป จะมีฟังก์ชันกระตุ้น ดังสมการที่ 4 และ 5

$$z_i^j = \sum_{k=1}^n w_{ik}^j x_k^{j-1} + b_i^j \quad (4)$$

$$x_i^j = g(y_i^j) \quad (5)$$

โดย g แทนฟังก์ชันกระตุ้น

n แทนจำนวนเพอร์เซ็ปตรอนในลำดับชั้นที่ $j-1$

2. การแพร่กระจายย้อนกลับและการสอน (Backpropagation and Training)

โครงข่ายประสาทเทียมแบบไปข้างหน้าสามารถหาค่าความคลาดเคลื่อนของเพอร์เซ็ปตรอนในชั้นสุดท้ายได้ง่ายจากการคำนวณเกรเดียนของฟังก์ชันต้นทุนเทียบกับค่าผลลัพธ์ของชั้นสุดท้าย แต่การหาค่าความคลาดเคลื่อนของเพอร์เซ็ปตรอนของชั้นก่อนหน้า เพื่อใช้ในการเรียนรู้ นั้นไม่สามารถทำได้โดยตรง จึงมีการใช้วิธีการแพร่กระจายย้อนกลับ ดังสมการที่ 6

$$\delta_j^l = \frac{\partial J}{\partial z_j^l} = \frac{\partial J}{\partial a_j^l} \frac{\partial a_j^l}{\partial z_j^l} = \frac{\partial J}{\partial a_j^l} g'(z_j^l) \quad (6)$$

โดย δ_j^l แทนค่าความคลาดเคลื่อนของเพอร์เซ็ปตรอนตัวที่ j ในลำดับชั้น l
 j แทนฟังก์ชันต้นทุน
 z แทนค่าที่คำนวณได้ก่อนจะผ่านฟังก์ชันกระตุ้น g

ในการหาค่า $\frac{\partial J}{\partial a_j^l}$ ในลำดับชั้นสุดท้ายสามารถคำนวณได้โดยตรงจากฟังก์ชันต้นทุนที่เลือกใช้ ส่วนในลำดับชั้นก่อนหน้าต้องหาโดยวิธีการแพร่กระจายย้อนกลับ โดยคำนวณคล้ายกับการป้อนไปข้างหน้า แต่กลับทิศ ดังสมการที่ 7

$$\frac{\partial J}{\partial a_j^l} = \sum_{k=1}^m \frac{\partial J}{\partial a_k^{l+1}} \frac{\partial a_k^{l+1}}{\partial a_j^l} = \sum_{k=1}^m \delta_k^{l+1} w_{kj}^{l+1} \quad (7)$$

โดย m แทนจำนวนเพอร์เซ็ปตรอนในลำดับชั้นที่ $l + 1$

จากสมการที่ 7 จะสามารถหาค่าความคลาดเคลื่อนของแต่ละระดับชั้นได้ ทำให้สามารถคำนวณค่าความคลาดเคลื่อนเทียบกับน้ำหนัก และค่าความเอนเอียงใด ๆ ได้ดังสมการที่ 8 และ 9

$$\frac{\partial J}{\partial w_{jk}^l} = \frac{\partial J}{\partial z_j^l} \frac{\partial z_j^l}{\partial w_{jk}^l} = \delta_j^l a_k^{l-1} \quad (8)$$

$$\frac{\partial J}{\partial b_j^l} = \frac{\partial J}{\partial z_j^l} \frac{\partial z_j^l}{\partial b_j^l} = \delta_j^l \quad (9)$$

เนื่องจากวิทยานิพนธ์นี้ใช้สโตแคสติกเกรเดียนเดสเซนท์ จึงใช้การปรับปรุงค่าน้ำหนัก w_{jk}^l ดังสมการที่ 10

$$w_{jk}^l = w_{jk,t-1}^l - \alpha a_{k,t}^{l-1} \delta_{j,t}^l \quad (10)$$

2.1.2 ฟังก์ชันกระตุ้น (Activation Function)

ข้อมูลส่งออกของแต่ละเพอร์เซ็ปตรอน จะมีการใช้ฟังก์ชันกระตุ้น $f(x)$ ที่ช่วยให้โครงข่ายประสาทเทียมสามารถแก้ปัญหาได้หลากหลายมากขึ้น ฟังก์ชันกระตุ้นที่นิยมกันมีดังรายละเอียดตามตารางที่ 2

ตารางที่ 2 แสดงประเภทของฟังก์ชันกระตุ้น ค่าผลลัพธ์ และสมการ

ลำดับที่	ฟังก์ชัน	ค่าผลลัพธ์	สมการของฟังก์ชัน
1	ฟังก์ชันซิกมอยด์ (Sigmoid Function)	0 ถึง 1	$\sigma(x) = \frac{1}{1+e^{-x}}$ (6)
2	ฟังก์ชันแทนเจนต์ไฮเพอร์โบลิก (Hyperbolic Tangent Function)	-1 ถึง 1	$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ (7)
3	ฟังก์ชันค่าสูงสุดอย่างอ่อน (Softmax Function)	0 ถึง 1	$f(x)_j = \frac{e^{x_j}}{\sum_{i=1}^K e^{x_i}}$ (8)
4	ฟังก์ชันเรคตีไฟด์เชิงเส้น (Rectified Linear Unit Function; ReLU)	1^+ หรือ 0	$f(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases}$ (9)
5	ฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้น (Exponential Linear Unit Function; ELU)	$(-\alpha, \infty)$	$f(\alpha, x) = \begin{cases} \alpha(e^x - 1), & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases}$ (10)

สำหรับฟังก์ชันค่าสูงสุดอย่างอ่อนในชั้นของโครงข่ายประสาทเทียมที่ต้องการคำนวณมีผลลัพธ์ทั้งหมด K ตัว ค่าของผลลัพธ์จะแทนด้วยสัญลักษณ์ x กล่าวคือ ฟังก์ชันค่าสูงสุดอย่างอ่อนของผลลัพธ์ตัวที่ j หรือแทนด้วยสัญลักษณ์ f_j

2.1.3 ฟังก์ชันต้นทุน (Cost Function หรือ Loss Function หรือ Objective Function)

เป็นฟังก์ชันของโครงข่ายประสาทเทียมที่แสดงถึงต้นทุนของโครงข่ายประสาทเทียม โดยในการเรียนรู้ของโครงข่ายประสาทเทียมนั้น จะมีการปรับค่าน้ำหนักเพื่อที่จะลดค่าผลลัพธ์ของฟังก์ชันต้นทุน สำหรับฟังก์ชันต้นทุนที่นิยมใช้คือ ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error; MSE) สามารถคำนวณได้จากสมการที่ 11

$$J = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (11)$$

โดย J แทนฟังก์ชันต้นทุน

n แทนจำนวนข้อมูลทั้งหมดที่ใช้ในการเรียนรู้

y_i แทนผลลัพธ์จริงที่ต้องการของข้อมูลชุดที่ i

\hat{y}_i แทนผลลัพธ์ที่ทำนายได้ของข้อมูลชุดที่ i

2.1.4 การหาค่าที่เหมาะสมที่สุด (Optimization)

จุดประสงค์ของการเรียนรู้ของโครงข่ายประสาทเทียมคือ เพื่อลดค่าของฟังก์ชันต้นทุนให้น้อยที่สุดด้วยการปรับปรุงน้ำหนักของเส้นเชื่อมในโครงข่ายประสาทเทียม สำหรับวิธีการที่ใช้ในการปรับปรุงน้ำหนักของเส้นเชื่อมที่ได้รับความนิยมดังต่อไปนี้

1. สโตแคสติกเกรเดียนเตสเซนซ์ (Stochastic Gradient Descent; SGD)

การเรียนรู้ด้วยสโตแคสติกเกรเดียนเตสเซนซ์ มีการปรับค่าของน้ำหนัก โดยนำค่าน้ำหนักก่อนหน้ามาคำนวณดังสมการที่ 12

$$w_t = w_{t-1} - \alpha \frac{\partial J_t}{\partial w} \quad (12)$$

โดย w แทนพารามิเตอร์ซึ่งเป็นน้ำหนักที่ต้องการจะปรับค่า

α แทนอัตราการเรียนรู้

$\frac{\partial J}{\partial w}$ แทนเกรเดียนของฟังก์ชันต้นทุนเทียบกับ w

การที่จะทำให้การเรียนรู้มีการลู่เข้าดีขึ้นและหลีกเลี่ยงการติดอยู่ที่โลคอลลอพติมา (Local Optima) ซึ่งเป็นค่าที่เหมาะสมที่สุดในช่วง ๆ หนึ่งแต่ไม่ใช่ทั้งหมด สามารถใช้โมเมนตัม (Momentum) แก้ปัญหาดังกล่าว ดังสมการที่ 13 และ 14

$$v_t = \gamma v_{t-1} + \alpha \frac{\partial J_t}{\partial w} \quad (13)$$

$$w_t = w_{t-1} - v_t \quad (14)$$

โดย v แทนค่าความเร็วซึ่งมีการปรับค่าพร้อมกับ w

γ แทนค่าสัมประสิทธิ์ของโมเมนตัม (Momentum Coefficient)

2. เกรเดียนแบบปรับตัวได้ (Adaptive Gradient Descent; AdaGrad)

วิธีนี้จะปรับอัตราการเรียนรู้จากค่าเริ่มต้นที่กำหนดด้วยตัวเอง จากนั้นจะนำค่าเกรเดียนในอดีตมาใช้ในการปรับค่าอัตราการเรียนรู้ ดังสมการที่ 15 และ 16

$$g_t = \frac{\partial J_t}{\partial w} \quad (15)$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{\sum_{k=1}^t g_k^2}} g_t \quad (16)$$

โดย g_t แทนเกรเดียนที่เวลา t

3) อาร์เอ็มเอสพรอป (RMSProp)

วิธีนี้จะเก็บค่าเกรเดียนของครั้งก่อนหน้ามาใช้ในรอบการเรียนรู้ปัจจุบัน และนำไปปรับปรุงอัตราส่วนของอัตราการเรียนรู้ โดยมีการใช้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) สำหรับเก็บค่าเฉลี่ยของเกรเดียน ดังสมการ 17 18 และ 19 ตามลำดับ

$$g_t = \frac{\partial J_t}{\partial w} \quad (17)$$

$$MeanSquare_t = \gamma MeanSquare_{t-1} + (1 - \gamma) g_t^2 \quad (18)$$

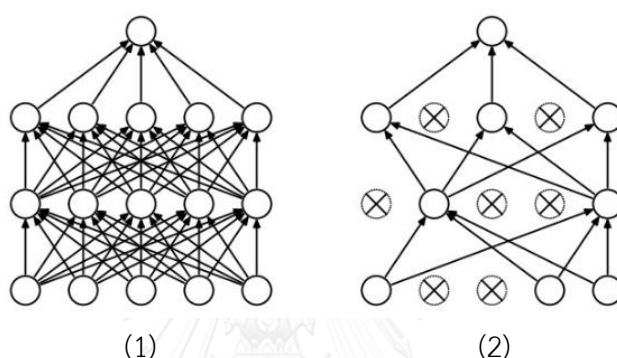
$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{MeanSquare_t}} g_t \quad (19)$$

โดย γ แทนอัตราการใช้เกรเดียนของอดีตในการเรียนรู้ นิยมกำหนดให้เป็น 0.9

2.1.5 การดรอปเอาต์ (Dropout)

เป็นวิธีป้องกันการอิงกับข้อมูลเรียนรู้มากเกินไป (Overfitting) และปัญหาการตรวจจับคุณลักษณะซ้ำ ๆ กันของนิวรอนมากกว่าสองตัวขึ้นไปในโครงข่ายประสาทเทียม เรียกว่า การปรับตัวร่วมกัน (Co-adaptation) ทำให้โครงข่ายประสาทเทียมไม่สามารถทำงานได้อย่างมี

ประสิทธิภาพและเป็นการสิ้นเปลืองทรัพยากรในคำนวณค่าของนิวรอนแต่ละตัวที่สุดท้ายแล้วได้ผลลัพธ์ออกมาเหมือนกัน การดรอปเอาต์มีหลักการทำงานคือ จะสุ่มตัดเส้นเชื่อมของโครงข่ายประสาทเทียมในระหว่างการเรียนรู้ ดังรูปที่ 3 ในทุก ๆ รอบการเรียนรู้ของข้อมูลแต่ละตัว โดยจะไม่มีการทำการดรอปเอาต์ในการทดสอบ ซึ่งในขั้นตอนการเรียนรู้ผู้วิจัยจะสุ่มค่าความน่าจะเป็นแต่ละการเชื่อมต่อของค่าน้ำหนัก ซึ่งค่าน้ำหนักแต่ละค่าจะถูกสุ่มให้อยู่ระหว่าง 0 ถึง 1 หากสุ่มได้เป็น 0 จะไม่ถูกนำไปใช้คำนวณในการเรียนรู้ และในขั้นตอนการทดสอบค่าน้ำหนักทุกค่าจะถูกคูณด้วยค่าความน่าจะเป็นที่กำหนดให้



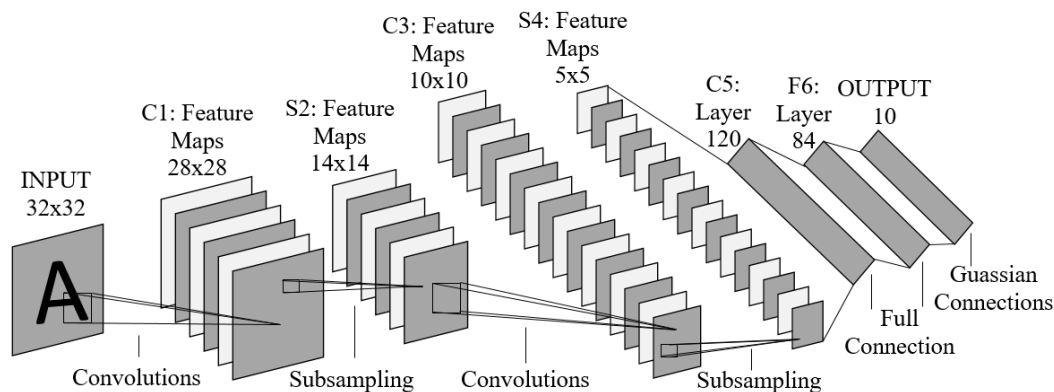
รูปที่ 3 เปรียบเทียบการดรอปเอาต์ในโครงข่ายประสาทเทียม

- (1) แสดงโครงข่ายประสาทเทียมแบบปกติ
- (2) แสดงการดรอปเอาต์ที่มีการตัดเส้นเชื่อมของโครงข่ายประสาทเทียม

[ที่มา: <http://cs231n.github.io/neural-networks-2/> Accessed: January 5, 2017]

2.1.6 โครงข่ายประสาทเทียมเชิงลึก (Deep Neural Network)

เป็นโครงข่ายประสาทเทียมที่มีหลาย ๆ ชั้นและหลากหลายรูปแบบ เช่น โครงข่ายประสาทเทียมความเชื่อเชิงลึก (Deep Belief Network หรือ DBN) โครงข่ายประสาทเทียมกองซ้อนของตัวเข้ารหัส (Stacked Auto Encoders) โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network) หน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory; LSTM) โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) มีความแตกต่างจากโครงข่ายประสาทเทียมแบบธรรมดาตรงที่ โครงข่ายประสาทเทียมแบบคอนโวลูชันเหมาะกับข้อมูลนำเข้าแบบ 2 หรือ 3 มิติ ในขณะที่โครงข่ายประสาทเทียมแบบธรรมดาเหมาะกับข้อมูลนำเข้า 1 มิติ โครงสร้างของโครงข่ายประสาทเทียมคอนโวลูชันมีได้หลายรูปแบบจากการนำชั้นต่าง ๆ ดังรูปที่ 4



รูปที่ 4 โครงข่ายประสาทเทียมคอนโวลูชันแบบ LeNet-5

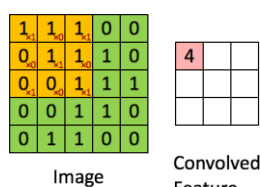
[ที่มา: <http://www.rsipvision.com/exploring-deep-learning/>

Accessed: December 18, 2016]

โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network) เป็นโครงข่ายประสาทเทียมเชิงลึกประเภทหนึ่ง มีจุดเริ่มต้นมาจากการวิจัยทางด้านการรู้จำภาพตัวอักษร โดยมักจะใช้ข้อมูลรับเข้าเป็นเมทริกซ์จากการแปลงมาจากรูปภาพ โครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชันมีรายละเอียดดังนี้

1. ชั้นคอนโวลูชัน (Convolutional Layer)

ประกอบด้วย เคอร์เนลฟิลเตอร์ (Kernel Filter) ซึ่งแต่ละเคอร์เนลจะถูกกำหนดโดยการสุ่ม ในขั้นตอนการเรียนรู้เริ่มต้น หลังจากนั้นจะปรับค่าโดยการเรียนรู้แบบแพร่กระจายย้อนกลับจำนวนผลลัพธ์ที่ได้จากชั้นคอนโวลูชัน จะเท่ากับจำนวนของเคอร์เนลฟิลเตอร์ เรียกว่า ฟีเจอร์แมพ (Feature map) ในชั้นของคอนโวลูชันมักจะตามด้วยฟังก์ชันกระตุ้น ซึ่งเป็นฟังก์ชันแบบไม่เชิงเส้น (Non-linear Function) ขั้นตอนการทำคอนโวลูชันเป็นดังรูปที่ 5



รูปที่ 5 แสดงขั้นตอนการทำคอนโวลูชันระหว่างภาพนำเข้า

[ที่มา: <http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/> Accessed: December 18, 2016]

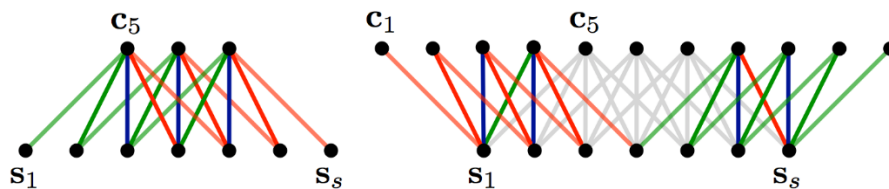
- ขนาดของตัวกรอง (Filter Size)

คือความกว้างและความสูงของตัวกรองที่จะนำมาใช้ในการทำคอนโวลูชัน

- การทำคอนโวลูชันแบบแคบ (Narrow Convolution)

นิยมใช้ในการทำคอนโวลูชันโดยทั่วไป ในการทำคอนโวลูชัน ตัวกรองที่นำไปทำ การต่อทเมทริกซ์นั้นจะไม่มีผลกระทบเลยขอบของเมทริกซ์รับเข้า กล่าวคือ การคอนโวลูชันที่มีข้อมูลรับเข้าขนาด $N \times N$ กับตัวกรองขนาด $m \times m$ จะได้เมทริกซ์ขนาด $(N-m+1) \times (N-m+1)$

นอกจากนี้ยังมีการทำคอนโวลูชันแบบกว้าง (Wide Convolution) จะมีการกระทำเลยขอบของเมทริกซ์รับเข้าออกไป โดยพื้นที่ที่เกินออกไปนั้นจะมีการแทนค่าของข้อมูลช่องนั้น ๆ ด้วย 0 เรียกว่า การเสริมด้วยศูนย์ (Zero Padding) กล่าวคือ การทำคอนโวลูชันแบบกว้างที่มีข้อมูลรับเข้าขนาด $N \times N$ กับตัวกรองขนาด $m \times m$ จะได้เมทริกซ์ขนาด $(N+m-1) \times (N+m-1)$ การทำคอนโวลูชันแบบกว้างนี้มีจุดประสงค์เพื่อป้องกันการสูญเสียข้อมูลตรงบริเวณขอบของข้อมูลรับเข้า โดยรูปที่ 6 เป็นการเปรียบเทียบระหว่างการทำคอนโวลูชันแบบแคบและแบบกว้าง

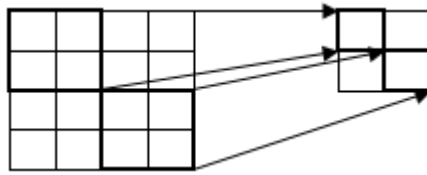


รูปที่ 6 เปรียบเทียบการทำคอนโวลูชันแบบแคบและแบบกว้าง

[ที่มา: <http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/> Accessed: December 18, 2016]

- ขนาดของการก้าวข้าม (Stride Size)

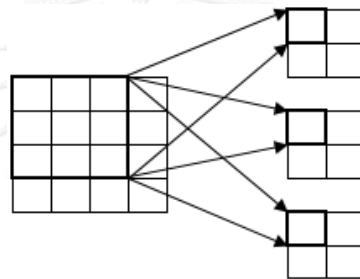
ขนาดของการก้าวข้าม คือจำนวนช่องของข้อมูลรับเข้า ที่จะถูกเลื่อนไปเมื่อทำการหาผลลัพธ์ของคอนโวลูชันในแต่ละช่อง ดังรูปที่ 7 แต่โดยทั่วไปมักใช้ขนาดของการก้าวข้ามเป็นหนึ่ง



รูปที่ 7 การทำคอนโวลูชันโดยมีข้อมูลรับเข้าขนาด 4x4 ตัวกรองขนาด 2x2

- จำนวนตัวกรอง (Number of Filters)

ในการแต่ละชั้นของคอนโวลูชันสามารถมีตัวกรองได้มากกว่าหนึ่ง ดังรูปที่ ๗ โดยน้ำหนักของตัวกรองแต่ละตัวจะแยกกัน จำนวนตัวกรองในชั้นคอนโวลูชันใด ๆ จะเป็นการกำหนดจำนวนช่องสัญญาณของข้อมูลรับเข้าในชั้นถัดไป



รูปที่ 8 การทำคอนโวลูชันโดยมีจำนวนตัวกรองเท่ากับ 3

- จำนวนช่องสัญญาณ (Channel)

จำนวนช่องสัญญาณ หรือความลึกของข้อมูลรับเข้า อาจเกิดจากจำนวนของตัวกรองในชั้นคอนโวลูชันก่อนหน้า หรืออาจมีค่ามากกว่าหนึ่ง เช่น การใช้ช่องสัญญาณทั้งหมด 3 ช่องสัญญาณแทนค่าแม่สีในงานวิจัยทางด้านการใช้รูปภาพ หรือสามารถเขียนสมการในการคำนวณผลลัพธ์ของชั้นคอนโวลูชันได้ดังสมการที่ 20

$$z_{ij}^l = \sum_{c=0}^{k-1} \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} w_{a,b}^l a_{c,i+a,j+b}^{l-1} + b^l \quad (20)$$

โดย k แทนจำนวนช่องสัญญาณ

หลังจากชั้นคอนโวลูชันจะมีการคำนวณฟังก์ชันกระตุ้น เพื่อนำผลลัพธ์ที่ได้ส่งต่อเป็นข้อมูลเข้าให้ชั้นต่อไปในเน็ตเวิร์ก โดยฟังก์ชันกระตุ้นที่เหมาะสมกับการทำนายอายุการใช้งานคงเหลือซึ่งมีลักษณะเป็นค่าตัวเลข ได้แก่ ฟังก์ชันเรคตีไฟต์เชิงเส้นและฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้น แต่ทั้งสองฟังก์ชันมีความแตกต่างกัน คือขอบเขตของผลลัพธ์ดังแสดงในหัวข้อที่ 3.1.2 ตารางที่ 1 ซึ่งถ้าข้อมูลเข้ามีค่าน้อยกว่าหนึ่ง ฟังก์ชันเรคตีไฟต์เชิงเส้นจะมีขอบเขตของผลลัพธ์เท่ากับศูนย์ ในขณะที่ฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้นจะให้ขอบเขตของผลลัพธ์ติดตามแนวโน้มแบบเอกซ์โพเนนเชียล จึงส่งผลให้ค่าที่แบบจำลองทำนายได้มีความแม่นยำมากกว่า

2. ชั้นพูลลิง (Pooling Layer หรือ Subsampling Layer)

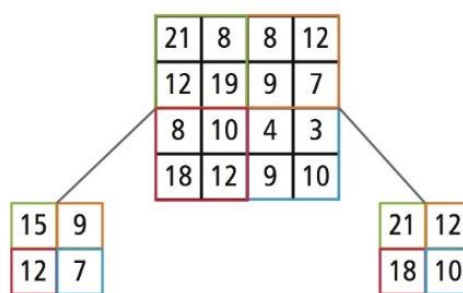
ในชั้นนี้มีจุดประสงค์เพื่อทำการลดขนาดของข้อมูลที่ผ่านมาจากการคอนโวลูชันมาแล้ว นิยมนำมาทำต่อกับชั้นคอนโวลูชัน แต่ก็อาจไม่จำเป็นต้องนำมาต่อกันเสมอไป ซึ่งจะขึ้นอยู่กับการออกแบบ การพูลลิงที่เป็นที่นิยมมีสองวิธีคือ

- การทำพูลลิงแบบหาค่ามากที่สุด (Max Pooling)

เป็นวิธีที่นิยมแพร่หลายมากในงานวิจัยด้านคอนโวลูชันแนลโครงข่ายประสาทเทียมในปัจจุบัน

- การทำพูลลิงแบบหาค่าเฉลี่ย (Average Pooling)

มีวิธีการพูลลิงเหมือนกับการพูลลิงแบบหาค่ามากที่สุด แต่ผลลัพธ์ที่ได้จากการพูลลิงจะเป็นค่าเฉลี่ย ในรูปที่ 9 แสดงให้เห็นถึงการเปรียบเทียบกันของการทำพูลลิงทั้งสองแบบ



Average Pooling

Max Pooling

รูปที่ 9 แสดงการพูลลิงแบบหาค่าเฉลี่ยและหาค่ามากที่สุด

[ที่มา: <http://www.embedded-vision.com/sites/default/files/technical-articles/CadenceCNN/Figure7.jpg> Accessed: December 18, 2016]

3. ชั้นการเชื่อมโยงเต็มรูปแบบ (Fully Connected Layer)

เป็นขั้นสุดท้ายของนิเวศวิทยาโครงข่ายประสาทเทียมคอนโวลูชันจะเป็นการเชื่อมโยงเต็มรูปแบบ หลังจากการประกอบกันของชั้นคอนโวลูชันและชั้นพูลลิง โดยในชั้นนี้ประกอบด้วย ชั้นย่อย ๆ ที่มีเพอร์เซ็ปตรอน อยู่จำนวนหนึ่ง ซึ่งเพอร์เซ็ปตรอนแต่ละตัวจะมีเส้นเชื่อมกับเพอร์เซ็ปตรอนทุกตัวในชั้นก่อนหน้าและเพอร์เซ็ปตรอนทุกตัวในชั้นถัดไป ทำให้การคำนวณแบบป้อนไปข้างหน้าและการแพร่กระจายย้อนกลับสามารถทำได้ด้วยวิธีการปกติ

2.1.7 การทำเรกูลาร์ไรเซชัน (Regularization)

หลังจากทำการหาคุณลักษณะพิเศษ (Feature Extraction) แล้ว ต้องมีการพิจารณาถึงสิ่งที่สำคัญ หรือไม่สำคัญ เช่น ข้อมูลรบกวน (Noise) จึงจำเป็นต้องมีการทำเรกูลาร์ไรเซชัน เพื่อจัดการกับความซับซ้อนของแบบจำลองด้วยการให้น้ำหนักและความสำคัญของข้อมูลที่จะใช้ในการสร้างแบบจำลอง ซึ่งถือเป็นการลดข้อมูลที่ไม่สำคัญออกไป เพื่อให้แบบจำลองมีความซับซ้อนน้อยลง

1. เรกูลาร์ไรเซชันแบบแอลหนึ่ง (L^1 Regularization)

คือการนำผลบวกค่าสัมบูรณ์ของค่าน้ำหนักรวมกับค่าฟังก์ชันต้นทุน ดังสมการที่ 21

$$C = C_0 + \frac{\lambda}{n} \sum_w |w| \quad (21)$$

โดย C_0 แทนฟังก์ชันต้นทุนเริ่มต้น

λ แทนค่าพารามิเตอร์เรกูลาร์ไรเซชัน

2. เรกูลาร์ไรเซชันแบบแอลสอง (L^2 Regularization)

คือการนำผลบวกกำลังสองของค่าน้ำหนักทั้งหมดรวมกับฟังก์ชันต้นทุน ดังสมการที่ 22 เพื่อลดค่าฟังก์ชันต้นทุนให้น้อยลง ซึ่งจะช่วยลดการอิงกับข้อมูลเรียนรู้มากเกินไป

$$C = C_0 + \frac{\lambda}{2n} \sum_w w^2 \quad (22)$$

โดย C_0 แทนฟังก์ชันต้นทุนเริ่มต้น

λ แทนค่าพารามิเตอร์เรกูลาร์ไรเซชัน

ในการถดถอยแบบเชิงเส้น (Linear Regression) ค่าฟังก์ชันต้นทุนจะเป็นผลบวกค่าความคลาดเคลื่อนกำลังสองดังสมการที่ 23 และเมื่อนำค่าเรกูลาร์ไรเซชันแบบแอลสองรวมกับฟังก์ชันต้นทุนดังกล่าวจะได้เป็นสมการที่ 24 โดยมีค่าน้ำหนักเป็นดังสมการที่ 25

$$C = (\mathbf{X}\mathbf{w} - \mathbf{y})^T (\mathbf{X}\mathbf{w} - \mathbf{y}) \quad (23)$$

$$C = (\mathbf{X}\mathbf{w} - \mathbf{y})^T (\mathbf{X}\mathbf{w} - \mathbf{y}) + \frac{1}{2} \alpha \mathbf{w}^T \mathbf{w} \quad (24)$$

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X} + \alpha \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y} \quad (25)$$

การทำเรกูลาร์ไรเซชันทั้งสองแบบมีความแตกต่างกันหลายประการ ดังแสดงในตารางที่ 3 โดยการเลือกใช้จะขึ้นกับจุดประสงค์ กรณีที่สนใจ และลักษณะของข้อมูลส่งออก

ตารางที่ 3 ความแตกต่างระหว่างเรกูลาร์ไรเซชันแบบแอลหนึ่งและแอลสอง

L ¹ Regularization	L ² Regularization
Robust	Not very robust
Unstable solution	Stable solution
Possibly multiple solutions	One solution
No feature selection	Built-in feature selection
sparse outputs	Non-sparse outputs
Computational inefficient on non-sparse cases	Computational efficient due to having analytical solutions

[ที่มา: <http://www.chioka.in/differences-between-l1-and-l2-as-loss-function-and-regularization/> Accessed: December 25, 2016]

2.1.8 การพยากรณ์อนุกรมเวลาด้วยวิธีอาร์มีนา (ARIMA)

การทำเหมืองข้อมูลมีความเกี่ยวข้องกับหลักการทางสถิติค่อนข้างมาก เนื่องจากมีการศึกษาการนำข้อมูลมาวิเคราะห์เพื่อทำนายเช่นเดียวกัน โดยสำหรับข้อมูลอนุกรมเวลาเป็นลำดับของข้อมูลที่มีการเก็บในช่วงเวลาใดช่วงเวลาหนึ่ง

การพยากรณ์ข้อมูลอนุกรมเวลานี้ นิยมใช้วิธีอาร์มีนาบ/ล หรือบ็อกซ์เจนกินส์ (Box-Jenkins) เป็นวิธีที่รวมวิธีการ Auto Regression (AR) และ Moving Average (MA) เข้าด้วยกัน ดังสมการที่ 26

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} + e_t \quad (26)$$

โดย y_t แทนค่าสังเกตช่วงเวลา t

c แทนค่าคงที่

e_t แทนความคลาดเคลื่อนสุ่มที่เวลา t ซึ่งเป็นอิสระกับ

ค่าเฉลี่ย 0 และความแปรปรวนคงที่

จากสมการที่ 23 ϕ_i ($i = 1, 2, \dots, p$) และ θ_j ($j = 1, 2, \dots, q$) คือ พารามิเตอร์แบบจำลองในส่วนของ p และ q ซึ่งถ้า $q = 0$ จะเป็นแบบจำลอง AR(p) และถ้า $p = 0$ จะเป็นแบบจำลอง MA(q) ในการสร้างแบบจำลองการพยากรณ์ข้อมูลอนุกรมเวลาด้วย ARIMA มี 3 ขั้นตอนดังนี้

1. การค้นหาแบบจำลอง (Model Estimation)

การค้นหาแบบจำลอง คือ การพิจารณาว่าข้อมูลคงที่ (Stationary) คือ มีค่าเฉลี่ย ความแปรปรวน และความสัมพันธ์ (Autocorrelation) คงที่ และมีการเปลี่ยนแปลงตามฤดูกาลหรือไม่ โดยดูจากกราฟ Autocorrelation และ Partial Autocorrelation หากพบว่าไม่คงที่ให้ทำการหาค่าความแตกต่าง (Differencing)

2. การประมาณค่าแบบจำลอง (Model Estimation)

ค่าพารามิเตอร์ของแบบจำลองจะมีการกำหนดมาในรูปแบบ ARIMA(p,d,q) โดยที่ p แทนจำนวนลำดับของ Auto regressive d แทนจำนวนลำดับของ Differencing และ q แทนจำนวนลำดับของ Moving average

3. การตรวจสอบความถูกต้องของแบบจำลอง (Model Validation)

มีจุดประสงค์เพื่อ ประเมินว่าสามารถทำนายข้อมูลได้ดีเพียงใด โดยพิจารณาค่าหลงเหลือ (Residual) คือ ความแตกต่างระหว่างค่าที่ได้จากการสังเกตและค่าที่ทำนายด้วยแบบจำลอง จากการสร้างกราฟแท่งความถี่ (Histogram) โดยค่าดังกล่าวจะต้องเป็นแบบสุ่ม คือ มีการกระจายปกติ หากไม่ใช่ให้เริ่มทำแบบจำลองใหม่ และแบบจำลองที่ดีต้องมีลักษณะเป็นแบบจำลองที่ประหยัด (Parsimonious Model) กล่าวคือมีจำนวนพารามิเตอร์น้อยที่สุด หรือรูปแบบง่ายที่สุด แต่ให้ค่าการทำนายที่แม่นยำสูงที่สุด

2.1.9 การพยากรณ์อนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โพเนนเชียล (Exponential smoothing)

เป็นการนำเทคนิคการปรับเรียบโดยคำนึงถึงอิทธิพลของข้อมูลในอดีตที่เปลี่ยนแปลงตามเวลาใช้ในการพยากรณ์ โดยจะให้ความสำคัญกับข้อมูลในปัจจุบันมากกว่าในอดีตที่ห่างออกไป กล่าวคือ เมื่อเวลาเปลี่ยนแปลงไปจะทำให้ข้อมูลในปัจจุบันมีผลต่อการพยากรณ์มากกว่าข้อมูลในอดีต จึงเป็นวิธีที่เหมาะสมกับการพยากรณ์ข้อมูลล่วงหน้าในระยะเวลาสั้นมากกว่ายาว โดยมีรูปแบบของค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก (Weighted Moving Average) ที่จะลดลงแบบเอกซ์โพเนนเชียล ซึ่งข้อมูลล่าสุดจะมีค่าถ่วงน้ำหนักมากที่สุดและจะลดลงไปเรื่อย ๆ ตามเวลาของข้อมูล นอกจากนี้ยังจำเป็นต้องมีค่าคงที่ที่ต้องการปรับเรียบ ซึ่งจะมีค่าตั้งแต่ 0 ถึง 1 ที่ถูกเลือกโดยการสุ่ม สมการของการพยากรณ์ด้วยวิธีการปรับเรียบแบบเอกซ์โพเนนเชียลเป็นดังสมการที่

27

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1}) \quad (27)$$

โดย F_t คือ ค่าที่ต้องการพยากรณ์ช่วงเวลา t

F_{t-1} คือ ค่าก่อนหน้า $t-1$

α คือ ค่าคงที่ค่าคงที่ที่ต้องการปรับเรียบมีค่า $0 \leq \alpha \leq 1$

สำหรับการสุ่มเลือกค่าคงที่ในการปรับเรียบควรเลือกตามการเปลี่ยนแปลงของข้อมูล หากมีการเปลี่ยนแปลงเร็วมาก ควรสุ่มเลือกค่าคงที่การปรับเรียบมาก ๆ หรือเข้าใกล้ 1 ในทางกลับกันหากการเปลี่ยนแปลงข้อมูลน้อย ควรสุ่มเลือกค่าคงที่การปรับเรียบให้มีค่าน้อย ๆ หรือเข้าใกล้ 0

2.1.10 การวัดประสิทธิภาพ (Performance Evaluation)

ในการประเมินว่าแบบจำลองทำนายได้แม่นยำเพียงใด นิยมการวิเคราะห์จากความคลาดเคลื่อนระหว่างค่าที่ทำนายได้กับค่าจริง หรือค่าผลเฉลยของชุดข้อมูล วิธีการวัดที่นิยมใช้ในการประเมินได้แก่ ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error; RMSE) ซึ่งเป็นค่าที่ใช้วัดความแตกต่างระหว่างค่าจริงกับค่าที่ทำนายได้ หากมีค่าน้อย แสดงว่าแบบจำลองทำนายได้ใกล้เคียงกับความเป็นจริง และเกิดความคลาดเคลื่อนน้อย ดังสมการที่ 28

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t^s - Y_t^a)^2} \quad (28)$$

โดย Y_t^s แทนค่าประมาณจากแบบจำลอง

Y_t^a แทนค่าจริง

T แทนจำนวนคาบเวลาที่ใช้ประมาณแบบจำลอง

2.2 งานวิจัยที่เกี่ยวข้อง

ในการศึกษางานวิจัยที่เกี่ยวข้อง ได้แบ่งออกเป็นสามด้าน โดยมีรายละเอียดดังต่อไปนี้

2.2.1 งานวิจัยเกี่ยวกับการบำรุงรักษาแบบพยากรณ์ล่วงหน้า

Susto, Schirru และคณะ [1] ได้ทำนายข้อมูลมิติขนาดใหญ่ (High dimensional Data) ด้วยวิธีการเรียนรู้ของเครื่อง ต่อมา Wong, Yang และคณะ [2] เสนอวิธีการ Extreme Learning Machine (ELM) เปรียบเทียบกับการใช้ Support Vector Machine โดยใช้วิธี Extreme Learning Machine ซึ่งเป็นเทคนิคการวิเคราะห์ข้อมูลทางสถิติมาจัดการข้อมูลก่อนเข้าสู่กระบวนการ (Data preprocessing) มาใช้ร่วมกับวิธีการโครงข่ายประสาทเทียม ทำให้มีความแม่นยำเพิ่มขึ้น

Darabad, Vakilian และคณะ [3] นำเสนอเทคนิค Self-organizing map (SOM) ซึ่งเป็นการนำวิธีการวิเคราะห์ทางสถิติรวมกับการจำแนกข้อมูลแบบแบ่งกลุ่ม (Clustering) โดยทั้ง [2] และ [3] ใช้ข้อมูลสัญญาณความสั่นสะเทือนเป็นข้อมูลอินพุต Prytz, Rune และคณะ [4] ได้นำข้อมูลการบำรุงรักษา (Maintenance Records) และข้อมูลประวัติการตรวจสอบสภาพ หรือซ่อมแซมยานพาหนะ (Logged Vehicle Data) มาพยากรณ์ว่า ยานพาหนะดังกล่าวสามารถใช้งานได้อีกเป็นเวลาเท่าใดด้วยการวิเคราะห์ซ้ำ 10 ครั้ง (10 folds cross validation) ด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree) มาวิเคราะห์ผ่านภาษา R ซึ่งวิธีการดังกล่าวสามารถลดการซ่อมนอกแผนงาน (Unplanned Repair)

งานวิจัยของ Langone, Alzate และคณะ [5] นำเสนอการใช้วิธี Least Squares Support Vector Machines (LS-SVM) ในการตรวจจับความผิดปกติของอุปกรณ์แบบ Real Time และวิธี Kernel Spectral Clustering (KSC) ใช้แบ่งกลุ่มข้อมูลปกติและผิดปกติ จากนั้นมาทำการพยากรณ์ด้วย nonlinear auto-regressive (NAR) Model

2.2.2 งานวิจัยเกี่ยวกับการวิเคราะห์ข้อมูลแบบอนุกรมเวลา

Ismail และ Hassan [6] ได้ทำการพยากรณ์ค่าโหลดทางไฟฟ้า (Electric Load) ซึ่งข้อมูลนำเข้าที่มีลักษณะเป็นอนุกรมเวลา (Time-Series Data) โดยนำเสนอวิธีการ REGARIMA เปรียบเทียบกับเทคนิคโครงข่ายประสาทเทียม วิธีการดังกล่าวได้นำรูปแบบการถดถอย (Regression) และอาร์มีน่า (ARIMA) รวมกัน เพื่อเพิ่มประสิทธิภาพในการพยากรณ์ข้อมูล เหมาะสำหรับการพยากรณ์ล่วงหน้าแบบระยะสั้น (Short Term Forecasting) ที่นิยมใช้ในการบำรุงรักษาโรงไฟฟ้า

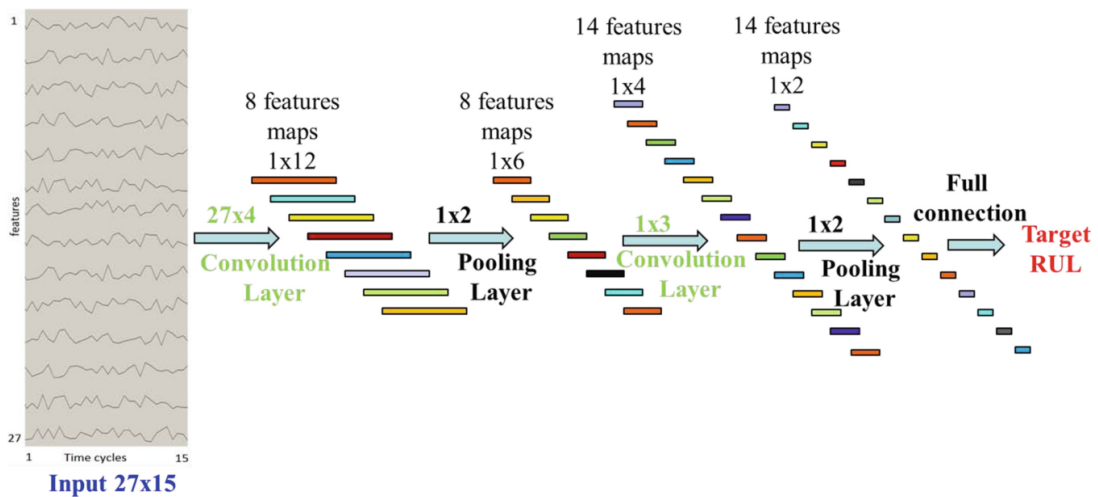
นอกจากนี้ Pereira, Pacheco และคณะ [7] ได้เสนอการตรวจจับความผิดปกติ (Fault Detection) ของกังหันก๊าซ (Gas Turbine) ในโรงไฟฟ้า โดยใช้การรู้จำรูปแบบ (Pattern Recognition) ซึ่งเหมาะสมกับข้อมูลนำเข้าที่เป็นอนุกรมเวลาแบบหลายตัวแปร (Multivariable Time Series Data) การรู้จำรูปแบบข้อมูลเมื่อเครื่องเริ่มต้นทำงานปกติ และหยุดทำงานด้วยเหตุผลบางประการ (Trip) ผลจากการวิเคราะห์ด้วยวิธีการดังกล่าวสามารถนำไปใช้เฝ้าระวัง (Monitoring) และทำนายโอกาสที่อุปกรณ์จะเสียได้

2.2.3 งานวิจัยเกี่ยวกับการทำนายอายุการใช้งานคงเหลือของเครื่องจักร

ในปี 2012 S. Porotsky [8] ได้เสนอวิธีการใช้โครงข่ายประสาทเทียมร่วมกับ Adaptive Network-based Fuzzy Inference System (ANFIS) สำหรับข้อมูลแบบ non-trendability ต่อมาในปี 2013 Baraldi, P., Compare, M., Saucio, S. และ Zio, E. [9] ทำนายด้วยการใช้หลักการ Particle Filtering (PF) ร่วมกับโครงข่ายประสาทเทียม และยังมีงานวิจัยที่นำโครงข่ายประสาทเทียมมาใช้ร่วมกับวิธีการอื่นเพื่อเพิ่มประสิทธิภาพในการทำนายอย่างหลากหลาย

จนกระทั่งในปี 2016 ได้มีการเสนอวิธีใหม่คือ โครงข่ายประสาทเทียมประสาทแบบคอนโวลูชันนอลเชิงลึกด้วยการถดถอย (Regression) โดย Giduthuri Sateesh Babu, Peilin Zhao, และ Xiao-Li Li [10] ซึ่งนับเป็นครั้งแรกในการนำหลักการดังกล่าวมาใช้ในการทำนายอายุการใช้งานคงเหลือของเครื่องจักร โดยใช้กับข้อมูลจำลองการเสื่อมของเครื่องยนต์กังหันไอพ่น (Turbofan Engine Degradation) วิธีการดังกล่าวสามารถดึงคุณสมบัติเฉพาะที่มีความซับซ้อนของข้อมูลนำเข้า หรือคุณสมบัติต่าง ๆ มาช่วยในการทำนายได้อย่างอัตโนมัติ โดยมีโครงสร้างในการทำนายอายุการใช้งานคงเหลือของเครื่องจักรเป็นดังรูปที่ 10 โดยประกอบด้วยอินพุตแบบอนุกรมเวลาหลายตัวแปร ชั้นคอนโวลูชัน 2 ชั้น ชั้นพูลลิง 2 ชั้น และชั้นการเชื่อมต่อ แบบสมบูรณ์ด้วยเพอร์เซ็ปตรอนหลายชั้น (Multi-Layer Perceptron; MLP) ซึ่งจัดเป็นแบบจำลองแบบดั้งเดิมของเน็ตเวิร์กคอนโวลูชัน ซึ่งได้นำมาประยุกต์ใช้งานกับการทำนายแบบถดถอยเพื่อทำนายอายุการใช้งานคงเหลือ ในขั้นตอนการสอนแบบจำลองจะทำการประมาณค่าพารามิเตอร์

ของเน็ตเวิร์กคอนโวลูชันด้วยขั้นตอนวิธีแบบแพร่กระจายย้อนกลับด้วยวิธีการสโตแคสติกเกรเดียนต์เดสเซนส์ในการหาค่าที่เหมาะสมที่สุดของฟังก์ชันต้นทุนที่ใช้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยสะสมของแบบจำลองเน็ตเวิร์กคอนโวลูชัน



รูปที่ 10 แสดงโครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก [10]

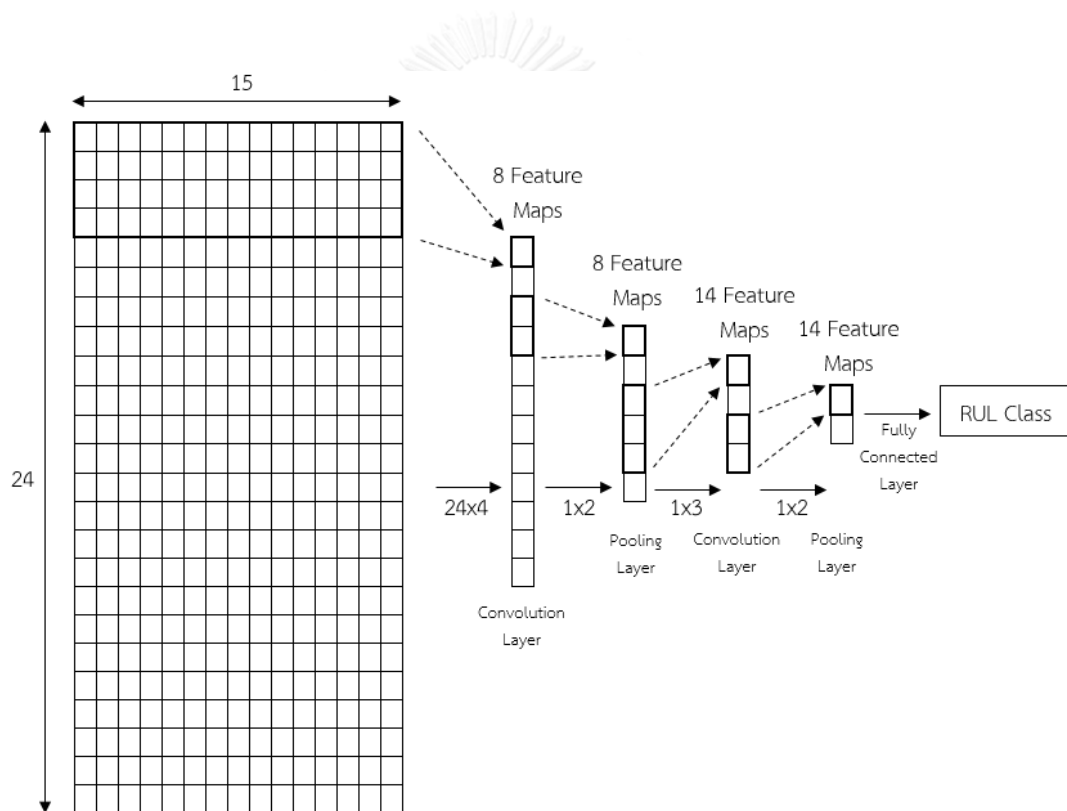
จากงานวิจัยส่วนใหญ่ได้มีการใช้ MATLAB, Python, JAVA และ R มาทำเหมืองข้อมูลร่วมกับวิธีการทางสถิติและการเรียนรู้ของเครื่อง เพื่อแสดงให้เห็นถึงความแม่นยำของการทำนายได้ และช่วยลดค่าใช้จ่ายในการซ่อมบำรุง โดยแต่ละวิธีการที่ใช้นั้นแตกต่างกันไปตามลักษณะข้อมูลนำเข้า ที่มีทั้งวิธีการที่มีประสิทธิภาพสูงและให้ค่าความคลาดเคลื่อนน้อย โดยการนำวิธีการทางด้าน การเรียนรู้ของเครื่องมาประยุกต์ใช้ได้แก่ การใช้เน็ตเวิร์กคอนโวลูชันเชิงลึก ซึ่งมีความสามารถในการหาคุณลักษณะพิเศษของข้อมูลอันได้แก่ ความสัมพันธ์ระหว่างข้อมูลนำเข้าแต่ละตัว ความสัมพันธ์ระหว่างข้อมูลนำเข้าแต่ละตัวและเวลา ความสัมพันธ์ระหว่างข้อมูลนำเข้าแต่ละตัว เวลา และค่าผลเฉลยของการทำนาย โดยถือว่าการวิเคราะห์ได้อย่างรอบด้านและแม่นยำ ส่งผลให้การทำนายมีประสิทธิภาพดีกว่าวิธีการเดิม

บทที่ 3

แบบจำลองการทำนายอายุการใช้งานคงเหลือของเครื่องจักรที่เพิ่มประสิทธิภาพ

3.1 แบบจำลองการทำนายอายุการใช้งานคงเหลือ

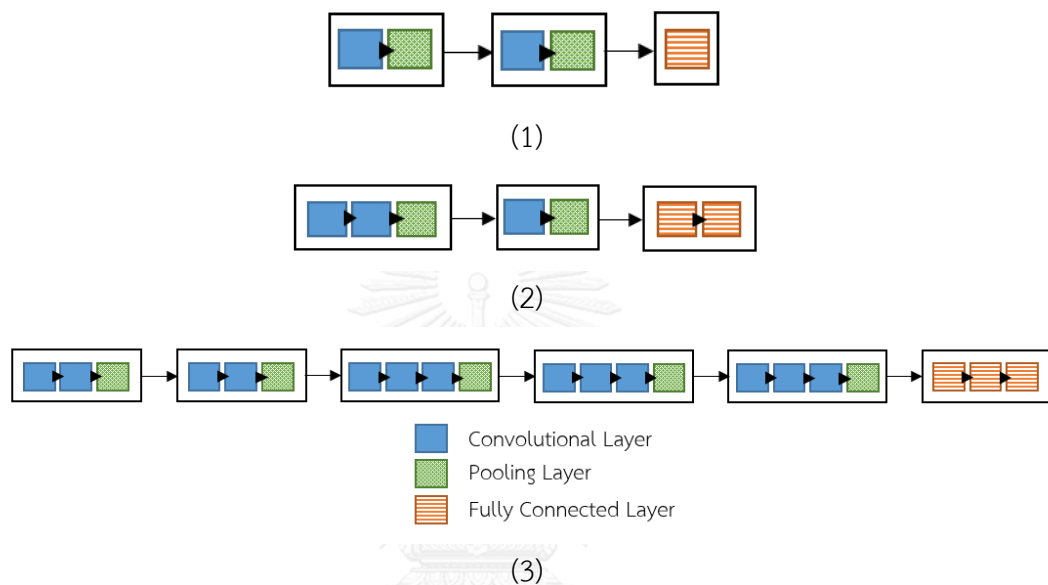
จากการศึกษางานวิจัยที่เกี่ยวข้องจะเห็นว่าโครงข่ายประสาทเทียมที่ดีที่สุดในการทำนายอายุการใช้งานคงเหลือคือ เน็ตเวิร์กคอนโวลูชันเชิงลึกแบบถดถอย ที่ใช้ฟังก์ชันซิกมอยด์เป็นฟังก์ชันกระตุ้นในการทำพีเอเจอร์แมพ การพูลลิงแบบหาค่าเฉลี่ย และสโตแคสติกเกรเดียนเดสเซนส์เป็นฟังก์ชันต้นทุน โดยมีโครงสร้างของการทำนายดังรูปที่ 11



รูปที่ 11 แสดงโครงสร้างการทำนายอายุการใช้งานคงเหลือด้วยโครงข่ายประสาทเทียมคอนโวลูชันเชิงลึกที่มีข้อมูลเข้า 24 จุดวัด ทีละ 15 ช่วงเวลา

เนื่องจากลักษณะชุดข้อมูลในการทำนายอายุการใช้งานคงเหลือของเครื่องจักรนั้นเป็นอนุกรมเวลาแบบหลายตัวแปร ซึ่งในการใช้งานจริงเครื่องจักรหรืออุปกรณ์มักมีจุดวัดจำนวนมาก ทำให้การใช้โครงสร้างพื้นฐานตามงานวิจัยที่เป็นแบบ LeNet และ AlexNet ไม่เหมาะสมและได้

ประสิทธิภาพไม่ดีเท่าที่ควร จึงมีการนำเสนอโครงสร้างที่มีความซับซ้อนมากขึ้น เพื่อให้แบบจำลองสามารถทำนายผลเฉลยสำหรับชุดข้อมูลที่มีลักษณะเป็นอนุกรมเวลาหลายตัวแปรได้แม่นยำมากขึ้น โดยได้โครงสร้างแบบ VGG-16 ในการทำนาย ซึ่งมีโครงสร้างแตกต่างจากโครงสร้างพื้นฐานตรงที่มีการเพิ่มชั้นคอนโวลูชัน ชั้นพูลลิง และชั้นการเชื่อมโยงเต็ม ทั้งนี้โครงสร้างเน็ตเวิร์กคอนโวลูชันทั้งสามแบบได้แสดงไว้ในรูปที่ 12



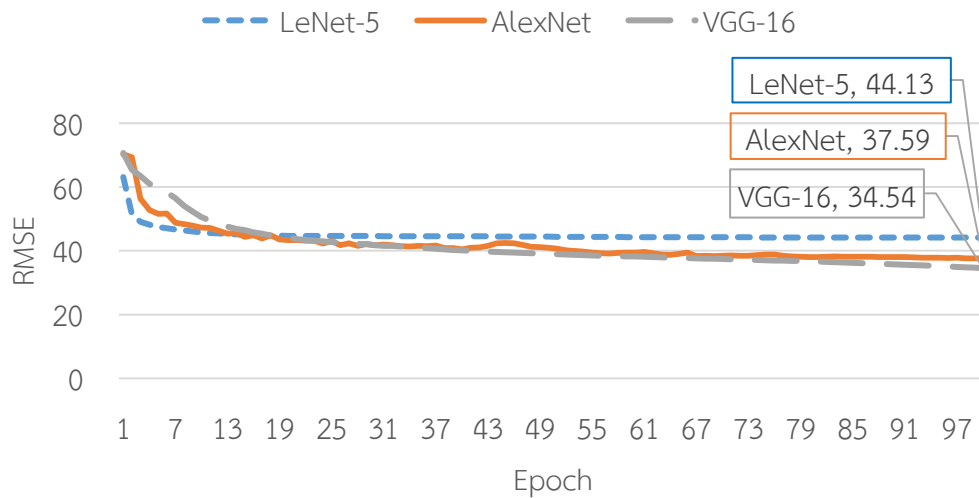
รูปที่ 12 ความแตกต่างของโครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึกแบบต่าง ๆ

- (1) โครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึกแบบ LeNet
- (2) โครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึกแบบ AlexNet
- (3) โครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึกแบบ VGG-16

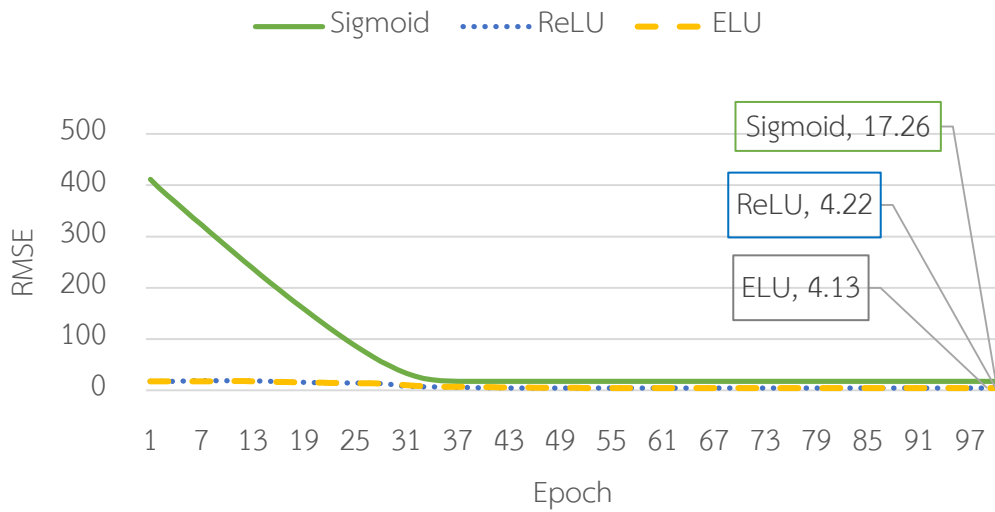
3.2 การเพิ่มประสิทธิภาพให้กับแบบจำลอง

นอกจากการนำเสนอโครงสร้างที่เหมาะสมในการทำนายอายุการใช้งานคงเหลือแล้ว ยังจำเป็นต้องมีการเพิ่มประสิทธิภาพให้กับแบบจำลอง โดยทำให้แบบจำลองที่นำเสนอมีค่าความคลาดเคลื่อนของการทำนายน้อยที่สุด ด้วยวิธีการและเทคนิคต่าง ๆ ดังนี้

1. การปรับโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันเชิงลึกในชั้นการทำคอนโวลูชัน ชั้นการพูลลิง และชั้นการเชื่อมโยงเต็มรูปแบบให้เหมาะสมอันได้แก่ โครงสร้างเน็ตเวิร์กคอนโวลูชันแบบ VGG-16

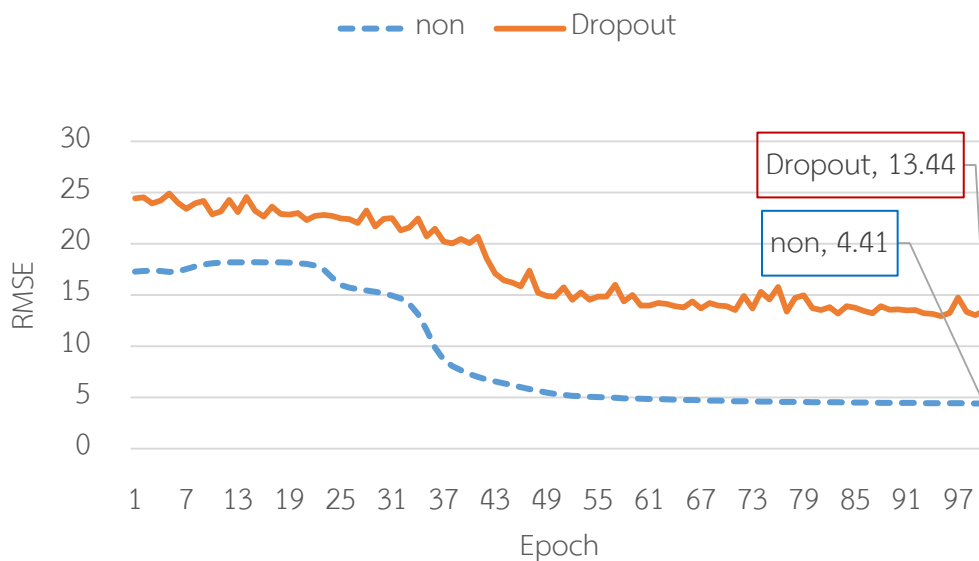


2. การเพิ่มประสิทธิภาพแบบจำลองโดยการเปลี่ยนฟังก์ชันกระตุ้นจากฟังก์ชันซิกมอยด์ เป็นฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้น ซึ่งจะเป็นการช่วยตัดแกนให้ลดลง เนื่องจากหากค่าที่ได้น้อยกว่าศูนย์ ค่าจากฟังก์ชันดังกล่าวจะติดลบตามแนวโน้มแบบเอกซ์โพเนนเชียล จึงเหมาะกับพฤติกรรมของค่าอายุการใช้งานคงเหลือที่เมื่อเครื่องจักรมีอาการผิดปกติ หรือเสียหาย ค่าอายุการใช้งานคงเหลือจะมีแนวโน้มลดลงจนเป็นศูนย์ และเปรียบเทียบกับฟังก์ชันที่เป็นที่นิยมใช้ในโครงสร้างนิเวศวิทยาประสาทเทียมคอนโวลูชันเชิงลึกได้แก่ ฟังก์ชันเรคทีไฟต์เชิงเส้นที่มีการตัดเป็นศูนย์ หากค่าที่ได้ต่ำกว่า หรือเท่ากับศูนย์ รูปที่ 13 แสดงให้เห็นถึงความแตกต่างของการใช้ฟังก์ชันกระตุ้นบนชุดข้อมูล C-MAPSS 1 ของรอบการเรียนรู้ 100 รอบการเรียนรู้แรก โดยที่แกนแนวตั้งคือค่า RMSE และแกนแนวนอนคือจำนวนรอบของการเรียนรู้ ซึ่งเห็นได้ว่า ฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้นทำให้ค่าประสิทธิภาพในการทำนายดีที่สุด รองลงมาคือ ฟังก์ชันเรคทีไฟต์เชิงเส้นและฟังก์ชันซิกมอยด์ตามลำดับ



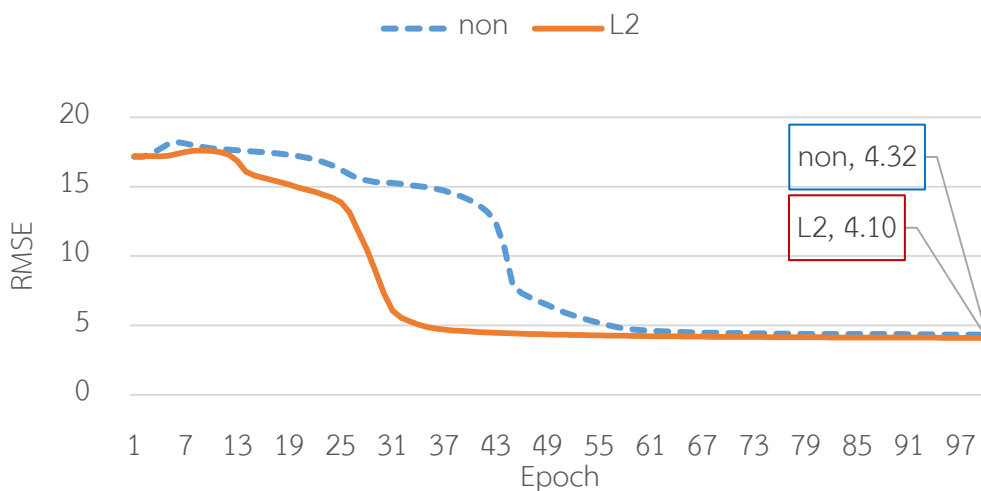
รูปที่ 13 เปรียบเทียบประสิทธิภาพจากการทำนายด้วยแบบจำลองที่ใช้ฟังก์ชันกระตุ้นซิกมอยด์ เรคตีไฟเชิงเส้น และเอกโพเนนเชียลเชิงเส้น 100 รอบการเรียนรู้แรกบนชุดข้อมูล C-MAPSS3

3. การนำเทคนิควิธีป้องกันการอิงกับข้อมูลเรียนรู้มากเกินไป หรือดรอปเอาต์ มาใช้ในการสุ่มตัดเส้นเชื่อมที่ไม่จำเป็นออกไป โดยเทคนิคนี้จะทำให้ประสิทธิภาพการทำนายของแบบจำลองลู่เข้าสู่ค่าที่น้อยขาลง ดังแสดงในรูปที่ 14 ซึ่งเป็นการทดลองให้เห็นถึงประสิทธิภาพของการใช้ดรอปเอาต์บนชุดข้อมูล C-MAPSS3 ของรอบการเรียนรู้ 100 รอบการเรียนรู้แรก



รูปที่ 14 เปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองด้วยการใช้เทคนิคการดรอปเอาต์ 100 รอบการเรียนรู้แรกบนชุดข้อมูล C-MAPSS3

4. การทำเรกูลาร์ไรเซชันแบบแอลสอง เพื่อลดความซับซ้อนของแบบจำลอง หลังจากทำการหาคุณลักษณะพิเศษ โดยการเพิ่มเทคนิคนี้สามารถเพิ่มประสิทธิภาพในการทำนายให้กับแบบจำลองได้ ดังแสดงในรูปที่ 15 ซึ่งเป็นการทดลองให้เห็นถึงประสิทธิภาพของการทำเรกูลาร์ไรเซชันแบบแอลสอง บนชุดข้อมูล C-MAPSS3 ของรอบการเรียนรู้ 100 รอบการเรียนรู้แรก



รูปที่ 15 เปรียบเทียบประสิทธิภาพการทำนายโดยใช้เทคนิคการทำเรกูลาร์ไรเซชันแบบแอลสอง 100 รอบการเรียนรู้แรกบนชุดข้อมูล C-MAPSS3

5. การหาค่าเหมาะที่สุดในการปรับปรุงน้ำหนักของเส้นเชื่อมด้วยเกรเดียนแบบปรับตัวได้ เนื่องจากสโตแคสติกเกรเดียนเดสเซนส์ที่ใช้ค่าอัตราการเรียนรู้ค่าเดียวกันเพียงค่าเดียวคุณกับค่าน้ำหนักทั้งหมดในการปรับค่าน้ำหนักตามสมการที่ 13 และ 14 ในขณะที่เกรเดียนแบบปรับตัวได้มีการปรับค่าอัตราการเรียนรู้สำหรับทุก ๆ พารามิเตอร์ตามกราฟาเดียนในอดีตทั้งหมดดังสมการที่ 15 และ 16 ดังจะเห็นได้จากรูปที่ 16 ซึ่งเป็นการเปรียบเทียบระหว่างการใช้สโตแคสติกเกรเดียนเดสเซนส์และเกรเดียนที่ปรับตัวได้ในการทำนายอายุการใช้งานคงเหลือของชุดข้อมูล C-MAPSS1 ซึ่งแสดงให้เห็นว่าการใช้สโตแคสติกเกรเดียนเดสเซนส์ทางซ้ายให้ค่า RMSE เป็น nan กล่าวคือไม่สามารถคำนวณค่าอายุการใช้งานคงเหลือหรือค่าผลเฉลี่ยของข้อมูลได้ ในขณะที่การใช้เกรเดียนที่ปรับตัวได้สามารถคำนวณออกมาทางด้านขวาของรูปที่ 16

Training set MSE	Training set MSE
epoch 0	epoch 0
Training set RMSE: nan	Training set RMSE: 69.6976536268
nan	4857.76292108
epoch 1	epoch 1
Training set RMSE: nan	Training set RMSE: 69.5601417236
nan	4838.6133166
epoch 2	epoch 2
Training set RMSE: nan	Training set RMSE: 69.4788621333
nan	4827.31228333
epoch 3	epoch 3
Training set RMSE: nan	Training set RMSE: 69.4291179663
nan	4820.40242158
epoch 4	epoch 4
Training set RMSE: nan	Training set RMSE: 69.3960703557
nan	4815.81458081
epoch 5	epoch 5
Training set RMSE: nan	Training set RMSE: 69.3726852295
nan	4812.56945595
epoch 6	epoch 6
Training set RMSE: nan	Training set RMSE: 69.3552340463
nan	4810.14848962
epoch 7	epoch 7
Training set RMSE: nan	Training set RMSE: 69.3418699692
nan	4808.29493082

รูปที่ 16 เปรียบเทียบประสิทธิภาพของแบบจำลองเมื่อใช้สโตแคสติกเกรเดียนเดสเซนส์และเกรเดียนต์ที่ปรับตัวในการทำนายอายุการใช้งานคงเหลือของชุดข้อมูล C-MAPSS1

ทั้งนี้การที่เราใช้สโตแคสติกเกรเดียนเดสเซนส์แล้วคำนวณค่าที่ได้เป็น nan นั้น มีสาเหตุมาจากการที่มีค่าอัตราการเรียนรู้มากเกินไป เมื่อนำมาคูณกับค่าน้ำหนักที่เยอะมาก ๆ ทำให้ได้ผลลัพธ์ที่มีค่าสูงมากจนคำนวณได้ยาก ส่วน เกรเดียนต์แบบปรับตัวได้ ซึ่งมีการปรับค่าอัตราการเรียนรู้ ทำให้มีการปรับจนเหมาะสมที่สุดจึงสามารถคำนวณออกมาเป็นค่าได้

เมื่อนำเทคนิคข้างต้นมาใช้ในแบบจำลอง VGG-16 เพื่อช่วยเพิ่มประสิทธิภาพ โดยมีหลักการทำงานเรียงลำดับกันดังแสดงในรูปที่ 17 ซึ่งเป็นแบบจำลองที่นำเสนอในงานวิจัยสำหรับการทำนายอายุการใช้งานคงเหลือของเครื่องจักร ในการกำหนดขนาดของข้อมูลเข้าในแต่ละชั้น ค่าพารามิเตอร์และฟังก์ชันต่าง ๆ ที่ใช้ในโครงสร้างที่นำเสนอจากรูปที่ 17 นั้น สามารถอธิบายได้ตามรหัสเทียมหรือ Pseudocode ที่แสดงไว้ดังตารางที่ 4 โดยประกอบด้วย อัลกอริธึมที่หนึ่ง การกำหนดโครงสร้างเน็ตเวิร์กคอนโวลูชันแบบ VGG-16 อันได้แก่ การกำหนดค่าชั้นคอนโวลูชัน พูลลิง และการเชื่อมโยงเต็มรูปแบบ อัลกอริธึมที่สอง การใช้เทคนิคเพื่อช่วยในการเพิ่มประสิทธิภาพ อัลกอริธึมที่สาม การกำหนดค่าน้ำหนัก ค่าความเอนเอียง และฟังก์ชันกระตุ้นของชั้นคอนโวลูชัน และอัลกอริธึมที่สี่ การกำหนดค่าการทำพูลลิงแบบหาค่ามากที่สุด ตามลำดับ

ตารางที่ 4 รหัสเทียมหรือ Pseudocode ของโครงสร้างเน็ตเวิร์กคอนโวลูชันที่เพิ่มประสิทธิภาพ

Algorithm 1. VGG-16 Structure
<pre> c <- apply_conv(X, kernel_height <- 2, kernel_width <- 2, num_channels <- 1, depth <- 8) c <- apply_conv(c, kernel_height <- 2, kernel_width <- 2, num_channels <- 8, depth <- 8) p <- apply_max_pool(c, kernel_height <- 1, kernel_width <- 2, stride_size <- 2) c <- apply_conv(c, kernel_height <- 2, kernel_width <- 2, num_channels <- 8, depth <- 8) c <- apply_conv(c, kernel_height <- 2, kernel_width <- 2, num_channels <- 8, depth <- 16) p <- apply_max_pool(c, kernel_height <- 2, kernel_width <- 2, stride_size <- 2) c <- apply_conv(p, kernel_height <- 2, kernel_width <- 2, num_channels <- 16, depth <- 16) c <- apply_conv(c, kernel_height <- 2, kernel_width <- 2, num_channels <- 16, depth <- 16) c <- apply_conv(c, kernel_height <- 2, kernel_width <- 2, num_channels <- 16, depth <- 32) p <- apply_max_pool(c, kernel_height <- 1, kernel_width <- 1, stride_size <- 2) c <- apply_conv(c, kernel_height <- 2, kernel_width <- 1, num_channels <- 32, depth <- 32) c <- apply_conv(c, kernel_height <- 2, kernel_width <- 1, num_channels <- 32, depth <- 32) c <- apply_conv(c, kernel_height <- 2, kernel_width <- 1, num_channels <- 32, depth <- 64) p <- apply_max_pool(c, kernel_height <- 2, kernel_width <- 1, stride_size <- 2) c <- apply_conv(p, kernel_height <- 2, kernel_width <- 1, num_channels <- 64, depth <- 64) c <- apply_conv(c, kernel_height <- 1, kernel_width <- 1, num_channels <- 64, depth <- 64) c <- apply_conv(p, kernel_height <- 1, kernel_width <- 1, num_channels <- 64, depth <- 64) p <- apply_max_pool(c, kernel_height <- 1, kernel_width <- 1, stride_size <- 2) shape <- p.get_shape().as_list() flat <- tf.reshape(p, [-1, shape[1] * shape[2] * shape[3]]) f_weights <- weight_variable([shape[1] * shape[2] * shape[3], num_hidden]) f_biases <- bias_variable([num_hidden]) f <- tf.nn.elu(tf.add(tf.matmul(flat, f_weights), f_biases)) f_weights <- weight_variable([shape[1] * shape[2] * shape[3], num_hidden]) f_biases <- bias_variable([num_hidden]) f <- tf.nn.elu(tf.add(tf.matmul(flat, f_weights), f_biases)) f_weights <- weight_variable([shape[1] * shape[2] * shape[3], num_hidden]) f_biases <- bias_variable([num_hidden]) f <- tf.nn.elu(tf.add(tf.matmul(flat, f_weights), f_biases)) f <- tf.nn.dropout(f, dropout) out_weights <- weight_variable([num_hidden, num_labels]) out_biases <- bias_variable([num_labels]) y_ <- tf.add(tf.matmul(f, out_weights), out_biases) </pre>
Algorithm 2. Enhancement Techniques
<pre> # Loss function using L2 Regularization loss <- tf.square(y_ - Y) regularizer <- tf.nn.l2_loss(out_weights) cost_function <- tf.reduce_mean(loss + beta * regularizer) </pre>

<pre>#Adagrad optimaiztion optimizer <- tf.train.AdagradOptimizer(learning_rate).minimize(cost_function)</pre>
Algorithm 3. Convolutional Layer
<pre>FUNCTION apply_conv(x,kernel_height,kernel_width,num_channels,depth): weights <- weight_variable([kernel_height, kernel_width, num_channels, depth]) biases <- bias_variable([depth]) RETURN tf.nn.elu(tf.add(tf.nn.conv2d(x, weights,[1,1,1,1],padding="SAME"),biases)) ENDFUNCTION</pre>
Algorithm 4. Pooling Layer
<pre>FUNCTION apply_max_pool(x,kernel_height,kernel_width,stride_size): RETURN tf.nn.max_pool(x, ksize=[1, kernel_height, kernel_width, 1], strides=[1, 1, stride_size, 1], padding = "VALID") ENDFUNCTION</pre>

3.3 การเตรียมข้อมูลด้วยการประมวลผลก่อน

เนื่องจากชุดข้อมูลแต่ละตัวแปรมีระยะห่างที่แตกต่างกันมาก เช่น จุดวัดที่หนึ่งอยู่ในช่วงหลัก ร้อย ในขณะที่จุดวัดที่แปดอยู่ในช่วงหลักพัน จึงจำเป็นอย่างยิ่งที่ต้องนำข้อมูลมาทำการนอร์มัลไรเซชัน เพื่อให้ชุดข้อมูลมีช่วงข้อมูลใกล้เคียงกัน เหมาะแก่การนำไปหาความสัมพันธ์ หรือวิเคราะห์ร่วมกัน แบบหลายตัวแปรในขั้นตอนต่อไป ในการแปลงระยะของข้อมูลให้อยู่ในระยะเดียวกัน จะใช้สมการที่ 29 ซึ่งเป็นค่าแปลงข้อมูลแบบ Z-Score ซึ่งเป็นการแทนข้อมูลเดิมในแต่ละจุดด้วยค่าใหม่ ซึ่งเป็นการปรับการกระจายของข้อมูลให้มีค่าเป็นศูนย์และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) เป็นหนึ่ง ทำได้โดยนำผลต่างระหว่างข้อมูลนำเข้ากับค่าเฉลี่ย (Mean) ทหารด้วยค่าเบี่ยงเบนมาตรฐานดังสมการที่ 29

$$z(x_i) = \frac{x_i - \mu}{\sigma} \quad (29)$$

โดย z แทนผลบวกของค่าสัมบูรณ์

μ แทนค่าเฉลี่ย

σ แทนค่าเบี่ยงเบนมาตรฐาน

สำหรับค่า μ และค่า σ สามารถคำนวณได้จากสมการที่ 30 และ 31 ตามลำดับ โดยที่ค่า μ คือค่าเฉลี่ยของข้อมูลอนุกรมเวลา X และค่า σ คือส่วนเบี่ยงเบนมาตรฐานของข้อมูลอนุกรมเวลา X

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad (30)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (31)$$

สำหรับการแปลงชุดข้อมูลในการทดลองซึ่งประกอบด้วยข้อมูลสอนล้ข้อมูลทดสอบ จะมีวิธีการใช้ข้อมูลเพื่อใช้การคำนวณไม่เหมือนกัน เนื่องจากข้อมูลทดสอบถือเป็นข้อมูลที่ต้องการทำนาย ซึ่งในทางความเป็นจริงแล้วจะไม่มีทางทราบถึงรายละเอียดหรือคุณสมบัติของข้อมูลเช่น ค่ามากที่สุด ค่าน้อยที่สุด ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน เป็นต้น จึงจะใช้ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของข้อมูลสอนในการแปลงช่วงข้อมูลทดสอบ

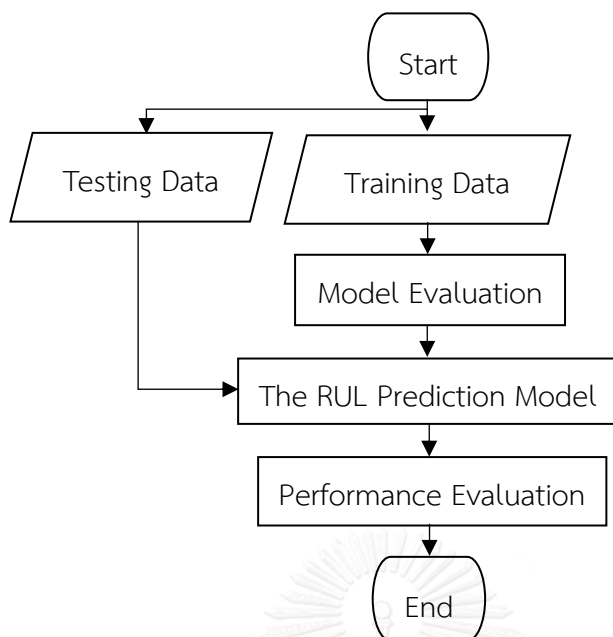
3.4 การทำนายอายุการใช้งานคงเหลือ

ในการดำเนินการทดลอง ผู้วิจัยใช้โอเพนซอร์สในภาษา Python และเฟรมเวิร์ก Tensorflow ในการพัฒนาโปรแกรมนิรอลเน็ตเวิร์กคอนโวลูชันสำหรับการทดลอง นอกจากนี้ได้มีการใช้ไลบรารีอื่น ๆ เพิ่มเติม ได้แก่ scikit, statsmodels, numpy, panda, sklearn และ math

สำหรับการทดลองได้มีการนำข้อมูลต่าง ๆ มาทำนายอายุการใช้งานคงเหลือและผลเฉลยของชุดนั้น ๆ ด้วยวิธีการพื้นฐานทางสถิติและวิธีการในงานวิจัย 2 วิธี ได้แก่ การทำนายแบบตัวแปรเดียว (Univariate Forecasting) และการทำนายแบบหลายตัวแปร (Multivariate Forecasting) มีรายละเอียดดังนี้

3.4.1 การทำนายแบบตัวแปรเดียว (Univariate Forecasting)

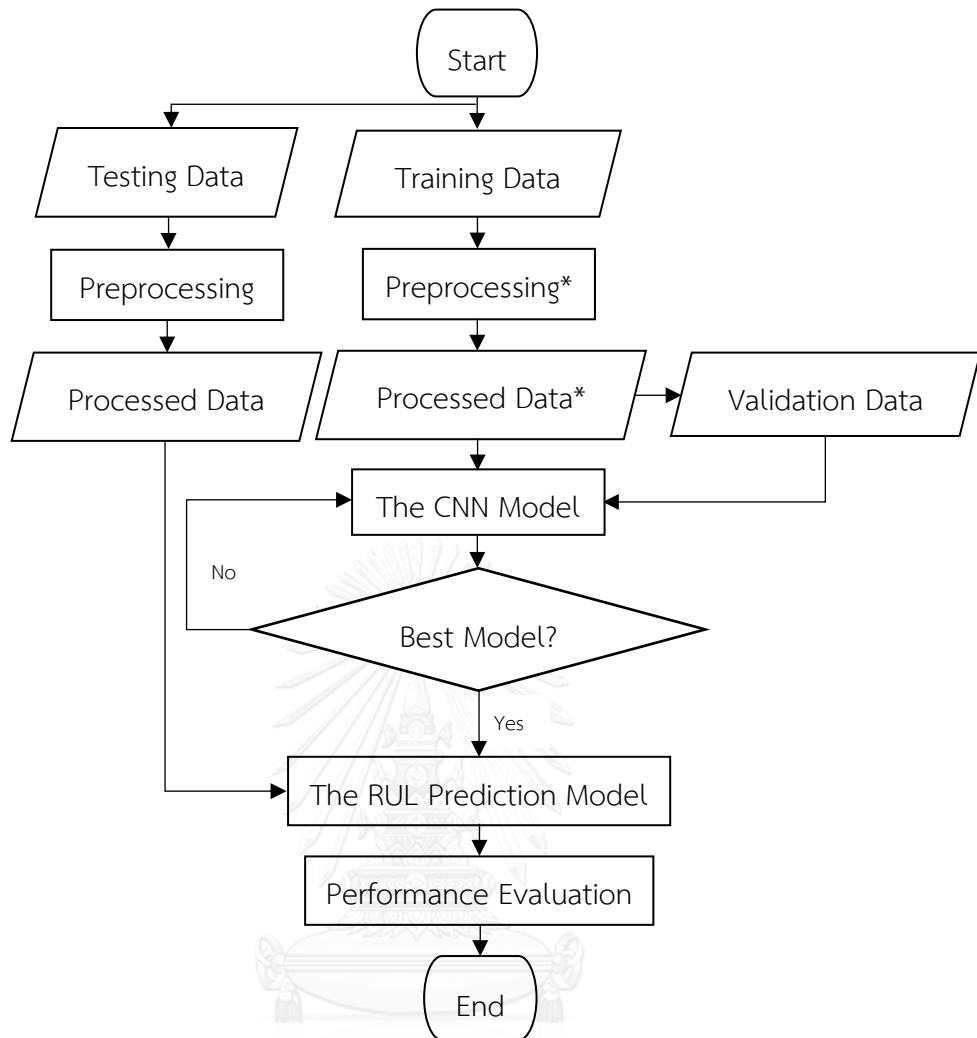
ได้ทำการทดลองโดยคำนวณสมการแบบจำลองที่เหมาะสมด้วยวิธีการอาร์มีมาและการปรับเรียบแบบเอกซ์โพเนนเชียล ในการพยากรณ์ค่าอายุการใช้งานคงเหลือและค่าผลเฉลยของทุก ๆ ชุดข้อมูล จากนั้นนำไปเปรียบเทียบกับประสิทธิภาพ โดยคำนวณค่าความคลาดเคลื่อนที่ได้จากค่าที่ทำนายได้กับค่าจริง การทดลองมีกระบวนการดังรูปที่ 18 ซึ่งประกอบด้วย การนำข้อมูลเข้า การประมาณแบบจำลอง การหาแบบจำลอง และการประเมินประสิทธิภาพของแบบจำลอง



รูปที่ 18 กระบวนการทำนายอายุการใช้งานคงเหลือแบบตัวแปรเดียว

3.4.2 การทำนายแบบหลายตัวแปร (Multivariate Forecasting) ด้วยการใช้โครงข่ายประสาทเทียมคอนโวลูชันเชิงลึก

ได้ทำการทดลองกับแบบจำลองที่มีโครงสร้างคอนโวลูชันเชิงลึกพื้นฐานได้แก่ LeNet-5, AlexNet และ VGG-16 นอกจากนี้ยังได้ทำการเพิ่มประสิทธิภาพเทคนิคต่าง ๆ ในแบบจำลองโครงสร้างคอนโวลูชันเชิงลึกพื้นฐานในการทดลอง เพื่อให้เห็นถึงความแตกต่างของประสิทธิภาพที่ดีขึ้นกว่าเดิม รูปที่ 19 แสดงกระบวนการในการทำนายอายุการใช้งานคงเหลือของเครื่องจักรด้วยโครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึก



รูปที่ 19 กระบวนการทำงานของแบบจำลองการทำนายอายุการใช้งานคงเหลือของเครื่องจักรด้วย
โครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึก

โดยกระบวนการที่มีเครื่องหมาย “*” ในรูปที่ 19 แสดงถึงกระบวนการที่แตกต่างจากการทำนายแบบตัวแปรเดียวในรูปที่ 18 ได้แก่ การประมวลผลก่อน (Preprocessing) และข้อมูลที่ได้หลังจากประมวลผลก่อน (Processed Data)

บทที่ 4

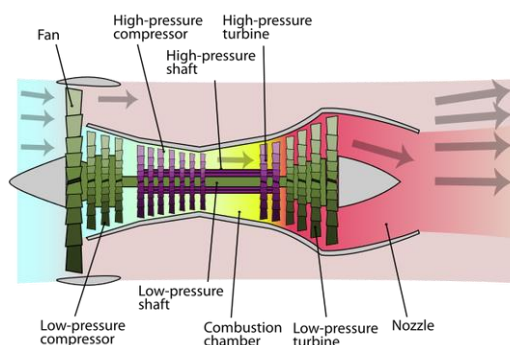
การทดลองและผลการทดลอง

4.1 ชุดข้อมูลที่ใช้ในการทดลอง

เป็นข้อมูลแบบอนุกรมเวลาที่เก็บจากการจำลองการทำงาน หรือการบันทึกค่าจริงของอุปกรณ์ต่างๆ โดยในการวิจัยนี้ทำการทดลองบนชุดข้อมูลที่นำเชื่อถือและนิยมใช้ทั้งหมด 14 ชุด ประกอบด้วย ชุดข้อมูลเผยแพร่ทั่วไป 13 ชุดและชุดข้อมูลจากเหตุการณ์จริงในโรงไฟฟ้า 1 ชุด ทั้งนี้ เพื่อเป็นการยืนยันประสิทธิภาพของแบบจำลองที่เสนอ ชุดข้อมูลแต่ละชุดมีรายละเอียดดังนี้

1. ข้อมูลจำลองของเครื่องเทอร์โบแฟน (Turbofan Engine Degradation Simulation)

ถูกจัดทำโดยศูนย์ความเป็นเลิศทางด้านการทำงาน (The Prognostics Center of Excellence หรือ PCoE) ที่ศูนย์วิจัยนาซาเอมส์ (NASA Ames Research Center) ด้วยเครื่องจำลองระบบการขับเคลื่อนทางการบินเชิงพาณิชย์ (Commercial Modular Aero-Propulsion System Simulation หรือ C-MAPSS) ซึ่งเป็นการจำลองการทำงานในส่วนของใบพัดมีลักษณะดังรูปที่ 12 ประกอบด้วย 4 ชุดข้อมูลย่อย นอกจากนี้ชุดข้อมูลดังกล่าวยังถูกนำไปใช้ในการแข่งขัน The Prognostics Challenge Competition ที่ The International Conference on Prognostics and Health Management (PHM08) โดยรายละเอียดของชุดข้อมูลข้างต้นได้ถูกบันทึกภายใต้เงื่อนไขในการทำงาน ค่าจุดวัดและอาการผิดปกติที่แตกต่างกัน ซึ่งในชุดข้อมูลนี้จะมีข้อมูลรบกวนปะปนอยู่ด้วย โดยมีรายละเอียดตัวแปรและสถิติดังตารางที่ 6 7 และ 8 ตามลำดับ



รูปที่ 20 ส่วนประกอบของเทอร์โบแฟน

[ที่มา: <https://datascientistinabox.com/2015/11/30/predictive-modeling-of-turbofan-degradation-part-1/> Accessed: December 25, 2016]

ตารางที่ 5 แสดงชื่อแอตทริบิวต์และคำอธิบายของชุดข้อมูลเทอร์โบแฟน

ลำดับที่	แอตทริบิวต์	คำอธิบาย
1	Unit no.	the engine number
2	Time	the operational cycle number
3	opt.set1	the three operating settings
4	opt.set2	
5	opt.set3	
6	sensor1	the 21 sensor values
7	Sensor2	
...	...	
26	Sensor21	

ตารางที่ 6 แสดงชื่อตัวแปรและคำอธิบายจุดวัดเพิ่มเติมของชุดข้อมูลเทอร์โบแฟน [11]

ลำดับที่	คำย่อ	คำอธิบาย	หน่วยวัด
1	T2	Total temperature at fan inlet	$^{\circ}R$
2	T24	Total temperature at LPC outlet	$^{\circ}R$
3	T30	Total temperature at HPC outlet	$^{\circ}R$
4	T50	Total temperature at LPT outlet	$^{\circ}R$
5	P2	Pressure at fan inlet	<i>psia</i>
6	P15	Total pressure in bypass-duct	<i>psia</i>
7	P30	Total pressure at HPC outlet	<i>psia</i>
8	Nf	Physical fan speed	<i>rpm</i>
9	Nc	Physical core speed	<i>rpm</i>
10	Epr	Engine Pressure ratio	-
11	Ps30	Static pressure at HPC outlet	<i>psia</i>
12	Phi	Ratio of fuel flow to Ps30	<i>pps/psi</i>
13	NRf	Corrected fan speed	<i>rpm</i>
14	NRc	Corrected core speed	<i>rpm</i>
15	BPR	Bypass ratio	-
16	farB	Burner fuel-air ratio	-

ลำดับที่	คำย่อ	คำอธิบาย	หน่วยวัด
17	htBleed	Bleed enthalpy	-
18	Nf_dmd	Demanded fan speed	<i>rpm</i>
19	PCNfR_d md	Demanded corrected fan speed	<i>rpm</i>
20	W31	HPT coolant bleed	<i>lpm/s</i>
21	W32	LPT coolant bleed	<i>lbm/s</i>

$^{\circ}R$ refers to The Rankine temperature scale.

psia refers to Pounds per square inch absolute.

rpm refers to Revolutions per minute.

pps refers to Pulse per second.

psi refers to Pounds per square inch.

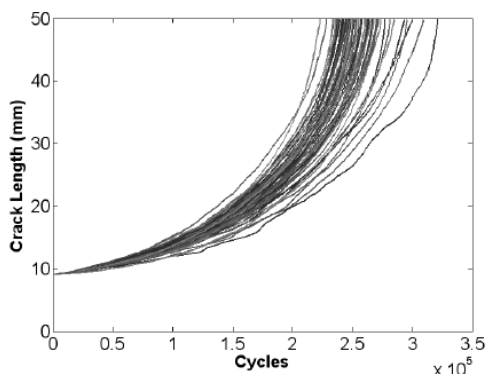
lbm/s refers to Pound mass per second.

ตารางที่ 7 แสดงจำนวนเครื่องเทอร์โบแฟนที่ใช้ในการจำลองการทำงานของแต่ละชุดข้อมูลย่อย

ชุดข้อมูล	C-MAPSS				PHM08
	Data1	Data2	Data3	Data4	
Training	100	260	100	249	218
Testing	100	259	100	248	218

2. ข้อมูลรอยแตกของโลหะ (Virkler)

เป็นชุดข้อมูลการเพิ่มขนาดของรอยแตกเมื่อโลหะมีค่าความเครียดแบบสม่ำเสมอในอดีตเป็นที่นิยมใช้สำหรับการวิเคราะห์ทางสถิติ [12] ได้จัดทำและรวบรวมขึ้นโดย Virkler และคณะ [13] การทดสอบความกว้างของโลหะผสมอลูมิเนียมได้ทำการตรวจสอบการขยายตัวของรอยแตกซึ่งมีหน่วยการวัดเป็นมิลลิเมตร ในชุดข้อมูลประกอบด้วยข้อมูลการขยายตัวของรอยแตก 68 ชุด และประกอบด้วยจุดวัด 164 จุด รูปที่ 21 แสดงความยาวของรอยแตกที่เปลี่ยนไปตามจำนวนรอบในขณะที่ทำการทดสอบ



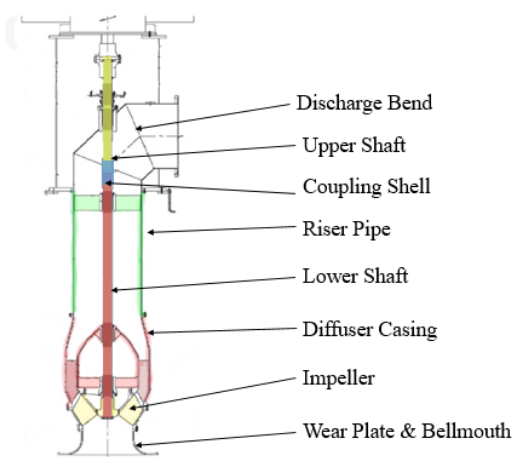
รูปที่ 21 ตัวอย่างการขยายตัวของรอยแตกในสภาพการรับน้ำหนักเดียวกัน

[ที่มา:

https://www.researchgate.net/publication/232709378_Major_Challenges_in_Prognostics_Study_on_Benchmarking_Prognostics_Datasets Accessed: April 19, 2017]

3. เครื่องปั๊มน้ำหมุนเวียน (Circulating Water Pump; CWP)

ในระบบเครื่องกำเนิดไฟฟ้ากังหันไอน้ำ (Steam Turbine Generator) มีตัวแลกเปลี่ยนความร้อนที่จ่ายผ่านเครื่องควบแน่น (Condenser) ซึ่งเป็นหน้าที่หลักของการหมุนเวียนน้ำ ปั๊มน้ำหมุนเวียน ซึ่งส่วนประกอบต่าง ๆ ดังแสดงในรูปที่ 14 จะทำหน้าที่สูบน้ำหมุนเวียน เพื่อดูดซับความร้อนจากระบบ และนำกลับไปยังหอทำความเย็น (Cooling Tower)



รูปที่ 22 โครงสร้างส่วนประกอบ CWP

ข้อมูลจุดวัดของเครื่องปั๊มน้ำหมุนเวียนที่นำมาทำการทดลองประกอบด้วยจุดวัดแปดจุด ได้แก่ ค่ากระแสไฟฟ้าหนึ่งจุด ค่าอุณหภูมิสามจุด และค่าการสั่นสะเทือนในแนวแกนตั้ง

และแกนนอนรวมกันสี่จุด ข้อมูลจุดวัดของอุปกรณ์ถูกเก็บรวบรวมจากเครื่องมือการจัดการข้อมูลเรียลไทม์ (Real Time) ที่เรียกว่าระบบพีไอ (PI) ซึ่งใช้ในการเก็บรวบรวมข้อมูลประวัติการวิเคราะห์ การนำเสนอและแสดงผล เพื่อใช้ในการบำรุงรักษาโรงไฟฟ้า ค่าจุดวัดถูกเก็บรวบรวมเป็นรายชั่วโมงต่อเนื่องกันเป็นเวลาหนึ่งปี จุดรับสัญญาณหนึ่ง ๆ จะมีค่า 8,760 (24×365) โดยทำการรวบรวมค่าเซ็นเซอร์แปดจุด สัญญาณเซ็นเซอร์จะมีค่าทั้งหมด 61,320 ค่า ข้อมูลเซ็นเซอร์ของอุปกรณ์ถูกเก็บรวบรวมจากเครื่องมือการจัดการข้อมูลแบบเรียลไทม์ที่เรียกว่าระบบพีไอ สำหรับเครื่องปัมน้ำหมุนเวียนในการทดลองตามวิธีที่นำเสนอ ซึ่งจัดเป็นชุดข้อมูลที่มีลักษณะหลายตัวแปรตามช่วงเวลา

สำหรับชุดข้อมูลที่นำมาใช้การทดสอบเป็นการรวบรวมข้อมูลจากโรงงานแห่งหนึ่งในประเทศไทย โดยนำข้อมูลช่วงเวลาประวัติการตรวจสอบความผิดปกติและการซ่อมอุปกรณ์ในระบบงานบำรุงรักษาจากระบบบริหารจัดการองค์กร (Enterprise Resource Planning; ERP) มาผสานกับข้อมูลจุดวัดแต่ละจุดตามช่วงเวลาการทำงานของอุปกรณ์ที่ถูกบันทึกไว้ด้วยระบบพีไอ ดังแสดงในตารางที่ 9

ตารางที่ 8 ข้อมูลประวัติเหตุการณ์ ช่วงเวลา และรายละเอียดที่อุปกรณ์ถูกตรวจสอบ ซ่อมแซม หรือมี
การทำงานผิดปกติในปี 2558

เหตุการณ์	ช่วงเวลา (ชั่วโมง)	รายละเอียด
1	98	Check Temperature
2	122	Electric Part
3	263	Vibration
4	61	Time-based Inspection
5	137	Misalignment
6	303	Vibration
7	385	Bent Shaft
8	29	Vibration
9	125	Misalignment
10	156	Electric Part
11	13	Check Temperature
12	170	Misalignment

4. ข้อมูลโรงไฟฟ้าพลังความร้อนร่วม (Combined Cycle Power Plant; CCPP)

ชุดข้อมูลนี้มีข้อมูล 9,568 จุดที่รวบรวมได้จากโรงไฟฟ้าพลังความร้อนร่วมห้าปี (พ.ศ. 2549-2553) ห้าชุดข้อมูล เมื่อโรงไฟฟ้าได้รับการตั้งค่าให้ทำงานเต็มกำลัง เพื่อทำนายพลังงานไฟฟ้าต่อชั่วโมง (EP) ของโรงไฟฟ้า ชุดข้อมูลมีรายละเอียดดังตารางที่ 10

ตารางที่ 9 แสดงแอตทริบิวต์ หน่วย และช่วงของข้อมูล CCPP

แอตทริบิวต์	หน่วย	ค่าน้อย-มากที่สุดของข้อมูล
อุณหภูมิ (T)	°C	1.81 - 37.11
ความดันบรรยากาศ (AP)	milibar	992.89 - 1033.30
ความชื้นสัมพัทธ์ (RH)	%	25.56 - 100.16
สัทยอากาศ (V)	cm Hg	25.36 - 81.56
พลังงานไฟฟ้าต่อชั่วโมง (EP)	MW	420.26 - 495.76

โรงไฟฟ้าพลังความร้อนร่วมประกอบด้วยกังหันแก๊ส (GT) กังหันไอน้ำ (ST) และเครื่องกำเนิดไอน้ำร้อน ในโรงไฟฟ้าพลังความร้อนร่วมนั้น ไฟฟ้าจะถูกสร้างขึ้นโดยกังหันก๊าซ และกังหันไอน้ำซึ่งรวมกันในรอบเดียวและจะถูกถ่ายโอนจากกังหันหนึ่งไปยังอีกกังหัน ในขณะที่เกิดสัทยอากาศขึ้นและมีผลกระทบต่อกังหันไอน้ำ และตัวแปรอีกสามมีผลต่อประสิทธิภาพการทำงานของกังหันแก๊สและข้อมูลการใช้พลังงาน (Energy)

5. ข้อมูลการใช้พลังงาน (Energy)

ชุดข้อมูลนี้ถูกรวบรวมเป็นระยะเวลาประมาณ 4.5 เดือน สภาวะอุณหภูมิและความชื้นของห้องถูกสังเกตการณ์ผ่านเครือข่ายเซนเซอร์ไร้สายซิกบี (ZigBee) โดยที่โหนดไร้สายแต่ละตัวจะส่งผ่านค่าอุณหภูมิและความชื้นประมาณ 3.3 นาที จากนั้นข้อมูลไร้สายถูกเฉลี่ยสำหรับคาบเวลา 10 นาที ข้อมูลพลังงานถูกบันทึกทุก ๆ 10 นาทีด้วยเครื่องวัดจากหน่วยงานของสนามบินเชฟในเบลเยียม (Chievres Airport, Belgium) ซึ่งเป็นสถานที่ที่ใกล้ที่สุด รายละเอียดของชุดข้อมูลดังแสดงในตารางที่ 10

ตารางที่ 10 แสดงชื่อแอตทริบิวต์คำอธิบาย และหน่วยการวัดของชุดข้อมูลการใช้พลังงาน

แอตทริบิวต์	คำอธิบาย	หน่วยการวัด
date time	year-month-day hour:minute:second	Time point

แอตทริบิวต์	คำอธิบาย	หน่วยการวัด
Appliances	energy use	Wh
lights	energy use of light fixtures in the house	Wh
T1	Temperature in kitchen area	Celsius
RH_1	Humidity in kitchen area	%
T2	Temperature in living room area	Celsius
RH_2	Humidity in living room area	%
T3	Temperature in laundry room area	Celsius
RH_3	Humidity in laundry room area	%
T4	Temperature in office room	Celsius
RH_4	Humidity in office room	%
T5	Temperature in bathroom	Celsius
RH_5	Humidity in bathroom	%
T6	Temperature outside the building (north side)	Celsius
RH_6	Humidity outside the building (north side)	%
T7	Temperature in ironing room	Celsius
RH_7	Humidity in ironing room	%
T8	Temperature in teenager room 2	Celsius
RH_8	Humidity in teenager room 2	%
T9	Temperature in parents room	Celsius
RH_9	Humidity in parents room	%
To	Temperature outside from Chievres weather station	Celsius
Pressure	Pressure from Chievres weather station	mm Hg
RH_out	Humidity outside from Chievres weather station	%
Wind speed	Wind speed from Chievres weather station	m/s
Visibility	Visibility from Chievres weather station	km
Tdewpoint	Tdewpoint from Chievres weather station	Â°C
rv1	Random variable 1	nondimensional
rv2	Random variable 2	nondimensional

6. ข้อมูลมลภาวะในปักกิ่ง (Beijing)

ชุดข้อมูลรายชั่วโมงนี้ประกอบด้วยข้อมูลมลภาวะ (particle Pollution; PM2.5) ของสถานทูตสหรัฐอเมริกาในปักกิ่งและข้อมูลทางอุตุนิยมวิทยาจากท่าอากาศยานนานาชาติปักกิ่ง

ช่วงเวลาของข้อมูลอยู่ระหว่างวันที่ 1 มกราคม 2553 ถึง 31 ธันวาคม 2557 ข้อมูลที่หายไปจะแสดงเป็น "NA" ชุดข้อมูลมีรายละเอียดดังตารางที่ 12

ตารางที่ 11 แสดงแอตทริบิวต์ คำอธิบาย และหน่วยการวัดของชุดข้อมูล Beijing

แอตทริบิวต์	คำอธิบาย	หน่วย
No	row number	-
year	year of data in this row	-
month	month of data in this row	-
day	day of data in this row	-
hour	hour of data in this row	-
pm2.5	PM2.5 concentration ()	ug/m ³
DEWP	Dew Point ()	°f
TEMP	Temperature (°f)	°f
PRES	Pressure ()	hPa
cbwd	Combined wind direction	-
lws	Cumulated wind speed ()	m/s
ls	Cumulated hours of snow	hours
lr	Cumulated hours of rain	hours

4.2 การดำเนินการทดลอง

ในการทำนายอายุการใช้งานคงเหลือของเครื่องจักรแบ่งออกเป็นกระบวนการย่อยได้แก่ การนำชุดข้อมูลเข้า ซึ่งแบ่งเป็นชุดข้อมูลสำหรับสอนและทดสอบ การประมวลผลก่อน (Data Preprocessing) แบบจำลองสำหรับการทำนาย การตรวจสอบแบบจำลอง และการวัดประสิทธิภาพ

4.2.1 การสอน (Training)

การสอนแบบจำลองถือเป็นขั้นตอนหลักที่สำคัญ วิทยาพจน์นี้ได้เลือกตัดจำนวนแถวของข้อมูลนำเข้าที่ละ 15 แถว หรือ 15 ช่วงเวลา ส่วนจำนวนหลักขึ้นอยู่กับจำนวนจุดวัดของข้อมูลเข้าในแต่ละชุด โดยกำหนดขนาดของแต่ละรุ่น (Batch Size) เป็น 10 และจำนวนรุ่นของการผลิตในแต่ละรอบของการคำนวณ (Number batches per epoch) อยู่ที่ 800

สำหรับการเลือกใช้แบบจำลอง ได้มีเงื่อนไขในแต่ละรอบการเรียนรู้ (Epoch) เพื่อเปรียบเทียบค่าความคลาดเคลื่อนบนชุดข้อมูลตรวจสอบ (Validation Dataset) ในรอบนั้น ๆ กับรอบที่ผ่านมาทั้งหมด และเลือกแบบจำลองที่มีค่าความคลาดเคลื่อนน้อยที่สุดมาใช้สอนในรอบต่อไป จากนั้นจะมีเงื่อนไขให้หยุดการเรียนรู้ของแบบจำลองเมื่อค่าความคลาดเคลื่อนบนชุดข้อมูลตรวจสอบไม่มีแนวโน้มเพิ่มขึ้นหรือมีค่าคงที่

4.2.2 การทดสอบ (Testing)

สำหรับการทดสอบจะใช้แบบจำลองที่มีประสิทธิภาพสูงสุดโดยวัดจากค่า RMSE บนข้อมูล ซึ่งเป็นค่าที่บอกถึงค่าความคลาดเคลื่อนในการทำนายข้อมูล สำหรับการตรวจสอบความถูกต้องของแบบจำลอง (Data Validation) เพื่อนำแบบจำลองที่ได้ไปใช้กับข้อมูลทดสอบต่อไป

ในการประเมินแบบจำลองในแต่ละรอบการเรียนรู้ ได้ทำการทดสอบบนชุดข้อมูลสำหรับตรวจสอบที่ได้แบ่งมาจากชุดข้อมูลสอนในอัตราส่วน 70 ต่อ 30 หลังจากนั้นทำการเปรียบเทียบค่า RMSE เพื่อเลือกแบบจำลองที่มีประสิทธิภาพดีกว่ามาทำการทดสอบในรอบต่อไป ตารางที่ 12 แสดงถึงจำนวนแอดทริบิวต์ ข้อมูลสอน และข้อมูลทดสอบที่มีในแต่ละชุดข้อมูลในการดำเนินการทดลอง

ตารางที่ 12 แสดงจำนวนแอดทริบิวต์ ข้อมูลสอน และข้อมูลทดสอบของแต่ละชุดข้อมูล

ชุดข้อมูล	แอดทริบิวต์	ข้อมูลสอน	ข้อมูลทดสอบ
C-MAPSS1	24	20,631	13,096
C-MAPSS2	24	53,759	33,991

ชุดข้อมูล	แอตทริบิวต์	ข้อมูลสอน	ข้อมูลทดสอบ
C-MAPSS3	24	24,720	16,596
C-MAPSS4	24	61,249	41,214
PHM08	24	45,917	22,144
Virkler	68	164	110
CWP	8	8,834	5,931
CCPP-2006	4	9,568	6,424
CCPP-2007	4	9,568	6,424
CCPP-2008	4	9,568	6,424
CCPP-2009	4	9,568	6,424
CCPP-2010	4	9,568	6,424
Energy	27	19,735	13,249
Beijing	7	43,824	29,422

4.2.3 การกำหนดค่าต่าง ๆ ของแบบจำลอง (Model Configuration)

ในการดำเนินการทดลอง ผู้วิจัยได้ทำการกำหนดโครงสร้าง ค่าพารามิเตอร์ต่าง ๆ ในกับแบบจำลองในแต่ละแบบได้แก่ ขนาด มิติ และฟังก์ชันต่าง ๆ สำหรับโครงสร้างเน็ตเวิร์กที่ใช้ในการดำเนินการทดลองดังตารางที่ 13 โดยได้อธิบายถึงรายละเอียดโครงสร้างของเน็ตเวิร์กในการทำนายแบบหลายตัวแปร ได้แก่ วิธีเพอร์เซ็ปตรอนหลายชั้น (Multi-Layer Perceptron; MLP) วิธี LeNet-5 Lite วิธี AlexNet Lite วิธี VGG-16 Lite วิธี Enhanced LeNet-5 วิธี Enhanced LeNet-5-2 และวิธี Enhanced VGG-16 ซึ่งโครงสร้างสุดท้ายเป็นโครงสร้างที่นำเสนอในงานวิจัย

วิธีการที่ต่อท้ายด้วยคำว่า “Lite” เป็นวิธีการดั้งเดิม ส่วนวิธีการขึ้นต้นด้วยคำว่า “Enhanced” เป็นวิธีการแบบที่ปรับเปลี่ยน เพิ่มเติมเทคนิคต่าง ๆ เข้าไปในแบบจำลองโครงสร้างเน็ตเวิร์กคอนโวลูชันเชิงลึกพื้นฐาน

ตารางที่ 13 แสดงขนาดและมิติที่ใช้ในโครงสร้างของแบบจำลองแบบต่าง ๆ

MLP	LeNet-5 Lite	AlexNet Lite	VGG-16 Lite	Enhanced LeNet-5	Enhanced LeNet-5-2	Enhanced VGG-16
10 hidden nodes	C = 2x2 P = 1x2	C = 2x2 C = 1x2 P = 1x2	C = hx2 C = 1x2 P = 1x2	C = 2x2 P = 1x2	C = 2x2 P = 1x2	C = 2x2 C = 2x2 P = 1x2
	C = 1x3 P = 1x2	C = 1x2 P = 1x2	C = 1x2 C = 1x2 P = 1x2	C = 1x3 P = 1x2	C = 1x3 P = 1x2	C = 2x2 C = 2x2 P = 2x2
			C = 1x1 C = 1x1 C = 1x1 P = 1x1			C = 2x2 C = 2x2 C = 2x2 P = 1x1
			C = 1x1 C = 1x1 C = 1x1 P = 1x1			C = 2x1 C = 2x1 C = 2x1 P = 2x1
			C = 1x1 C = 1x1 C = 1x1 P = 1x1			C = 2x1 C = 1x1 C = 1x1 P = 1x1
	F = ReLU	F = ReLU	F = ReLU F = ReLU F = ReLU	F = ELU	F = ELU	F = ELU F = ELU F = ELU
Sigmoid	ReLU	ReLU	ReLU	ELU	ELU	ELU

ในการทดลองที่ใช้โครงสร้างที่เป็นเน็ตเวิร์กคอนโวลูชัน จะมีการใช้เทคนิคเพื่อเพิ่มประสิทธิภาพในการทำนายต่าง ๆ ดังแสดงในตารางที่ 14 ซึ่งเทคนิคที่มีการนำมาใช้ได้แก่ การดรอปเอาต์ การทำเรกูลารีเซชันแบบแอลสอง และเทคนิคที่ใช้ในการหาค่าที่ดีที่สุด โดยเครื่องหมาย “✓” หมายถึงมีการใช้เทคนิคนั้น ๆ ในโครงสร้าง และเครื่องหมาย “✗” หมายถึงไม่มีการนำเทคนิคนั้น ๆ ในมาใช้

นอกจากนี้ จะเห็นว่าเทคนิคที่ใช้เหมือนกันในทุกแบบจำลองคือ การหาค่าเหมาะสมที่สุดได้ใช้ เกรเดียนแบบปรับตัวได้ เนื่องจากการใช้สโตแคสติกเกรเดียนเดสเซนชันนั้นไม่สามารถปรับค่าน้ำหนัก เพื่อลดค่าของฟังก์ชันต้นทุนให้น้อยที่สุดได้ จึงทำให้แบบจำลองไม่สามารถทำนายค่าอายุการใช้งาน คงเหลือหรือผลเฉลยในชุดข้อมูลที่ทำการทดลองได้

ตารางที่ 14 เทคนิคที่ใช้ในแต่ละการทดลองการทำนายด้วยโครงข่ายประสาทเทียม

Methods	Dropout	L ² regularization	Optimization	
The standard ANN	MLP	✗	✗	✗
The standard CNN	LeNet-5 Lite	✗	✗	AdaGrad
	AlexNet Lite	✗	✗	AdaGrad
	VGG-16 Lite	✗	✗	AdaGrad
The enhanced CNN	Enhanced LeNet-5	✓	✓	AdaGrad
	Enhanced LeNet-5-2	✗	✓	AdaGrad
	Enhanced VGG-16	✗	✓	AdaGrad

4.3 ผลการทดลอง

การทดลองเป็นการเปรียบเทียบประสิทธิภาพในการทำนายค่าอายุการใช้งานคงเหลือของ อุปกรณ์ในชุดข้อมูล 14 ชุดด้วยวิธีการทำนายสองวิธีได้แก่ การทำนายแบบตัวแปรเดียวและการทำนายแบบหลายตัวแปร โดยแต่ละวิธีได้ทำการทดลองด้วยแบบจำลองแบบต่าง ๆ เพื่อนำมา เปรียบเทียบกับแบบจำลองในการทำนายข้อมูลแบบหลายตัวแปรที่น่าเสนอ และได้ทำการประเมินผล การทดลองเป็นค่า RMSE ดังตารางที่ 15 และ 16 ทั้งนี้ตารางที่ 15 เป็นค่า RMSE m ได้จากการ ทดลองกับชุดข้อมูลทางด้านการทำนายอายุการใช้งานคงเหลือของเครื่องจักร และตารางที่ 16 เป็นค่า RMSE ที่ได้จากการทดลองแบบจำลองต่าง ๆ กับข้อมูลอนุกรมเวลาหลายตัวแปรที่เป็นที่นิยมและใช้ กันแพร่หลาย โดยวิธีการในแต่ละชุดข้อมูลที่ให้ค่าความคลาดเคลื่อนของการทำนายน้อยที่สุดจะแทน ด้วยเครื่องหมาย “*”

ทั้งนี้ได้แสดงผลการทดลองในรูปแบบกราฟแท่ง โดยที่แกนอนแสดงชุดข้อมูลทั้ง 14 ชุดและ แกนตั้งแสดงค่า RMSE ของการทดลองด้วยแบบจำลองทั้งเก้าแบบ จากกราฟดังกล่าวจะเห็นได้ว่า แบบจำลองที่ให้ค่าความคลาดเคลื่อนในการทำนายน้อยที่สุดคือแบบจำลอง Enhanced VGG-16 ที่

นำเสนอ โดยที่ตัวเลขบนแท่งกราฟในแต่ละชุดข้อมูลแสดงถึงค่า RMSE ของแบบจำลอง Enhanced VGG-16 ในรูปที่ 23 และ 24 ภาคผนวก ข

ตารางที่ 15 ค่า RMSE จากการทดลองแบบจำลองกับชุดข้อมูลด้านการทำนายอายุการใช้งานคงเหลือ

Method		C-MAPSS1	C-MAPSS2	C-MAPSS3	C-MAPSS4	PHM08	Virkler	CWP
Univariate Forecasting	ARIMA	6.93	15.59	15.70	16.36	15.94	30.27	19.05
	Exponential Smoothing	7.16	16.56	16.66	17.37	16.87	31.94	20.15
Multivariate Forecasting	MLP	19.04	27.08	31.69	36.49	31.11	17.60	30.27
	LeNet-5 Lite	21.67	29.37	48.86	37.53	36.42	14.63	44.21
	AlexNet Lite	17.06	15.77	17.76	18.08	15.52	10.84	41.89
	VGG-16 Lite	17.19	26.21	28.47	52.66	22.19	5.91	19.92
	Enhanced LeNet-5	21.98	30.18	29.15	32.04	35.38	19.13	51.77
	Enhanced LeNet-5-2	19.15	25.30	21.54	29.03	34.59	16.96	47.52
	Enhanced VGG-16	5.26*	8.57*	9.65*	8.76*	7.14*	2.03*	16.17*

ตารางที่ 16 ค่า RMSE จากการทดลองแบบจำลองกับชุดข้อมูลอนุกรมเวลาหลายตัวแปร

Method		CCPP-2006	CCPP-2007	CCPP-2008	CCPP-2009	CCPP-2010	Energy	Beijing
Univariate Forecasting	ARIMA	16.97	16.64	17.03	17.04	17.13	84.49	94.16
	Exponential Smoothing	17.30	16.99	17.5	17.44	17.52	91.27	39.76
Multivariate Forecasting	MLP	26.12	18.31	17.44	17.76	17.63	149.18	39.24
	LeNet-5 Lite	17.3	17.06	17.33	17.49	17.27	72.93	55.12
	AlexNet Lite	17.03	16.98	16.09	16.7	17.11	65.77	31.17
	VGG-16 Lite	17.28	17.09	17.15	17.18	17.13	105.08	65.23
	Enhanced LeNet-5	50.68	17.66	17.37	17.98	17.64	101.29	69.39
	Enhanced LeNet-5-2	23.99	18.27	18.04	17.82	19.55	97.95	66.55
	Enhanced VGG-16	3.53*	4.20*	3.97*	3.78*	3.84*	4.64*	3.72*

4.4 การวิเคราะห์ผลการทดลอง

จากผลการทดลองข้างต้นในตารางที่ 15 และ 16 แสดงให้เห็นว่า วิธีการที่นำเสนอคือ วิธี Enhanced VGG-16 ให้ค่าความคลาดเคลื่อนของการทำนายน้อยที่สุดบนชุดข้อมูลทั้ง 14 ชุด กล่าวคือค่าความคลาดเคลื่อนในการทำนายของวิธีการที่นำเสนอมีน้อยกว่าวิธีการแบบอื่น ดังนั้นในการทำนายอายุการใช้งานคงเหลือของเครื่องจักรซึ่งข้อมูลเข้ามีลักษณะเป็นอนุกรมเวลาแบบหลายตัวแปรจึงเหมาะกับแบบจำลอง Enhanced VGG-16 ที่นำเสนอ เนื่องจากค่าความคลาดเคลื่อนที่ได้จากการทดลองบนชุดข้อมูลที่มีลักษณะดังกล่าวด้วยวิธีการที่นำเสนอ มีค่าน้อยที่สุดเมื่อเทียบกับแบบจำลองที่มีลักษณะโครงสร้างพื้นฐานและเป็นที่นิยมทั้งหมดแบบ อีกทั้งผลการทดลองในตารางที่ 16 ยังแสดงให้เห็นว่า สามารถนำแบบจำลองที่นำเสนอไปทดลองกับข้อมูลชุดอื่น ๆ ที่ไม่ใช่ชุดข้อมูลเกี่ยวกับการทำนายอายุการใช้งานคงเหลือได้ เนื่องจากมีลักษณะเป็นอนุกรมหลายตัวแปร เช่นเดียวกันและยังคงมีประสิทธิภาพในการทำนายดีกว่าแบบจำลองอื่น ๆ ที่ได้ทำการทดลองอีกด้วย

นอกจากนี้ได้มีการคำนวณเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองเป็นคู่ ๆ ด้วยวิธีการพื้นฐานทางสถิติเรียกว่า Paired T Test เพื่อตรวจสอบว่าแบบจำลองที่นำเสนอมีประสิทธิภาพดีกว่าแบบจำลองอื่น ๆ อย่างมีนัยยะสำคัญหรือไม่ โดยคำนวณค่า T Value และ P Value ของแบบจำลองในแต่ละคู่ ดังแสดงในตารางที่ 16 และ 17 ภาคผนวก ค โดยเครื่องหมาย “*” หมายถึงแบบจำลองที่ทดสอบมีประสิทธิภาพดีกว่าอย่างมีนัยยะสำคัญ

บทที่ 5

สรุปการวิจัยและแนวทางการวิจัยในขั้นถัดไป

5.1 สรุปการวิจัย

ในวิทยานิพนธ์นี้ได้มุ่งเน้นการปรับโครงสร้างโครงข่ายประสาทเทียมให้เหมาะสมกับชุดข้อมูลหลายตัวแปรที่มีลักษณะเป็นช่วงเวลาโดยได้ดำเนินการตามขอบเขตดังนี้

1. ทำการทำนายอายุการใช้งานคงเหลือของเครื่องจักรที่มีข้อมูลนำเข้าเป็นอนุกรมเวลาแบบหลายตัวแปรด้วยเน็ตเวิร์กคอนโวลูชันเชิงลึก
2. เปรียบเทียบประสิทธิภาพเมื่อมีการปรับโครงสร้างโครงข่ายประสาทเทียมคอนโวลูชันเชิงลึกใหม่กับโครงสร้างแบบเดิม ได้แก่ LeNet-5 VGG-16 และ AlexNet
3. เปรียบเทียบประสิทธิภาพด้านโครงสร้างระหว่างวิธีการแบบเดิมกับวิธีการใหม่ที่มีการใช้เทคนิคต่าง ๆ ได้แก่

- การป้องกันการอิงกับข้อมูลมากเกินไปด้วยการดรอปเอาต์
- การลดความซับซ้อนของแบบจำลองด้วยการทำเรกูลาร์ไลเซชันแบบแอลสอง
- การหาค่าที่เหมาะสมที่สุดด้วยเกรเดียนแบบปรับตัวได้ ในการปรับอัตราการเรียนรู้และค่าน้ำหนักในแบบจำลอง เพื่อลดค่าความคลาดเคลื่อนในการทำนาย

ซึ่งเทคนิคต่าง ๆ ข้างต้นล้วนส่งผลทำให้แบบจำลองที่ได้มีประสิทธิภาพในการทำนายผลเฉลยดีขึ้นกว่าเดิมเมื่อเทียบกับวิธีการพื้นฐาน

สำหรับการทดลอง ผู้วิจัยได้แบ่งวิธีการออกเป็นสองวิธีการใหญ่ ๆ ได้แก่ วิธีการทำนายแบบตัวแปรเดียวและวิธีการทำนายแบบหลายตัวแปร ซึ่งวิธีการทำนายแบบตัวแปรเดียว ได้ทดลองทำนายด้วยวิธีการพื้นฐานทางสถิติ ได้แก่ วิธีอาร์มาและวิธีการปรับเรียบแบบเอกซ์โพเนนเชียล ส่วนวิธีการทำนายแบบหลายตัวแปร ได้ทดลองด้วยวิธีการโครงข่ายประสาทเทียม ได้แก่

1. วิธีเพอร์เซ็ปตรอนหลายชั้น (Multi-Layer Perceptron; MLP)
2. วิธี LeNet-5 Lite
3. วิธี AlexNet Lite
4. วิธี VGG-16 Lite
5. วิธี Enhanced LeNet-5
6. วิธี Enhanced LeNet-5-2
7. วิธี Enhanced VGG-16

โดยที่วิธีที่หนึ่ง ถือเป็นวิธีการที่ใช้โครงสร้างโครงข่ายประสาทเทียมแบบพื้นฐาน วิธีที่สองถึงสี่ เป็นวิธีที่ใช้โครงสร้างพื้นฐานของเน็ตเวิร์กคอนโวลูชัน และวิธีที่ห้าถึงเจ็ดเป็นวิธีการที่นำโครงสร้างของเน็ตเวิร์กคอนโวลูชันพื้นฐานมาเพิ่มเติมเทคนิคต่าง ๆ เพื่อเพิ่มประสิทธิภาพในการทำนายให้ดีขึ้น

จากผลการทดลองเพื่อเปรียบเทียบประสิทธิภาพในการทำนายอายุการใช้งานคงเหลือและค่าผลเฉลยในชุดข้อมูลอื่น ๆ ที่เป็นที่ยอมรับและเผยแพร่ทั่วไปด้วยวิธีที่นำเสนอคือ วิธี Enhanced VGG-16 ให้ค่าความคลาดเคลื่อนในการทำนายน้อยที่สุด กล่าวคือประสิทธิภาพดีที่สุด เมื่อทดลองด้วยชุดข้อมูลที่เป็นที่ยอมรับในด้านการทำนายอายุการใช้งานคงเหลือของเครื่องจักรและชุดข้อมูลอนุกรมเวลาแบบหลายตัวแปรทั้งหมด 14 ชุด ซึ่งวิธีการดังกล่าว เป็นการทำนายด้วยการใช้โครงข่ายประสาทเทียมคอนโวลูชันเชิงลึกแบบ VGG-16 ที่มีการใช้ฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้นเป็นฟังก์ชันกระตุ้น เนื่องจากคุณสมบัติในการคำนวณผลลัพธ์กรณีที่ค่าข้อมูลเข้าน้อยกว่าศูนย์ จะให้ผลลัพธ์ที่ลดลงตามฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้น ทำให้สามารถทำนายผลเฉลยได้แม่นยำมากขึ้น การทำเรกูลาร์ไลเซชันแบบแอลสอง เพื่อช่วยลดความซับซ้อนของแบบจำลอง และการหาค่าเหมาะสมที่สุดแบบเกรเดียนแบบปรับตัวได้

สำหรับวิธีการที่นำเสนอนี้สามารถนำไปใช้ได้กับการทำนายอายุการใช้งานคงเหลือของเครื่องจักร เพื่อประกอบการตัดสินใจและดำเนินการบำรุงรักษาแบบทำนายล่วงหน้าในโรงไฟฟ้าหรือโรงงาน นอกจากนี้เนื่องด้วยประสิทธิภาพของวิธีการดังกล่าวยังแสดงให้เห็นถึงความสามารถในการทำนายข้อมูลอนุกรมเวลาแบบหลายตัวแปรที่มีคุณสมบัติตัวแปรจำนวนมากได้อย่างแม่นยำ จึงสามารถนำไปประยุกต์ใช้ได้กับงานด้านอื่น ๆ ไม่เพียงแต่ด้านการบำรุงรักษาโรงไฟฟ้าเพียงอย่างเดียว

5.2 ข้อจำกัดของงานวิจัย

ในการทำนายอายุการใช้งานคงเหลือของอุปกรณ์ด้วยเน็ตเวิร์กคอนโวลูชันมีข้อจำกัดของงานวิจัยดังนี้

1. ชุดข้อมูลในการจำลองที่มีการเผยแพร่มีจำนวนน้อย เนื่องจากชุดข้อมูลด้านการทำนายอายุการใช้งานคงเหลือส่วนใหญ่เป็นชุดข้อมูลที่เหมาะกับแบบจำลองการวิเคราะห์ทางด้านฟิสิกส์ (Physics-Based models) ซึ่งต้องอาศัยความรู้เฉพาะด้านในการนำข้อมูลมาวิเคราะห์ นอกจากนี้ชุดข้อมูลที่เผยแพร่ส่วนใหญ่ไม่มีผลเฉลย เนื่องจากนิยมนำมาทำการจำลองอาการหรือสถานะต่างๆของเครื่องจักรหรืออุปกรณ์เพื่อประกอบการสร้างแบบจำลอง
2. ชุดข้อมูลจากการทำงานของเครื่องจักรในโรงไฟฟ้าหรือโรงงานมีความไม่สมบูรณ์ เนื่องจากต้องมีการใช้ข้อมูลจากหลายแหล่งมาประกอบกันเพื่อให้ได้ชุดข้อมูลและผลเฉลยที่

สมบูรณ์ที่สุดเท่าที่จะเป็นไปได้เหมาะแก่การนำไปทดลอง ทั้งนี้มีสาเหตุมาจากกระบวนการในการบันทึกและจัดเก็บข้อมูลในการดำเนินการต่าง ๆ

3. ข้อจำกัดทางด้านฮาร์ดแวร์ เนื่องจากแบบจำลองที่นำเสนอมีโครงสร้างค่อนข้างซับซ้อน จึงต้องใช้หน่วยความจำ (RAM) มากในการประมวลผลและทดลอง

5.3 แนวทางการวิจัยในอนาคต

เนื่องจากลักษณะข้อมูลจุดวัดของอุปกรณ์หรือระบบในโรงไฟฟ้าหรือโรงงานมีลักษณะเป็นข้อมูลขนาดใหญ่ (Big Data) คือ ข้อมูลมีจำนวนมาก หลากหลาย และเกิดขึ้นใหม่ตลอดเวลา ทำให้ต้องพัฒนาขั้นตอนวิธีในการทำนายสถานะของเครื่องจักรออกมาได้อย่างมีประสิทธิภาพ จึงมีความต้องการพัฒนาการบำรุงรักษาแบบพยากรณ์ล่วงหน้าด้วยการใช้แบบจำลองและระบบที่มีความเหมาะสม ถูกต้องแม่นยำ และมีประสิทธิภาพในการดำเนินการทำนายข้อมูลขนาดใหญ่สำหรับใช้งานในระบบที่ใหญ่ขึ้น ไม่ใช่เพียงแค่ระดับอุปกรณ์เพียงอย่างเดียว

5.4 ผลงานตีพิมพ์จากงานวิจัย

ส่วนหนึ่งของการศึกษาเบื้องต้นของงานวิจัยได้รับการตีพิมพ์ดังรายละเอียดต่อไปนี้

- Sanayha, M., & Vateekul, P. (2017). Fault detection for circulating water pump using time series forecasting and outlier detection. Paper presented at the 2017 9th International Conference on Knowledge and Smart Technology: Crunching Information of Everything, KST 2017, 193-198. 2017.

เนื้อหาของงานวิจัยได้แสดงไว้ในภาคผนวก ก

- Sanayha, M., & Vateekul, P. (2017). Remaining Useful Life Prediction Using Enhanced Convolutional Neural Network on Multivariate Time Series Sensor Data. Paper presented at the 2017 14th International Joint Conference on Computer Science and Software Engineering.

อยู่ระหว่างการรอตีพิมพ์ เนื้อหาของงานวิจัยได้แสดงไว้ในภาคผนวก ข

ดัชนีศัพท์

[A]

Auto Regression Integrated Moving Average (ARIMA)	การพยากรณ์อนุกรมเวลาด้วยวิธีอาร์มา, 21
AlexNet	เน็ตเวิร์กคอนโวลูชันแบบ AlexNet, 6
Activation Function	ฟังก์ชันกระตุ้น, 10
Adaptive Gradient Descent (AdaGrad)	เกรเดียนแบบปรับตัวได้, 14
Average Pooling	การทำพูลลิงแบบหาค่าเฉลี่ย, 19

[B]

Batch Size	ขนาดของแต่ละรุ่น, 49
Backpropagation	การแพร่กระจายย้อนกลับ, 11

[C]

Corrective Maintenance	การบำรุงรักษาเชิงแก้ไข, 4
Cost Function	ฟังก์ชันต้นทุน, 13
Co-adaptation	การปรับตัวร่วมกัน, 14
Convolutional Neural Network	เน็ตเวิร์กคอนโวลูชัน, 16
Combined Cycle Power Plant	โรงไฟฟ้าพลังความร้อนร่วม, 45
Circulating Water Pump	เครื่องปั้มน้ำหมุนเวียน, 43

[E]

Exponential Linear Unit Function	ฟังก์ชันเอกซ์โพเนนเชียลเชิงเส้น, 13
Exponential smoothing	การพยากรณ์อนุกรมเวลาด้วยวิธีการปรับเรียบแบบเอกซ์โพเนนเชียล, 23
Epoch	รอบการเรียนรู้, 48

[F]	
Feature Extraction	การหาคุณลักษณะพิเศษ, 21
Fault Detection	การตรวจจับความผิดปกติ, 25
Feature Map	จำนวนผลลัพธ์จากชั้นคอนโวลูชันที่เท่ากับ จำนวนของเคอร์เนลฟิลเตอร์, 16
[H]	
Historical Data	ข้อมูลประวัติ, 4
Hidden Layer	ชั้นซ่อน, 9
Hyperbolic Tangent Function	ฟังก์ชันแทนเจนต์ไฮเพอร์โบลิก, 13
[I]	
Image Classification	การจัดกลุ่มข้อมูลรูปภาพ, 5
[L]	
LeNet-5	เน็ตเวิร์กคอนโวลูชันแบบ LeNet-5, 6
Loss Function	ฟังก์ชันต้นทุน, 13
Local Optima	โลคอลออปติมา, 13
Long-Short Term Memory	หน่วยความจำระยะสั้นแบบยาว, 15
L^2 Regularization	การทำเรกูลาร์ไรเซชันแบบแอลสอง, 20
Linear Regression	การถดถอยแบบเชิงเส้น, 4
[M]	
Measuring Point	จุดวัด, 4
Mean Squared Error (MSE)	ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย, 14
Momentum	โมเมนตัม, 13
Multi-Layer Perceptron (MLP)	เพอร์เซ็ปตรอนหลายชั้น, 25
Max Pooling	การทำพูลลิงแบบหาค่ามากที่สุด, 19
Multivariate Forecasting	การทำนายแบบหลายตัวแปร, 37

[N]	
Neural Network	โครงข่ายประสาทเทียม, 5
Noise	ข้อมูลรบกวน, 20
[O]	
Objective Function	ฟังก์ชันต้นทุน, 13
Optimization	การหาค่าที่เหมาะสมที่สุด, 13
Overfitting	การอิงกับข้อมูลเรียนรู้มากเกินไป, 14
[P]	
Preventive Maintenance	การบำรุงรักษาเชิงป้องกัน, 4
Predictive Maintenance	การบำรุงรักษาแบบพยากรณ์ล่วงหน้า, 4
Pooling Layer	ชั้นพูลลิง, 19
Performance Evaluation	การวัดประสิทธิภาพ
Prognostics	การทำนายล่วงหน้า, 5
Preprocessing	การประมวลผลก่อน, 24
[R]	
Remaining Useful Life	อายุการใช้งานคงเหลือ, 5
Rectified Linear Unit Function	ฟังก์ชันเรคทีไฟต์เชิงเส้น, 13
Regularization	การทำเรกูลาร์ไรเซชัน, 20
Root Mean Square Error (RMSE)	ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย, 23
[S]	
Signal	สัญญาณ, 4
Sigmoid Function	ฟังก์ชันซิกมอยด์, 13
Softmax Function	ฟังก์ชันค่าสูงสุดอย่างอ่อน, 13
Stochastic Gradient Descent (SGD)	สโตแคสติกเกรเดียนเดสเซนท์, 13
Standard Deviation (SD)	ส่วนเบี่ยงเบนมาตรฐาน, 36

[T]

Turbofan Engine Degradation Simulation	ชุดข้อมูลจำลองของเครื่องเทอร์โบแฟน, 40
Training	การสอน, 48
Testing	การทดสอบ, 48

[U]

Univariate Forecasting	การทำนายแบบตัวแปรเดียว, 37
------------------------	----------------------------

[V]

Vibration Signal	สัญญาณการสั่นสะเทือน, 4
VGG-16	เน็ตเวิร์กคอนโวลูชันแบบ VGG-16, 6
Validation	การตรวจสอบความถูกต้องของแบบจำลอง, 22

[W]

Weight	ค่าน้ำหนัก, 9
Weighted Moving Average	ค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก, 23

[Z]

Zero Padding	การเสริมด้วยศูนย์, 17
--------------	-----------------------

รายการอ้างอิง

- [1] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine Learning for Predictive Maintenance: A Multiple Classifier Approach," *IEEE Transactions on Industrial Informatics*, vol. 11, pp. 812-820, 2015.
- [2] P. K. Wong, Z. Yang, C. M. Vong, and J. Zhong, "Real-time fault diagnosis for gas turbine generator systems using extreme learning machine," *Neurocomputing*, vol. 128, pp. 249-257, 2014/03/27/ 2014.
- [3] V. P. Darabad, M. Vakilian, T. R. Blackburn, and B. T. Phung, "An efficient PD data mining method for power transformer defect models using SOM technique," *International Journal of Electrical Power & Energy Systems*, vol. 71, pp. 373-382, 2015.
- [4] R. Prytz, S. Nowaczyk, T. Rönngvaldsson, and S. Byttner, "Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data," *Engineering Applications of Artificial Intelligence*, vol. 41, pp. 139-150, 5// 2015.
- [5] R. Langone, C. Alzate, B. De Ketelaere, J. Vlasselaer, W. Meert, and J. A. K. Suykens, "LS-SVM based spectral clustering and regression for predicting maintenance of industrial machines," *Engineering Applications of Artificial Intelligence*, vol. 37, pp. 268-278, 2015/01/01/ 2015.
- [6] M. M. Ismail and M. A. M. Hassan, "Artificial neural network based approach compared with stochastic modelling for electrical load forecasting," in *2013 5th International Conference on Modelling, Identification and Control (ICMIC)*, 2013, pp. 112-118.
- [7] O. José Pereira, L. de Almeida Pacheco, S. Sá Barreto, W. Emanuel, C. Fontes, and C. Arthur M. Teixeira Cavalcante, *Pattern Recognition using Multivariate Time Series for Fault Detection in a Thermoelectric Unit* vol. 31, 2012.

- [8] S. Porotsky, "Remaining useful life estimation for systems with non-trendability behaviour," in *2012 IEEE Conference on Prognostics and Health Management*, 2012, pp. 1-6.
- [9] P. Baraldi, M. Compare, S. Saucio, and E. Zio, "Ensemble neural network-based particle filtering for prognostics," *Mechanical Systems and Signal Processing*, vol. 41, pp. 288-300, 12// 2013.
- [10] G. Sateesh Babu, P. Zhao, and X.-L. Li, "Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life," in *Database Systems for Advanced Applications: 21st International Conference, DASFAA 2016, Dallas, TX, USA, April 16-19, 2016, Proceedings, Part I*, S. B. Navathe, W. Wu, S. Shekhar, X. Du, X. S. Wang, and H. Xiong, Eds., ed Cham: Springer International Publishing, 2016, pp. 214-228.
- [11] J. Xu, Y. Wang, and L. Xu, "PHM-Oriented Integrated Fusion Prognostics for Aircraft Engines Based on Sensor Data," *IEEE Sensors Journal*, vol. 14, pp. 1124-1132, 2014.
- [12] O. F. Eker, F. Camci, and I. K. Jennions, "Major Challenges in Prognostics: Study on Benchmarking Prognostics Datasets," presented at the First European Conference of the Prognostics and Health Management Society 2012, 2012.
- [13] D. A. Virkler, B. M. Hillberry, and P. K. Goel, "The Statistical Nature of Fatigue Crack Propagation," *Journal of Engineering Materials and Technology*, vol. 101, pp. 148-153, 1979.



ภาคผนวก ก รายละเอียดของข้อมูลเทอร์โบแฟนชุด C-MAPSS

Data Set: FD001

Train trajectories: 100

Test trajectories: 100

Conditions: ONE (Sea Level)

Fault Modes: ONE (HPC Degradation)

Data Set: FD002

Train trajectories: 260

Test trajectories: 259

Conditions: SIX

Fault Modes: ONE (HPC Degradation)

Data Set: FD003

Train trajectories: 100

Test trajectories: 100

Conditions: ONE (Sea Level)

Fault Modes: TWO (HPC Degradation, Fan Degradation)

Data Set: FD004

Train trajectories: 248

Test trajectories: 249

Conditions: SIX

Fault Modes: TWO (HPC Degradation, Fan Degradation)



Experimental Scenario

Data sets consists of multiple multivariate time series. Each data set is further divided into training and test subsets. Each time series is from a different engine – i.e., the data can be considered to be from a fleet of engines of the same type. Each engine starts with different degrees of initial wear and manufacturing variation which is unknown to the user. This wear and variation is considered normal, i.e., it is not considered a fault condition. There are three operational settings that have a substantial effect on engine performance. These settings are also included in the data. The data is contaminated with sensor noise.

The engine is operating normally at the start of each time series, and develops a fault at some point during the series. In the training set, the fault grows in magnitude until system failure. In the test set, the time series ends some time prior to system failure. The objective of the competition is to predict the number of remaining operational cycles before failure in the test set, i.e., the number of operational cycles after the last cycle that the engine will continue to operate. Also provided a vector of true Remaining Useful Life (RUL) values for the test data.

The data are provided as a zip-compressed text file with 26 columns of numbers, separated by spaces. Each row is a snapshot of data taken during a single operational cycle, each column is a different variable. The columns correspond to:

- 1) unit number
- 2) time, in cycles
- 3) operational setting 1
- 4) operational setting 2
- 5) operational setting 3
- 6) sensor measurement 1
- 7) sensor measurement 2
- ...
- 26) sensor measurement 26

Reference: A. Saxena, K. Goebel, D. Simon, and N. Eklund, “Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation”, in the Proceedings of the 1st International Conference on Prognostics and Health Management (PHM08), Denver CO, Oct 2008.

รายละเอียดข้อมูลเทอร์โบแฟนชุด PHM08

PHM08 Prognostics Data Challenge Dataset

Description

This dataset was used for the prognostics challenge competition at the International Conference on Prognostics and Health Management (PHM08). The challenge is still open for the researchers to develop and compare their efforts against the winners of the challenge in 2008. References to the three winner papers are provided below.

[1] Heimes, F.O., “Recurrent neural networks for remaining useful life estimation”, in the Proceedings of the 1st International Conference on Prognostics and Health Management (PHM08), Denver CO, Oct 2008.

[2] Tianyi Wang, Jianbo Yu, Siegel, D., Lee, J., “A similarity-based prognostics approach for Remaining Useful Life estimation of engineered systems”, in the Proceedings of the 1st International Conference on Prognostics and Health Management (PHM08), Denver CO, Oct 2008.

[3] Peel, L., “Recurrent neural networks for remaining useful life estimation”, in the Proceedings of the 1st International Conference on Prognostics and Health Management (PHM08), Denver CO, Oct 2008.

Experimental Scenario

Data sets consist of multiple multivariate time series. Each data set is further divided into training and test subsets. Each time series is from a different engine – i.e., the data can be considered to be from a fleet of engines of the same type. Each engine starts with different degrees of initial wear and manufacturing variation which is unknown to the user. This wear and variation is considered normal, i.e., it is not considered a fault condition. There are three operational settings that have a substantial effect on engine

performance. These settings are also included in the data. The data are contaminated with sensor noise.

The engine is operating normally at the start of each time series, and starts to degrade at some point during the series. In the training set, the degradation grows in magnitude until a predefined threshold is reached beyond which it is not preferable to operate the engine. In the test set, the time series ends some time prior to complete degradation. The objective of the competition is to predict the number of remaining operational cycles before in the test set, i.e., the number of operational cycles after the last cycle that the engine will continue to operate properly.

Usage

The data are provided as a zip-compressed text file with 26 columns of numbers, separated by spaces. Each row is a snapshot of data taken during a single operational cycle; each column is a different variable. The columns correspond to:

- 1) unit number
- 2) time, in cycles
- 3) operational setting 1
- 4) operational setting 2
- 5) operational setting 3
- 6) sensor measurement 1
- 7) sensor measurement 2
- ...
- 26) sensor measurement 26

Users are expected to train their algorithms using data in the file named train.txt. They must then evaluate the RUL prediction performance on data provided in file test.txt. Associated true RUL values are not being revealed just like in the competition. Very soon the users will be able to use a web application to upload their results and get an aggregate score feedback. Please check back in June 2010 to get more details on the web application for automated scoring. Until then they can get a feedback by emailing their results through simple text files to any of the email addresses provided below. A set of representative top 20 scores obtained during the competition are included here to provide a reference to everyone.

No.	Score
1	436.841
2	512.426
3	737.769
4	809.757
5	908.588
6	975.586
7	1,049.57
8	1,051.88
9	1,075.16
10	1,083.91
11	1,127.95
12	1,139.83
13	1,219.61
14	1,263.02
15	1,557.61
16	1,808.75
17	1,966.38
18	2,065.47
19	2,399.88
20	2,430.42



Evaluation

The final score is a weighted sum of RUL errors. The scoring function is an asymmetric function that penalizes late predictions more than the early predictions. (Please see attached documentation for details)

Once algorithms are trained to satisfaction, users can apply them to the final test dataset contained in the file named `final_test.txt`. Users should send the vector of RULs for the final test set to the PHM Society for evaluation. A score will be mailed back soon. Researchers are encouraged to publish their results regardless of the absolute performance if they believe there is novelty in their algorithm. The intent is to develop innovative approaches for prognostics.

Note: Any team or individual is allowed to submit their results on the final test set ONLY ONCE.

Data Set: train.txt, test.txt

Train trajectories: 218

Test trajectories: 218

final_test trajectories: 435

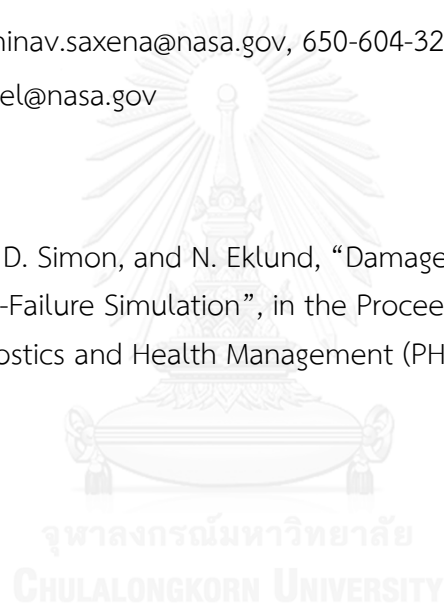
Contacts

Abhinav Saxena – abhinav.saxena@nasa.gov, 650-604-3208

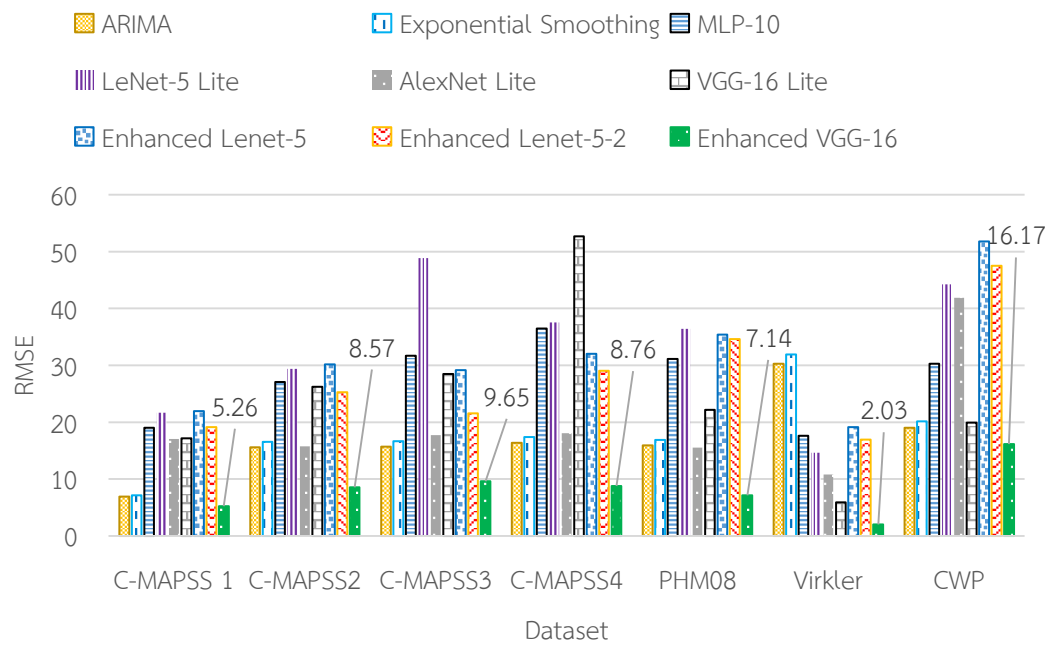
Kai Goebel – kai.goebel@nasa.gov

References

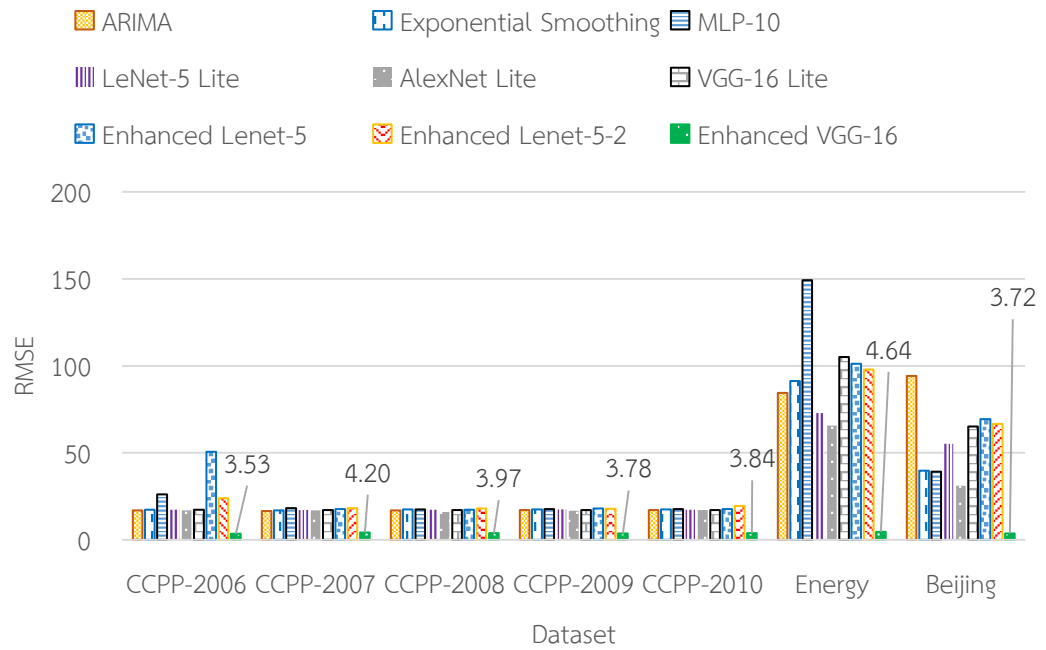
A. Saxena, K. Goebel, D. Simon, and N. Eklund, “Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation”, in the Proceedings of the 1st International Conference on Prognostics and Health Management (PHM08), Denver CO, Oct 2008.



ภาคผนวก ข ผลการทดลองค่า RMSE ในรูปแบบกราฟแท่งของชุดข้อมูลทั้ง 14 ชุด
เมื่อทดลองกับแบบจำลองทั้งเก้าแบบ



รูปที่ 23 ค่า RMSE ในรูปแบบกราฟแท่งจากการทดลองด้วยแบบจำลองเก้าแบบกับชุดข้อมูลด้านการ
ทำนายอายุการใช้งานคงเหลือ



รูปที่ 24 ค่า RMSE ในรูปแบบกราฟแท่งจากการทดลองด้วยแบบจำลองแก้แบบกับชุดข้อมูลอนุกรมเวลาหลายตัวแปร

ภาคผนวก ค ผลการทดสอบการเปรียบเทียบประสิทธิภาพของแบบจำลองทั้งเก้าแบบ
ด้วยวิธี Paired T Test

ตารางที่ 17 แสดงค่า T Value จากการทำการทดสอบ Pair T Test

Methods	ARIMA	Exponential Smoothing	MLP	LeNet-5 Lite	AlexNet Lite	VGG-16 Lite	Enhanced Lenet-5	Enhanced Lenet-5-2
ARIMA		0.69	-1.02	-0.9	0.89	-0.74	-2.15	-1.43
Exponential Smoothing	0.69		-2.19*	-1.83	0.62	-1.58	-3.3	-2.7
MLP	-1.02	-2.19*		0.37	1.95	0.87	-0.48	0.35
LeNet-5 Lite	-0.9	-1.83	0.37		3.29*	0.34	-1.29	-0.21
AlexNet Lite	0.89	0.62	1.95	3.29*		-1.73	-3.9	-3.29
VGG-16 Lite	-0.74	-1.58	0.87	0.34	-1.73		-1.6	-0.64
Enhanced Lenet-5	-2.15	-3.3	-0.48	-1.29	-3.9	-1.6		2.11
Enhanced Lenet-5-2	-1.43	-2.7	0.35	-0.21	-3.29	-0.64	2.11	
Enhanced VGG-16	-2.85*	-3.19*	-3.08*	-5.72*	4.37*	-3.42*	-4.69*	-4.25*

ตารางที่ 18 แสดงค่า P Value จากการทำการทดสอบ Pair T Test

Methods	ARIMA	Exponential Smoothing	MLP	LeNet-5 Lite	AlexNet Lite	VGG-16 Lite	Enhanced Lenet-5	Enhanced Lenet-5-2
ARIMA		0.5	0.32	0.38	0.39	0.47	0.05	0.17
Exponential Smoothing	0.5		0.05*	0.089	0.54	0.14	0.005	0.02
MLP	0.32	0.05*		0.71	0.070	0.4	0.64	0.73
LeNet-5 Lite	0.38	0.089	0.71		0.006*	0.73	0.22	0.83
AlexNet Lite	0.39	0.54	0.070	0.006*		0.11	0.002	0.005*
VGG-16 Lite	0.47	0.14	0.4	0.73	0.11		0.13	0.53
Enhanced Lenet-5	0.05	0.005	0.64	0.22	0.002	0.13		0.05
Enhanced Lenet-5-2	0.17	0.02	0.73	0.83	0.005*	0.53	0.05	
Enhanced VGG-16	0.061*	0.037*	0.048*	0.013*	0.268*	0.041*	0.005*	0.036*

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวมนัสกานต์ เสน่หา เกิดเมื่อวันที่ 5 สิงหาคม พ.ศ. 2532 ที่กรุงเทพมหานคร สำเร็จการศึกษาระดับปริญญาตรีหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ในปีการศึกษา 2554 และเข้าศึกษาในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2558

