

การพยากรณ์กำลังผลิตไฟฟ้ารายภาคจากพลังงานแสงอาทิตย์โดยใช้โครงข่ายประสาทเชิงลึกร่วมกับ
การปรับจูนไฮเปอร์พารามิเตอร์



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2562
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

REGIONAL SOLAR POWER FORECASTING USING DEEP NEURAL NETWORK AND
HYPERPARAMETER TUNING



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

Department of Electrical Engineering

FACULTY OF ENGINEERING

Chulalongkorn University

Academic Year 2019

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การพยากรณ์กำลังผลิตไฟฟ้ารายภาคจากพลังงาน แสงอาทิตย์โดยใช้โครงข่ายประสาทเชิงลึกร่วมกับการปรับ จูนไฮเปอร์พารามิเตอร์
โดย	นายสุกฤษฎ์ ใจดี
สาขาวิชา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.วันเฉลิม โปรา

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

..... คณะบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุรียะ พุ่มรินทร์)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.วันเฉลิม โปรา)

..... กรรมการภายนอกมหาวิทยาลัย
(รองศาสตราจารย์ ดร.เกียรติยศ กวีญาณ)

สุกฤษฎ์ ใจดี : การพยากรณ์กำลังผลิตไฟฟ้ารายภาคจากพลังงานแสงอาทิตย์โดยใช้
 โครงข่ายประสาทเชิงลึกร่วมกับการปรับจูนไฮเปอร์พารามิเตอร์. (REGIONAL SOLAR
 POWER FORECASTING USING DEEP NEURAL NETWORK AND
 HYPERPARAMETER TUNING) อ.ที่ปรึกษาหลัก : รศ. ดร.วันเฉลิม โปรา

เทคโนโลยีพลังงานปัจจุบันนำไปสู่การขยายตัวของโครงข่ายสมาร์ตกริดซึ่งเป็นผลให้เกิด
 การเปลี่ยนแปลงในหลายแง่มุม เช่น การเพิ่มขึ้นของสัดส่วนพลังงานแสงอาทิตย์ ความก้าวหน้า
 ระบบสำรองพลังงาน และความสมดุลระหว่างการผลิตและใช้พลังงาน เป็นต้น การเปลี่ยนแปลง
 เหล่านี้ส่งผลให้เกิดข้อดีหลายแง่มุม เช่น การผลิต การสำรอง และการใช้พลังงานภายในเน็ตเวิร์ก
 เอง ดังนั้นโครงข่ายสมาร์ตกริดจึงไม่จำเป็นต้องพึ่งพาระบบจำหน่ายหลัก แต่อย่างไรก็ตามเน็ตเวิร์ก
 ส่วนใหญ่พึ่งพาพลังงานแสงอาทิตย์ซึ่งมีความผันผวนตามสภาพอากาศ ดังนั้นการมีระบบพยากรณ์
 จะช่วยรักษาเสถียรภาพของเน็ตเวิร์ก วิทยานิพนธ์ฉบับนี้จึงเสนอวิธีการพยากรณ์กำลังผลิตไฟฟ้า
 จากพลังงานแสงอาทิตย์โดยใช้โครงข่ายประสาทเทียมแบบวนกลับร่วมกับการปรับปรุงความ
 แม่นยำในการพยากรณ์ด้วยเทคนิคต่างๆ ได้แก่ การปรับจูนไฮเปอร์พารามิเตอร์ การคัดเลือกตัว
 แปร การหาค่าเฉลี่ยเชิงพื้นที่ การหาค่าเฉลี่ยเคลื่อนที่แบบเอกซ์โพเนนเชียลของความเข้ม
 แสงอาทิตย์ และการเอนเซมเบิลค่าพยากรณ์พลังงานไฟฟ้าจากแต่ละแบบจำลอง ข้อมูลอินพุต
 สำหรับวิทยานิพนธ์ฉบับนี้ประกอบไปด้วย ข้อมูลจากแบบจำลองพยากรณ์อากาศและข้อมูลจาก
 อุปกรณ์ตรวจวัด วิธีที่นำเสนอศึกษาแยกเป็น 2 กรณี ได้แก่ การพยากรณ์กำลังผลิตรวมของ 7
 โรงไฟฟ้า การพยากรณ์พลังงานแสงอาทิตย์บนตึกภาคควิชา ผลลัพธ์ที่ได้พบว่าแบบจำลองเอนเซม
 เบิลให้ค่าความผิดพลาดอยู่ที่ 6.94% RMSE ซึ่งลดลง 1.71% เมื่อเทียบกับแบบจำลองอินพุตที่ดี
 ที่สุด เมื่อเราเปรียบเทียบชนิดของโหนดจาก GRU เป็น CuDNNGRU ค่า RMSE เพิ่มขึ้นน้อยกว่า 1%
 แต่เวลาในการฝึกสอนลดลงมากกว่าเท่าตัว

สาขาวิชา วิศวกรรมไฟฟ้า

ปีการศึกษา 2562

ลายมือชื่อนิสิต

ลายมือชื่อ อ.ที่ปรึกษาหลัก

6170306321 : MAJOR ELECTRICAL ENGINEERING

KEYWORD: Solar Power Forecast, Genetic Algorithm, Recurrent Neural Network,
hyperparameter tuning, Numerical Weather Prediction

Sukrit Jaidee : REGIONAL SOLAR POWER FORECASTING USING DEEP
NEURAL NETWORK AND HYPERPARAMETER TUNING . Advisor: Assoc. Prof.
WANCHALERM PORA, Ph.D.

The up-to-date energy technologies lead to the expansion of the Smart Grid Networks (SGN), which brings about a variety of changes, such as, the increment of the renewable energy ratio, progression in the storage system, and the real-time balance of energy demand and supply. In Thailand, the solar power is a dominant part of renewable energy, but it is subject to fluctuation due to weather conditions. Therefore, having a solar power forecast (SPF) helps to maintain the stability of the network. This thesis proposes SPF methods using Recurrent Neural Network (RNN) with a few accuracy improvements using several techniques, namely hyperparameter tuning by GA, feature selection, spatial averaging, exponential moving average, and the ensemble of forecasted solar power. The input data for this thesis are composed of forecast weather data and measured data. The proposed method applies to two study cases, namely forecasting the sum of solar power of 7 solar plants and forecasting solar power at EECU building. Test results show that the ensemble of the forecast power provides 6.94 % RMSE, which decreased by 1.71% compared to the best input model. When we change the recurrent node type from GRU to CuDNNGRU, RMSE increases less than 1% but the training time is reduced more than half.

Field of Study: Electrical Engineering

Student's Signature

Academic Year: 2019

Advisor's Signature

กิตติกรรมประกาศ

การที่วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีนั้นมาจากความช่วยเหลืออย่างมากของบุคคลหลายๆท่านที่ให้ความช่วยเหลือในการทำวิทยานิพนธ์ฉบับนี้ขึ้นมา ข้าพเจ้ารู้สึกซาบซึ้งใจเป็นอย่างมาก

ขอขอบพระคุณ รองศาสตราจารย์ ดร. วันเฉลิม โปธา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้ที่คอยให้ความช่วยเหลือและสนับสนุนทั้งในด้าน วิชาการ งานวิจัย ความรู้ คำแนะนำ และเทคนิคต่างๆ รวมทั้งโอกาสต่างๆที่มอบให้แก่ข้าพเจ้า

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. สุรีย์ พุ่มรินทร์ และ รองศาสตราจารย์ ดร. เกียรติยุทธ กวีญาณ ที่ได้ให้เกียรติมาเป็นกรรมการวิทยานิพนธ์รวมทั้งการให้คำแนะนำ ข้อเสนอแนะ แนวทางงานวิจัย พร้อมทั้งตรวจสอบ แก้ไขให้วิทยานิพนธ์ฉบับนี้ผ่านลุล่วงไปได้ด้วยดี

ขอขอบพระคุณอาจารย์ทุกท่าน คณะกรรมการหลักสูตรสมาร์ทกริด คลัสเตอร์ระบบอัจฉริยะ คณะบุคลากร และเจ้าหน้าที่ทุกท่าน ทั้งจากภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย สำหรับความรู้ในวิชาต่างๆ คำแนะนำและโอกาสได้พัฒนาตนเอง ขอขอบคุณเพื่อนๆในห้องปฏิบัติการคลัสเตอร์ระบบอัจฉริยะทุกคนสำหรับทัศนคติ มุมมอง ความรู้ใหม่ๆ รวมทั้งประสบการณ์อันมีค่ามากมาย

ขอขอบพระคุณคณะกรรมการผู้คัดเลือกผู้ได้รับทุนสมาร์ทกริด กฟผ. และคณะผู้บริหารสายงาน รวฟ. ชฟผ. อชว. ที่อนุญาตให้ลามาศึกษา ความรู้ที่ได้จากหลักสูตรและงานวิจัยนี้จะนำไปใช้ให้เกิดประโยชน์มากที่สุด

สุดท้ายนี้ข้าพเจ้าขอขอบพระคุณ คุณพ่อ คุณแม่ และครอบครัว ที่ให้การสนับสนุน ทั้งในด้านการศึกษา การใช้ชีวิต และเป็นกำลังใจให้ตลอดมา

สุกฤษฎ์ ใจดี

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ญ
สารบัญรูป.....	ฎ
บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญ.....	1
1.2 วัตถุประสงค์.....	2
1.3 ขอบเขตวิทยานิพนธ์.....	3
1.4 ขั้นตอนและวิธีการดำเนินงาน.....	4
1.5 ประโยชน์ที่คาดว่าจะได้รับจากวิทยานิพนธ์.....	4
1.6 เนื้อหาของวิทยานิพนธ์.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 งานวิจัยที่เกี่ยวข้อง.....	5
2.1.1 ปัญหาและวิธีการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์.....	5
2.2.2 การปรับจูนไฮเปอร์พารามิเตอร์ของเน็ตเวิร์ก.....	7
2.2 ความรู้พื้นฐานที่เกี่ยวข้อง.....	9
2.2.1 การรู้จำรูปแบบ (Pattern Recognition).....	10
2.2.1.1 ปัญหาของ Fully Connected Layer (FCL).....	10
2.2.1.2 สมการ Recurrent Neural Network (RNN).....	13

2.2.1.3 การคำนวณ Forward Propagation ของเน็ตเวิร์กรีเคอร์ซีฟ.....	14
2.2.1.4 การคำนวณ Backpropagation ของเน็ตเวิร์กรีเคอร์ซีฟ	15
2.2.1.5 สมการ Gated Recurrent Unit (GRU)	15
2.2.1.6 สมการ Long Short-Term Memory (LSTM).....	18
2.2.1.7 การต่อ GRU เลเยอร์ (Gated Recurrent Unit (GRU) layer)	21
2.2.1.8 ข้อดีข้อเสียของเน็ตเวิร์กแบบวนกลับ.....	23
2.2.2 ขั้นตอนวิธีเชิงวิวัฒนาการ (Evolutionary Algorithm)	23
2.2.2.1 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm; GA)	25
2.2.2.2 การเข้ารหัสโครโมโซม	26
2.2.2.3 การแทนแบบเลขฐานสอง.....	27
2.2.2.4 การแทนแบบเลขจำนวนเต็ม	27
2.2.2.5 การแทนแบบเลขจำนวนจริง	27
2.2.2.6 การแทนแบบเรียงสับเปลี่ยน	28
2.2.2.7 การคัดเลือกประชากรพ่อแม่ (Parent Selection).....	28
2.2.2.8 การคัดเลือกแบบวงล้อรูเล็ต (Roulette Wheel Selection).....	28
2.2.2.9 การคัดเลือกแบบการแข่งขัน (Tournament Selection).....	29
2.2.2.10 การคัดเลือกผู้อยู่รอด (Survivor Selection).....	30
2.2.3 การเรียนรู้แบบรวมกลุ่ม (Ensemble Learning Method).....	30
2.2.4 แบบจำลองพยากรณ์อากาศ	31
2.2.4.1 Pysolar	32
2.2.4.2 National Renewable Energy Laboratory (NREL).....	32
2.2.4.3 National Centers for Environmental Prediction (NCEP)	32
2.2.4.4 Measured Data (ค่าวัดจริง).....	32

2.2.5 การจัดกลุ่มข้อมูล.....	32
2.2.6 ตัวชี้วัดพื้นฐาน (Metrics) สำหรับการวัดผลแบบจำลอง	33
2.2.7 ภาษาและแพลตฟอร์มสำหรับการพัฒนาแบบจำลอง.....	34
บทที่ 3 วิธีการดำเนินการวิจัย	36
3.1 การพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ด้วยแบบจำลองการพยากรณ์อากาศเชิงตัวเลข (Very Short-Term SPF using data from NWP model)	38
3.1.1 ภาพรวมวิธีการดำเนินการวิจัย.....	38
3.1.2 ข้อมูลอินพุต	39
3.1.3 พารามิเตอร์ของเน็ตเวิร์ก.....	42
3.1.4 โครงสร้างเน็ตเวิร์ก FCL และ RNN.....	44
3.1.5 เทคนิคการฝึกสอน.....	46
3.2 การพยากรณ์กำลังผลิตไฟฟ้าระยะสั้นโดยใช้นิวรอนเน็ตเวิร์กและขั้นตอนวิธีเชิงพันธุกรรม (Very Short-Term SPF using GA Based Neural Network)	47
3.2.1 ภาพรวมของวิธีการดำเนินการวิจัย	47
3.2.2 ข้อมูลอินพุต	48
3.2.3 พารามิเตอร์ต่างๆที่ใช้ในการวิจัย	49
3.2.4 โครงสร้างเน็ตเวิร์ก.....	50
3.2.5 พารามิเตอร์ที่ใช้ในการฝึกสอน.....	52
3.3 การพยากรณ์กำลังผลิตไฟฟ้ารายภาคโดยใช้นิวรอนเน็ตเวิร์กพร้อมกับการปรับจูนไฮเปอร์พารามิเตอร์กรณีโรงไฟฟ้า 7 แห่ง (Neural Network based on GA and Ensemble for RSPF).....	53
3.3.1 ข้อมูลอินพุต	53
3.3.1.1 ข้อมูลจากแบบจำลองพยากรณ์อากาศ	53
3.3.1.2 ข้อมูลค่าวัดจากอุปกรณ์ตรวจวัด	55

3.3.1.3 ข้อมูลกำลังผลิตไฟฟ้า.....	55
3.3.1.4 การจัดกลุ่มข้อมูล.....	56
3.3.2 พารามิเตอร์ของเน็ตเวิร์ก.....	57
3.3.3 ขั้นตอนการรวมกำลังผลิตรวม (ผลรวม 7 โรง) ของแต่ละแบบจำลอง.....	58
3.3.4 ขั้นตอนการวัดประสิทธิภาพ.....	59
บทที่ 4 ผลการวิจัย.....	62
4.1 แบบจำลอง Single Solar Power Forecast (SSPF).....	62
4.1.1 การพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ด้วยแบบจำลองการพยากรณ์ อากาศเชิงตัวเลข (Very Short-Term SPF using data from NWP model).....	62
4.1.1.1 การทดสอบบนเน็ตเวิร์ก FCL.....	62
4.1.1.2 การทดสอบบนเน็ตเวิร์กรีเคอร์ซีฟชนิด RNN.....	65
4.1.2 การพยากรณ์กำลังผลิตไฟฟ้าระยะสั้นโดยใช้นิวรอนเน็ตเวิร์กและขั้นตอนวิธีเชิง พันธุกรรม (Very Short-Term SPF using GA Based Neural Network).....	70
4.2 แบบจำลอง Regional Solar Power Forecast (RSPF).....	74
4.2.1 ผลลัพธ์การค้นหาไฮเปอร์พารามิเตอร์ที่เหมาะสม.....	74
4.2.2 การเรียนรู้แบบรวมกลุ่ม.....	75
บทที่ 5 สรุปผลการวิจัย.....	76
บรรณานุกรม.....	79
ประวัติผู้เขียน.....	84

สารบัญตาราง

	หน้า
ตารางที่ 1 ข้อดีและข้อเสียของเน็ตเวิร์กแบบวงกลับ.....	23
ตารางที่ 2 วิวัฒนาการของขั้นตอนวิธีเชิงวิวัฒนาการ	24
ตารางที่ 3 ข้อดีและข้อเสียของวิธีการคัดเลือกประชากรพ่อแม่	30
ตารางที่ 4 ข้อดีของตัวชี้วัดแต่ละประเภท	34
ตารางที่ 5 ข้อดีและข้อเสียของแพลตฟอร์ม.....	35
ตารางที่ 6 การประยุกต์ใช้วิธีการต่างๆในการปรับปรุงประสิทธิภาพ	37
ตารางที่ 7 พารามิเตอร์ต่างๆของเน็ตเวิร์กสำหรับกรณีศึกษา.....	44
ตารางที่ 8 โครงสร้างของเน็ตเวิร์กประเภทต่างๆที่ได้จากขั้นตอนวิธีเชิงพันธุกรรม.....	51
ตารางที่ 9 ตัวแปรของแบบจำลองพยากรณ์อากาศทั้ง 3	54
ตารางที่ 10 รายการแหล่งข้อมูลแบบจำลองพยากรณ์อากาศที่ให้บริการฟรี.....	54
ตารางที่ 11 ตัวแปรต่างๆที่ได้จากอุปกรณ์ตรวจวัดที่ถุกวัดในแต่ละโรงไฟฟ้า	55
ตารางที่ 12 กำลังผลิตติดตั้งและตำแหน่งที่ตั้งของโรงไฟฟ้าทั้ง 7 โรง	56
ตารางที่ 13 หมายกลุ่มข้อมูลของโรงไฟฟ้าแต่ละโรง.....	57
ตารางที่ 14 โครงสร้างของเน็ตเวิร์กที่ได้จาก GA.....	58
ตารางที่ 15 ชุดข้อมูล 4 ชุดที่ใช้ในการทดสอบประสิทธิภาพ.....	62
ตารางที่ 16 การเปรียบเทียบ RMSE ระหว่างชุดข้อมูลที่ 1 และ 3 ในแต่ละ time-steps.....	64
ตารางที่ 17 ความผิดพลาดในแต่ละ time-steps ในแต่ละเน็ตเวิร์ก	71
ตารางที่ 18 ค่าความผิดพลาดในแต่ละช่วงเวลาของวันของแต่ละเน็ตเวิร์ก	71
ตารางที่ 19 ค่า MBE ในแต่ละช่วงเวลาของวันของแต่ละเน็ตเวิร์ก	71
ตารางที่ 20 เวลาที่ใช้สำหรับการฝึกสอนของเน็ตเวิร์กแต่ละประเภท	71

ตารางที่ 21 เปอร์เซ็นต์ RMSE ในแต่ละ Time-step ของแบบจำลองแต่ละประเภท75



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญรูป

	หน้า
รูปที่ 1 ที่มาและความสำคัญของปัญหา.....	2
รูปที่ 2 ระบบพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์	3
รูปที่ 3 วิธีการพยากรณ์โดยการใช้ ANN ร่วมกับ SOM	5
รูปที่ 4 วิธีการปรับปรุงประสิทธิภาพแบบจำลอง	10
รูปที่ 5 ข้อเสียของเน็ตเวิร์ก FCL	11
รูปที่ 6 การต่อกันของแต่ละเซลล์ RNN	11
รูปที่ 7 ปัญหาการจำสถานะเก่าที่ไม่เกี่ยวข้องกับสถานะปัจจุบันของเน็ตเวิร์ก RNN	12
รูปที่ 8 โครงสร้างทางสถาปัตยกรรมของเซลล์ GRU เทียบกับเซลล์ RNN.....	13
รูปที่ 9 ขั้นตอนวิธีการแพร่ไปข้างหน้าของเน็ตเวิร์ก RNN	14
รูปที่ 10 ขั้นตอนวิธีการแพร่ย้อนกลับของเน็ตเวิร์ก RNN.....	15
รูปที่ 11 โครงสร้างทางสถาปัตยกรรมของเซลล์ GRU เทียบกับเซลล์ RNN.....	16
รูปที่ 12 ฟังก์ชันซิกมอยด์ (sigmoid)	16
รูปที่ 13 ฟังก์ชัน tanh	17
รูปที่ 14 ฟังก์ชันเส้นตรงที่ถูกปรับแก้ Rectified (Rectified Linear Unit ; ReLU).....	18
รูปที่ 15 โครงสร้างของเซลล์ LSTM เทียบกับ GRU	19
รูปที่ 16 การต่อ GRU layer (แบบละเอียด Unrolling).....	22
รูปที่ 17 การต่อ GRU layer (แบบย่อ).....	23
รูปที่ 18 แผนภาพการทำงานของขั้นตอนวิธีเชิงวิวัฒนาการ	25
รูปที่ 19 การคัดเลือกแบบวงล้อรูเล็ต	29
รูปที่ 20 การคัดเลือกแบบการแข่งขัน	29

รูปที่ 21	กระบวนการคัดเลือกประชากรผู้อยู่รอด.....	30
รูปที่ 22	Bias-Variance trade-off.....	31
รูปที่ 23	ภาพรวมอินพุตประเภทต่างๆของแบบจำลอง.....	32
รูปที่ 24	การจัดกลุ่มโดยให้ความแปรผันภายในกลุ่มมีค่าน้อยและระหว่างกลุ่มมีค่ามาก.....	33
รูปที่ 25	ภาษาและแพลตฟอร์มสำหรับการพัฒนาแบบจำลอง.....	34
รูปที่ 26	ข้อดีและข้อเสียของแต่ละวิธีการพยากรณ์.....	36
รูปที่ 27	ข้อดีและข้อเสียของแต่ละวิธีการปรับปรุงประสิทธิภาพแบบจำลอง.....	36
รูปที่ 28	ภาพรวมของวิธีการพยากรณ์ทั้งสามวิธีการ.....	37
รูปที่ 29	แนวโน้มของการปรับปรุงประสิทธิภาพแบบจำลอง.....	38
รูปที่ 30	ภาพรวมของวิธีการพยากรณ์และการปรับปรุงประสิทธิภาพวิธีที่หนึ่ง.....	39
รูปที่ 31	อาร์เรย์ของแผงโซลาร์เซลล์ที่ติดตั้งบนอาคารบูรอดบดินทสันต์.....	40
รูปที่ 32	ตัวอย่างโดเมนของแบบจำลอง WRF ในพื้นที่ประเทศไทย.....	41
รูปที่ 33	โดเมนของแบบจำลอง WRF ซึ่งครอบคลุมกรณีศึกษา.....	41
รูปที่ 34	ข้อมูลอินพุต 3 ประเภท.....	42
รูปที่ 35	ลักษณะของชุดข้อมูลฝึกสอนในมิติต่างๆ.....	43
รูปที่ 36	การพยากรณ์ไปล่วงหน้า 8 HORIZON (8 time-steps).....	43
รูปที่ 37	จำนวนพารามิเตอร์ของเน็ตเวิร์กในแต่ละเลเยอร์สำหรับการกรณีศึกษา.....	44
รูปที่ 38	ภาพรวมของเน็ตเวิร์ก FCL และข้อมูลอินพุต.....	45
รูปที่ 39	ข้อมูลการมองย้อนกลับเทียม (artificial lookback data).....	45
รูปที่ 40	การต่อเน็ตเวิร์กในลักษณะต่างๆของ RNN.....	45
รูปที่ 41	การต่อเน็ตเวิร์ก RNN สำหรับกรณีศึกษา.....	46
รูปที่ 42	ภาพรวมการพยากรณ์และวิธีการปรับปรุงประสิทธิภาพด้วยวิธีที่สอง.....	47
รูปที่ 43	ขั้นตอนการค้นหาไฮเปอร์พารามิเตอร์ของเน็ตเวิร์ก.....	48

รูปที่ 44 อินพุตของเน็ตเวิร์กในลักษณะสามมิติ	49
รูปที่ 45 โครงสร้างโครงข่ายประสาทเทียมสำหรับกรณีสึกษา.....	50
รูปที่ 46 ภาพรวมขั้นตอนการพยากรณ์รวม 7 โรงไฟฟ้า.....	53
รูปที่ 47 ภาพรวมการพยากรณ์และการปรับปรุงประสิทธิภาพแบบจำลองด้วยวิธีที่สาม.....	54
รูปที่ 48 แบบจำลองพยากรณ์อากาศประเภทต่างๆ	55
รูปที่ 49 กำลังผลิตไฟฟ้าของแต่ละโรงไฟฟ้าในแต่ละพื้นที่.....	56
รูปที่ 50 การจัดกลุ่มข้อมูลด้วยเคมีนสำหรับกรณีสึกษา.....	57
รูปที่ 51 การค้นหาไฮเปอร์พารามิเตอร์ของเน็ตเวิร์กด้วยขั้นตอนวิธีเชิงพันธุกรรม	58
รูปที่ 52 ขั้นตอนการผสมผสานค่าพยากรณ์จากแต่ละแบบจำลอง.....	59
รูปที่ 53 ค่า MAE ของแบบจำลอง NCEP ในแต่ละโรงไฟฟ้า.....	60
รูปที่ 54 ค่า MAE ของแบบจำลอง NREL ในแต่ละโรงไฟฟ้า	60
รูปที่ 55 ค่า MAE ของแบบจำลอง Pysolar ในแต่ละโรงไฟฟ้า	61
รูปที่ 56 ค่า MAE ของแบบจำลอง measured data ในแต่ละโรงไฟฟ้า	61
รูปที่ 57 เปรียบเทียบ RMSE ของ 4 ชุดข้อมูลในแต่ละขนาดพื้นที่	62
รูปที่ 58 เปรียบเทียบ RMSE ตอน 6 โมงเช้าระหว่างชุดข้อมูล 1 และ 3 (พื้นที่ขนาด 1x1).....	63
รูปที่ 59 เปรียบเทียบ RMSE ระหว่างชุดข้อมูลที่ 1 และ 3 ในแต่ละ time-steps อนาคต.....	63
รูปที่ 60 เปรียบเทียบ MBE ระหว่างข้อมูลชุดที่ 1 และ 3 สำหรับขนาดพื้นที่ 1x1	64
รูปที่ 61 เปรียบเทียบ RMSE(%) ของแต่ละชุดข้อมูลและแต่ละขนาดพื้นที่ระหว่าง RNN และ FCL.....	65
รูปที่ 62 RMSE (%) ของข้อมูลชุดที่ 3 สำหรับพื้นที่ขนาด 1x1 ในแต่ละ time-steps	65
รูปที่ 63 การเปรียบเทียบ RMSE ของเน็ตเวิร์กทั้งสอง (1x1) ในแต่ละช่วงเวลาของวัน	66
รูปที่ 64 การเปรียบเทียบ MBE (%) สำหรับ 1x1 ของข้อมูลชุดที่ 3 ตั้งแต่เวลา 6:00 - 18.00	66
รูปที่ 65 ค่า MBE (%) ของแบบจำลอง FCL ในแต่ละ time-steps	67
รูปที่ 66 ค่า MBE (%) ของแบบจำลอง FCL ในแต่ละช่วงเวลาของวัน	67

รูปที่ 67 ค่า MBE (%) ของแบบจำลอง RNN ในแต่ละ time-steps	68
รูปที่ 68 ค่า MBE (%) ของแบบจำลอง RNN ในแต่ละช่วงเวลาของวัน	68
รูปที่ 69 การเปรียบเทียบการปรับปรุงประสิทธิภาพของเน็ตเวิร์กทั้งสอง	69
รูปที่ 70 เปรียบเทียบ RMSE ของแบบจำลอง FCL ที่มีและไม่มี time-delay	69
รูปที่ 71 การเปรียบเทียบกำลังผลิตไฟฟ้าที่เป็นค่าจริงและค่าพยากรณ์ของแต่ละเน็ตเวิร์ก	69
รูปที่ 72 ค่า RMSE ในแต่ละ time-steps ของเน็ตเวิร์กแต่ละประเภท	72
รูปที่ 73 ค่า RMSE ในแต่ละช่วงเวลาของวันของเน็ตเวิร์กแต่ละประเภท	72
รูปที่ 74 ค่า MBE ในแต่ละช่วงเวลาของวันของเน็ตเวิร์กแต่ละประเภท	72
รูปที่ 75 การเปรียบเทียบค่าจริงและค่าพยากรณ์กำลังผลิตไฟฟ้าจากเน็ตเวิร์กแต่ละประเภท	73
รูปที่ 76 การเปรียบเทียบค่า RMSE ของเน็ตเวิร์กแต่ละประเภท	73
รูปที่ 77 RMSE ของแบบจำลองที่ไม่มีการคัดเลือกไฮเปอร์พารามิเตอร์	74
รูปที่ 78 RMSE ของแบบจำลองที่มีการคัดเลือกไฮเปอร์พารามิเตอร์ด้วย GA	74
รูปที่ 79 RMSE ของแบบจำลองแต่ละประเภทในแต่ละ time-Step	75
รูปที่ 80 RMSE ของแบบจำลองแต่ละประเภทเมื่อเปรียบเทียบระหว่างการใช่และไม่ใช้ GA	76
รูปที่ 81 การเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์จากแต่ละแบบจำลอง	76

บทที่ 1

บทนำ

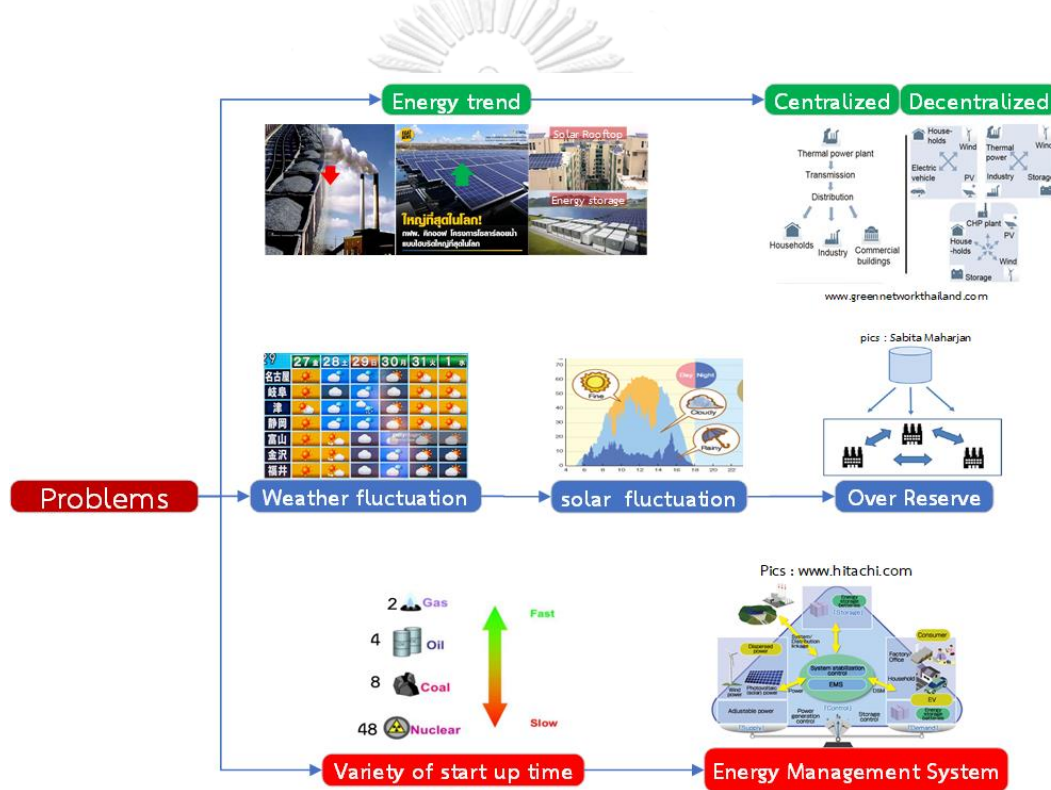
เนื้อหาของบทนี้จะนำเสนอที่มาและความสำคัญของวิทยานิพนธ์ วัตถุประสงค์ ขอบเขตของวิทยานิพนธ์ ขั้นตอนและวิธีการดำเนินงาน ประโยชน์ที่คาดว่าจะได้รับจากวิทยานิพนธ์ งานวิจัยในอดีตที่เกี่ยวข้องกับวิทยานิพนธ์และเนื้อหาของวิทยานิพนธ์

1.1 ที่มาและความสำคัญ

กำลังการผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ที่ป้อนให้กับระบบกริดของประเทศไทยมีแนวโน้มเพิ่มสูงขึ้นทุกปีทำให้เกิดปัญหาจำนวนมากขึ้นดังรูปที่ 1 เช่น ความแตกต่างของระยะเวลาที่ต้องใช้ในการเริ่มเดินเครื่องของพลังงานแต่ละประเภท ความยากในการบริหารจัดการพลังงานแต่ละประเภท ร่วมกันและความผันผวนของพลังงานแสงอาทิตย์ตามสภาพอากาศที่เปลี่ยนแปลง เป็นต้น ดังนั้นหากไม่มีระบบการพยากรณ์กำลังผลิตไฟฟ้าของโรงไฟฟ้าพลังงานแสงอาทิตย์จะต้องพิจารณาปรับเพิ่มกำลังผลิตสำรอง Spinning Reserve (กำลังไฟฟ้าที่มีไว้สำหรับรองรับโรงไฟฟ้าในระบบหลุดออกจากระบบฉับพลัน) เพื่อให้รองรับความผันผวนสูงสุดที่เพิ่มขึ้นในแต่ละปี การปรับเพิ่มกำลังผลิตสำรองจะทำให้ต้นทุนผลิตไฟฟ้าโดยรวมสูงขึ้น ทั้งนี้ในความเป็นจริงความผันผวนของโรงไฟฟ้าไม่ได้ผันผวนสูงสุดตลอดเวลา ทำให้การปรับ Spinning Reserve ไว้ที่ค่าผันผวนสูงสุดตลอดเวลาจะเกิดต้นทุนเสียโอกาสในช่วงที่ความผันผวนของโรงไฟฟ้าไม่ได้อยู่ที่ค่าสูงสุดตามสถิติ ดังนั้นเพื่อลดต้นทุนจาก Spinning Reserve การมีระบบการพยากรณ์กำลังผลิตไฟฟ้าของโรงไฟฟ้าพลังงานแสงอาทิตย์ที่มีความแม่นยำทั้งระยะยาว (รายสัปดาห์), ระยะกลาง (1 วันล่วงหน้า) และระยะสั้น (ทุก 15 นาทีถึง 3 ชั่วโมง) จะช่วยสนับสนุนการทำงานของศูนย์ควบคุมระบบกำลังไฟฟ้าแห่งชาติ (NCC) ดังรูปที่ 2 เพื่อบริหารจัดการ Spinning Reserve ให้เหมาะสมตามเวลาและสภาพความผันผวนจริงของระบบไฟฟ้า ได้มีการนำเสนอวิธีการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ในหลากหลายงานวิจัย [1], [2], และ [3] โดยส่วนใหญ่แล้ววิธีการพยากรณ์จะขึ้นอยู่กับกรอบเวลาในการพยากรณ์เป็นสำคัญ สำหรับการพยากรณ์ระยะสั้นอินพุตที่ใช้จะเป็นปริมาณเมฆที่ปกคลุมและทิศทางเคลื่อนที่ของเมฆ ในบริเวณใกล้เคียงกับพื้นที่พยากรณ์ การพยากรณ์ระยะกลางอินพุตที่ใช้จะเป็นข้อมูลสภาพอากาศในบริเวณพื้นที่นั้น เช่น อุณหภูมิ ความชื้น ความเข้มแสงอาทิตย์ ความเร็วลม และ ทิศทางลม เป็นต้น และการพยากรณ์ในระยะยาวอินพุตที่ใช้จะเป็นข้อมูลพยากรณ์อากาศ (weather forecast data) จากการพยากรณ์อากาศเชิงตัวเลข (Numerical Weather Prediction; NWP) [4]

วิทยานิพนธ์นี้จึงเสนอการปรับปรุงประสิทธิภาพแบบจำลองเดี่ยว (แบบจำลองเฉพาะสำหรับแต่ละโรงไฟฟ้า) และวิธีการพยากรณ์กำลังผลิตไฟฟ้ายุคใหม่ด้วยโครงข่ายประสาทเทียม (Neural

Network ; NN) โดยประยุกต์ใช้ 4 เทคนิค ได้แก่ 1) ประยุกต์ใช้อินพุตจากหลายแบบจำลองพยากรณ์อากาศเพื่อผสมผสานข้อดีของแต่ละแบบจำลองด้วยวิธีการเรียนรู้แบบรวมกลุ่ม (Ensemble Learning Method ; ELM) เพื่อปรับปรุงค่าความเอนเอียงและค่าความแปรปรวนของแบบจำลอง 2) ประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm ; GA) ในการคัดเลือกตัวแปร (Feature Selection ; FS) เพื่อคัดเลือกตัวแปรที่เกี่ยวข้องกับการพยากรณ์กำลังผลิตไฟฟ้าเนื่องจากตัวแปรที่ไม่เกี่ยวข้องจะทำให้ประสิทธิภาพของแบบจำลองแย่ลง 3) ประยุกต์ใช้เคมีนคลัสเตอร์ริง (K-means clustering) เพื่อคัดเลือกค่าพยากรณ์อากาศรอบโรงไฟฟ้าเพื่อลดค่าความผันผวน 4) ประยุกต์ใช้ GA ปรับจูนไฮเปอร์พารามิเตอร์ของแบบจำลองเพื่อปรับปรุงประสิทธิภาพแบบจำลองและ 5) ผสมผสานค่าพยากรณ์จากแต่ละโมเดลเข้าด้วยกันด้วย NN เพื่อลดความผันผวนของค่าพยากรณ์

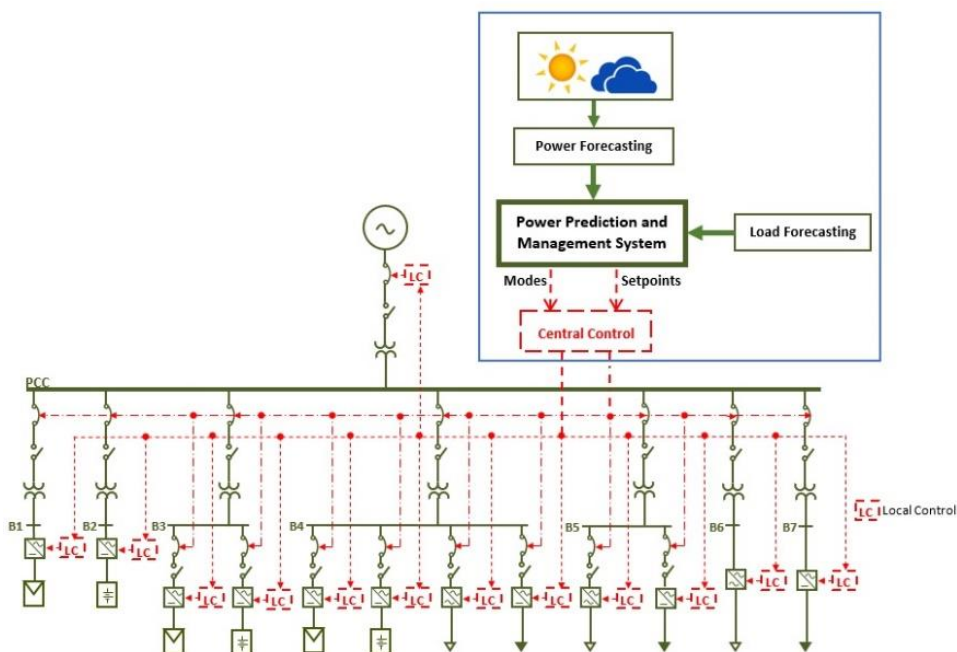


รูปที่ 1 ที่มาและความสำคัญของปัญหา

1.2 วัตถุประสงค์

- 1.2.1 การพัฒนาแบบจำลองเดี่ยว (Single Solar Power Forecast, SSPF) สำหรับการพยากรณ์พลังงานแสงอาทิตย์ของโรงไฟฟ้า
- 1.2.2 การพัฒนาแบบจำลองภูมิภาค (Regional Solar Power Forecast, RSPF) สำหรับการพยากรณ์พลังงานแสงอาทิตย์รวมของ 7 โรงไฟฟ้า

- 1.2.3 การปรับปรุงประสิทธิภาพแบบจำลองด้วยการปรับไฮเปอร์พารามิเตอร์และการเรียนรู้แบบรวมกลุ่ม
- 1.2.4 การประยุกต์ใช้แบบจำลอง SSPF เพื่อพยากรณ์พลังงานแสงอาทิตย์บนอาคารบุญรอดบิณทสัณฑ์
- 1.2.5 การประยุกต์ใช้แบบจำลอง RSPF เพื่อพยากรณ์พลังงานแสงอาทิตย์ของ 7 โรงไฟฟ้า



รูปที่ 2 ระบบพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์

ที่มา : https://smu.ca/engineeringlab/lcsm_research.html

1.3 ขอบเขตวิทยานิพนธ์

- 1.3.1 การพัฒนาแบบจำลอง SSPF สำหรับการพยากรณ์พลังงานแสงอาทิตย์ของโรงไฟฟ้าเดี่ยวและแบบจำลอง RSPF สำหรับการพยากรณ์พลังงานแสงอาทิตย์แบบรายภาค โดยการโปรแกรมลงบนภาษา Python ผ่านเครื่องมือ jupyter Notebook, TensorFlow, และ Keras โดยการประมวลผลส่วนใหญ่ทำบน Google Cloud และ Google Colab
- 1.3.2 แบบจำลอง SSPF สามารถให้รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error; RMSE) ที่ดีกว่าแบบจำลองของงานวิจัย “Influential

Variable Selection for Improving Solar Forecasts from Numerical Weather Prediction ปี 2018” อย่างน้อย 3 เปอร์เซนต์

- 1.3.3 วิธีการปรับปรุงประสิทธิภาพที่นำเสนอสามารถช่วยปรับปรุงประสิทธิภาพของแบบจำลอง SSPF และ RSPF ได้
- 1.3.4 ทดสอบ วิเคราะห์ และประเมินผลแบบจำลอง SSPF และ RSPF

1.4 ขั้นตอนและวิธีการดำเนินงาน

- 1.4.1 การศึกษาบทความทางวิชาการและงานวิจัยที่เกี่ยวข้องกับวิทยานิพนธ์
- 1.4.2 การสร้างแบบจำลอง SSPF และ RSPF
- 1.4.3 การพัฒนาวิธีการเพื่อปรับปรุงประสิทธิภาพแบบจำลอง
- 1.4.4 การทดสอบประสิทธิภาพแบบจำลอง
- 1.4.5 การสรุปและวิเคราะห์ประสิทธิภาพแบบจำลอง

1.5 ประโยชน์ที่คาดว่าจะได้รับจากวิทยานิพนธ์

- 1.5.1 แบบจำลอง SSPF และ RSPF
- 1.5.2 การนำแบบจำลองที่ได้ไปใช้เป็นเครื่องมือสำหรับการพยากรณ์

1.6 เนื้อหาของวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้ได้แบ่งออกเป็น 5 บท โดยแบ่งเป็นเนื้อหา ดังนี้

- บทที่ 1 บทนำ แบ่งเป็น 6 หัวข้อ ได้แก่ 1) ที่มาและความสำคัญของปัญหา 2) วัตถุประสงค์ 3) ขอบเขตวิทยานิพนธ์ 4) ขั้นตอนการศึกษาและวิธีการดำเนินงาน 5) ประโยชน์ที่คาดว่าจะได้รับจากวิทยานิพนธ์ และ 6) เนื้อหาวิทยานิพนธ์
- บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง แบ่งเป็น 2 หัวข้อ ได้แก่ 1) งานวิจัยที่เกี่ยวข้อง 2) ความรู้พื้นฐานที่เกี่ยวข้อง
- บทที่ 3 วิธีดำเนินการวิจัย แบ่งเป็น 5 หัวข้อ ได้แก่ 1) ขั้นตอนการเตรียมข้อมูล 2) ขั้นตอนการฝึกสอนแบบจำลอง 3) ขั้นตอนการรวมกำลังผลิตจากแต่ละโรงไฟฟ้า 4) ขั้นตอนการรวมกำลังผลิตรวมจากแต่ละแบบจำลอง 5) ขั้นตอนการวัดผล
- บทที่ 4 ผลการวิจัย แบ่งเป็น 3 หัวข้อ ได้แก่ 1) แบบจำลอง Single Solar Power Forecast (SSPF) 2) แบบจำลอง Regional Solar Power Forecast (RSPF) 3) ผลการปรับปรุงประสิทธิภาพ
- บทที่ 5 สรุปผลการวิจัย

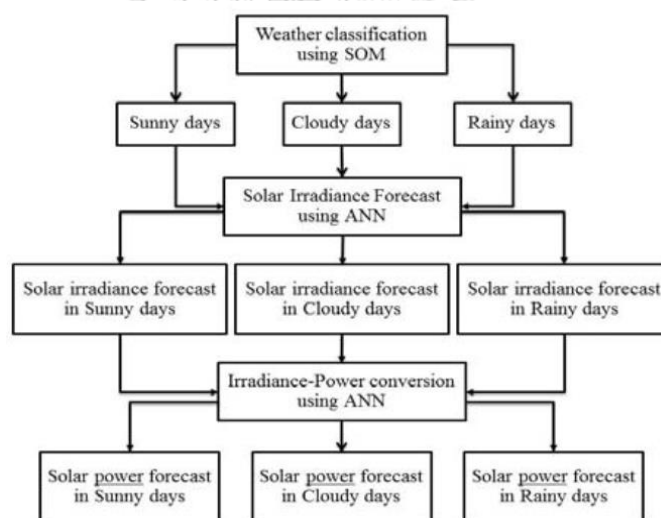
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับวิทยานิพนธ์ฉบับนี้สามารถแบ่งได้เป็น 2 หัวข้อ คือ 1) ปัญหาและวิธีการพยากรณ์ และ 2) การปรับจูนไฮเปอร์พารามิเตอร์ของแบบจำลอง SSPF และ RSPF

2.1.1 ปัญหาและวิธีการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์

งานวิจัย [5] ได้เสนอวิธีการพยากรณ์กำลังผลิตไฟฟ้าของพลังงานแสงอาทิตย์ (500 kW) ที่แม่ฮ่องสอน โดยใช้ Self-Organized Map (SOM) ในการจำแนกสภาพอากาศของแต่ละช่วงเวลา (00.00 – 12.00 น. และ 12.00 – 24.00 น.) ออกเป็น 3 กลุ่ม เพื่อปรับปรุงความแม่นยำในการพยากรณ์ และใช้โมเดล ANN ในการพยากรณ์กำลังผลิตไฟฟ้า (ล่วงหน้า 3 ชั่วโมง) ที่ผ่านการฝึกสอนด้วยชุดข้อมูลที่สอดคล้องกับผลการจำแนกสภาพอากาศ พบว่าค่า %RMSE ในกรณีที่ไม่พิจารณาปัจจัยเชิงฤดูกาลเท่ากับ 40.14% ส่วนกรณีที่พิจารณาปัจจัยเชิงฤดูกาลพบว่ามีค่าเท่ากับ 34.71%



รูปที่ 3 วิธีการพยากรณ์โดยการใช้ ANN ร่วมกับ SOM

ที่มา : S. Nitisanon and N. Hoonchareon,

2017 IEEE Power & Energy Society General Meeting

งานวิจัย [1] ได้เสนอวิธีทางสถิติ 3 วิธี (Partial Correlation, Stepwise Regression, และ Subset Regression) สำหรับการคัดเลือกตัวแปรสภาพอากาศที่สำคัญสำหรับการพยากรณ์ความเข้มแสงอาทิตย์เพื่อปรับปรุงค่าพยากรณ์ความเข้มแสงอาทิตย์ของ NWP และใช้ค่าเฉลี่ยเชิงพื้นที่เพื่อลด

ความแปรปรวนของค่าพยากรณ์ความเข้มแสงอาทิตย์บริเวณพื้นที่สนใจและใช้โมเดลถดถอยเชิงเส้นอธิบายความสัมพันธ์ระหว่างความเข้มแสงอาทิตย์และตัวแปรสภาพอากาศที่เกี่ยวข้อง พบว่าการเฉลี่ยเชิงพื้นที่ของค่าพยากรณ์อากาศครอบคลุม 7x7 grid points ให้ค่า NRMSE ลดลง 3.42% และ แบบถดถอยเชิงเส้นที่มีและไม่มีการคัดเลือกตัวแปรสามารถปรับปรุงประสิทธิภาพเพิ่มเติมจากการเฉลี่ยเชิงพื้นที่ ทำให้แบบจำลองมีค่า NRMSE อยู่ที่ 41.32% และ 41.89% NRMSE ตามลำดับ และตัวแปรที่มีความเกี่ยวข้องได้แก่ ความเข้มแสงอาทิตย์, อุณหภูมิ ความชื้นสัมพัทธ์ และ UV index

งานวิจัย [6] ได้เสนอวิธีผสมผสานแบบอนาล็อก (Analog Ensemble Method) สำหรับพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ในระดับภูมิภาค (Southeastern Massachusetts) ล่วงหน้า 1 วันด้วยความละเอียดรายชั่วโมง โดยใช้วิธีการผสมผสาน (blending) และการจัดกลุ่ม (Clustering) เพื่อปรับปรุงความแม่นยำซึ่งข้อมูลที่ใช้ในการพยากรณ์ประกอบไปด้วย ข้อมูลพยากรณ์อากาศจาก 3 แบบจำลองการพยากรณ์อากาศเชิงตัวเลข (NWP) ที่แตกต่างกัน ได้แก่ The North American Mesoscale Forecast System (NAM Forecast), The Global Forecast System (GFS Forecast), และ The Short-Range Ensemble Forecast (SREF Forecast) ข้อมูลกำลังผลิตไฟฟ้าที่วัดได้และข้อมูลเกี่ยวกับดาราศาสตร์ (solar time และ earth declination angle) พบว่าวิธีที่บทความนี้แนะนำเสนอให้ค่า NRMSE ลดลง 13.80% เป็น 61.21% เมื่อเปรียบเทียบกับ 3 baselines ที่ทำการทดสอบ บทความ [4] ได้อธิบายถึงระบบตัวอย่างที่ถูกนำมาใช้ในการพยากรณ์สำหรับการไฟฟ้าแต่ละแห่งและวิธีกำหนดค่าเพื่อตอบสนองความต้องการของการไฟฟ้าและพิจารณาเกี่ยวกับปัญหาของ Big Data

งานวิจัย [7] ได้ประยุกต์ใช้วิธีการเรียนรู้เครื่องจักรสำหรับการพยากรณ์กำลังผลิตไฟฟ้าในระดับภูมิภาค (Southeastern Massachusetts (SEMA) load zone) ล่วงหน้า 2 วันด้วยความละเอียด 1 ชั่วโมง โดยนำแบบจำลองกายภาพ (Physical PV Power Model) ของแต่ละโซลาร์ฟาร์มมาแบ่งเป็นคลัสเตอร์ตามลักษณะทางภูมิศาสตร์แล้วหาผลรวมของทุกคลัสเตอร์ซึ่งข้อมูลที่ใช้ในการพยากรณ์ประกอบไปด้วย ข้อมูลพยากรณ์อากาศจาก NWP สำหรับคำนวณ Plane of Array (POA) Irradiance และ PV Output Power และประยุกต์ใช้วิธีถดถอยเชิงเส้น (Linear Regression Method) สำหรับปรับค่า Bias และประยุกต์ใช้เทคนิคต่างๆ ได้แก่ การคัดเลือกพารามิเตอร์ที่สำคัญเพื่อปรับปรุงการพยากรณ์ ผสมผสานค่าพยากรณ์จาก NWP ประเภทต่างๆตามค่าน้ำหนักการผสมที่เหมาะสมเพื่อให้ได้ค่าพยากรณ์กำลังผลิตไฟฟ้าที่สามารถปรับเปลี่ยนได้ตามช่วงเวลา พบว่าค่า NRMSE อยู่ที่ 5.28% สำหรับการพยากรณ์ล่วงหน้า 2 วัน ซึ่งลดลง 30.6% เมื่อเปรียบเทียบกับ baseline ซึ่งเป็นแบบจำลองเดี่ยวที่ดีที่สุด

งานวิจัย [8] ได้นำเสนอวิธีการพยากรณ์กำลังผลิตไฟฟ้าในระดับภูมิภาครัฐโอคลาโฮมาด้วย 4 แบบจำลอง คือ Simple Linear Regression, Lasso Regression, Ridge Regression และ

Random Forests และประยุกต์ใช้เทคนิคต่างๆ ได้แก่ 1) ใช้วิธีการหาค่าประมาณเชิงพื้นที่ (Spatial Interpolation) เพื่อประมาณค่าข้อมูลพยากรณ์อากาศรอบโรงไฟฟ้าจากแบบจำลอง GEFS โดยใช้จุดพยากรณ์อากาศ 4 จุดใกล้ที่สุด 2) นำข้อมูลพยากรณ์อากาศในแต่ละจุด (4 จุดที่ใกล้ที่สุด) ไปฝึกสอนแบบจำลองแยกกันอย่างอิสระแล้วนำค่าพยากรณ์จากแต่ละแบบจำลองมาหาค่าเฉลี่ยและ 3) การทำฟีเจอร์เอ็นจิเนียริงเพื่อเพิ่มตัวแปร (Time, Location, และ Altitude) โดยข้อมูลอินพุตใช้ประกอบไปด้วย ข้อมูลค่าวัดจริงของพลังงานไฟฟ้าที่ผลิตได้จากแต่ละโซลาร์ฟาร์ม (98 แห่ง) และข้อมูลพยากรณ์อากาศจากแบบจำลอง GEFS พบว่า Random Forest methods ให้ความแม่นยำในการพยากรณ์มากกว่า Linear Regression อยู่ 6% และการทำฟีเจอร์เอ็นจิเนียริงช่วยลดค่า MAE ลง 9% เมื่อเปรียบเทียบกับแบบจำลอง Random Forests ที่ใช้เพียงแค่ข้อมูล GEFS เพียงอย่างเดียว

งานวิจัย [9] ได้เสนอวิธี Real-Time Reforecast (Artificial Neural Network; ANN) สำหรับพยากรณ์กำลังผลิตไฟฟ้าของโรงไฟฟ้าพลังงานแสงอาทิตย์ (48 MW) แบบ Intra-Hour Prediction และประยุกต์ใช้ ANN เพื่อปรับปรุงประสิทธิภาพของ 3 แบบจำลองพยากรณ์พื้นฐาน (Baseline Prediction Models) ได้แก่ 1) Physical Deterministic Model Based on Cloud Tracking Techniques; 2) Autoregressive Moving Average (ARMA) model; และ 3) k-th Nearest Neighbor (kNN) model โดยการใช้ Measured Power Data พบว่า 3 แบบจำลองพยากรณ์พื้นฐานที่นำเสนอมีประสิทธิภาพเพิ่มขึ้นสำหรับการพยากรณ์ 5, 10, และ 15 นาที

งานวิจัย [10] ได้เสนอรูปแบบการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์โดยใช้การพยากรณ์ความเข้มแสงอาทิตย์ล่วงหน้า 1 ชั่วโมงด้วยปัจจัยทางสภาพอากาศต่างๆ รวมไปถึงการปกคลุมของเมฆ (Cloud Cover) โดยการใช้แบบจำลอง Support Vector Machine (SVM) และมีการประยุกต์ใช้ K-Means Clustering Algorithm เพื่อแบ่งกลุ่มข้อมูลของสภาพอากาศรายวันที่มีความคล้ายคลึงกันเป็น 3 คลัสเตอร์ และจะใช้ข้อมูลที่อยู่ในคลัสเตอร์เดียวกันฝึกสอนแบบจำลอง SVM Regression พบว่าวิธีการที่นำเสนอได้ปรับปรุงความแม่นยำในการพยากรณ์

2.2.2 การปรับจูนไฮเปอร์พารามิเตอร์ของเน็ตเวิร์ก

งานวิจัย [11] ได้ประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรม (Neural-Genetic Algorithm) สำหรับการคัดเลือกตัวแปร (Feature Selection) ให้กับแบบจำลองโครงข่ายประสาทเทียมจำแนกประเภท (Neural Network Based Classifier) พบว่าความแม่นยำในการจำแนกข้อมูลเพิ่มขึ้น 3.3% เมื่อเทียบกับการใช้ตัวแปรทั้งหมดงานวิจัย [12] ได้ประยุกต์ใช้ MGA (GA ที่เขาได้ทำการออกแบบเอง) สำหรับการคัดเลือกตัวแปรเพื่อพยากรณ์ความต้องการของผู้ป่วยสำหรับแหล่งข้อมูลสำคัญต่าง ๆ ใน OPD โดยประยุกต์ใช้วิธีการคัดเลือกตัวแปรและวิธีการเรียนรู้แบบลึกโดยเปรียบเทียบกับ GA และ PCA พบว่า MGA ปรับปรุงประสิทธิภาพการคัดเลือกตัวแปรทำให้มีประสิทธิภาพดีขึ้นในขณะที่ใช้ตัว

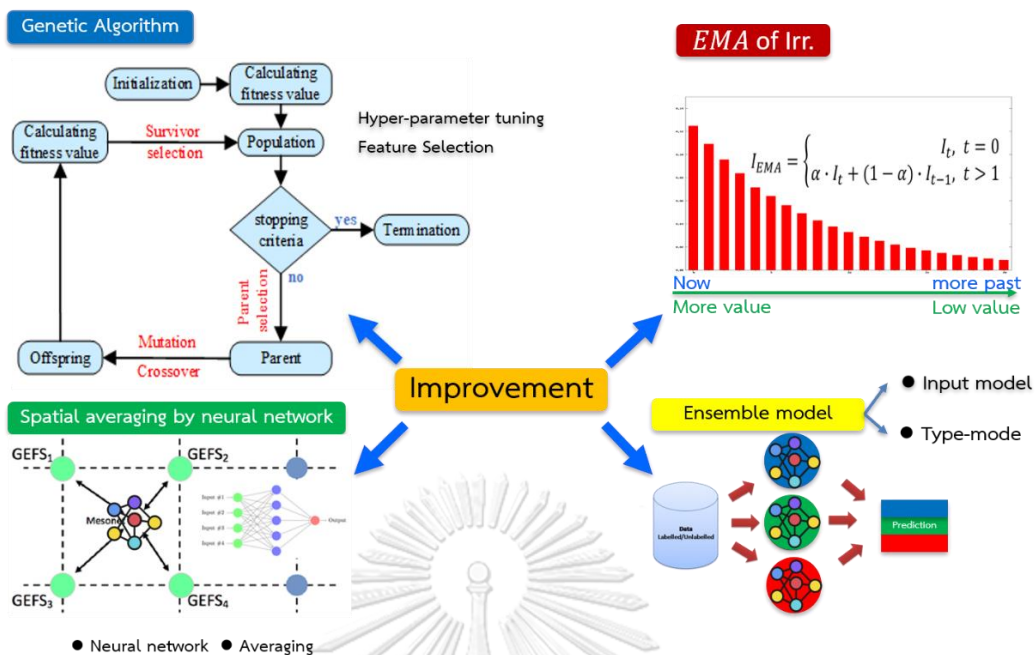
แปรน้อยลง งานวิจัย [13] เสนอวิธีการแบบผสมผสานเพื่อพยากรณ์ความต้องการพลังงานไฟฟ้าในอนาคตซึ่งประยุกต์ใช้ GA เพื่อหาชุดตัวแปรย่อยที่เหมาะสมให้กับระบบอนุกรมระบบประสาท-ฟัซซีแบบปรับตัว (Adaptive Neuro-Fuzzy Inference System) งานวิจัย [14] ประยุกต์ใช้ GA และ Ant Colony Optimization (ACO) สำหรับการคัดเลือกตัวแปรในปัญหาการพยากรณ์ความต้องการไฟฟ้า โดยประยุกต์ใช้ ANN เป็นโมเดลพยากรณ์ งานวิจัย [15] เสนอเฟรมเวิร์คของ GA-SVR สำหรับพยากรณ์ความต้องการห้องพักของโรงแรมซึ่งประยุกต์ใช้ GA ในการคัดเลือกตัวแปร ปรับพารามิเตอร์ และ เลือกรูปแบบที่ประหยัด งานวิจัย [16] เสนอวิธีพยากรณ์ความต้องการของอุตสาหกรรมค้าปลีกในจีน โดยประยุกต์ใช้ GA ในการคัดเลือกตัวแปรและค้นหาความรู้ในพื้นที่ยอดฮิตอื่น ๆ เช่น การวิเคราะห์ตลาดการเงิน งานวิจัย [17] พบว่าการคัดเลือกตัวแปรด้วย GA มีประสิทธิภาพเหนือกว่า CATLRN ในการระบุตัวแปรที่เกี่ยวข้องเพื่อประเมินสถานะการเงินของธนาคาร งานวิจัย รวม GA เข้ากับ PCA และ Decision Trees เพื่อสร้างตัวเลือกตัวแปรสำหรับการทำนายราคาหุ้น งานวิจัย [18] เสนอการคัดเลือกตัวแปรด้วย GA แบบใหม่เพื่อทำนายความทุกข์ทางการเงินโดยใช้ GA-based wrapper ร่วมกับ Te-Sheng Li (2006) ประยุกต์ใช้ GA ในการคัดเลือกตัวแปรโดยการปรับเวกเตอร์ตัวแปรให้เหมาะสม ลบตัวแปรที่ไม่เกี่ยวข้องและซ้ำซ้อนออกและค้นหาชุดตัวแปรที่ดีที่สุดเพื่อป้อนให้กับโมเดลการจำแนกข้อมูลด้วย NN พบว่าโมเดลจำแนกข้อมูลด้วย GA มีความแข็งแกร่งและมีประสิทธิภาพในการค้นหาตัวแปรย่อยที่เหมาะสมจากชุดข้อมูลขนาดใหญ่ งานวิจัย [19] พบว่าความสามารถของระบบ Hybrid GA/ANN ในการค้นหาตัวแปรที่สำคัญขึ้นอยู่กับการกำหนดค่าของจำนวนประชากร การประเมินค่าความเหมาะสมและการเลือกฟังก์ชันกระตุ้นของ NN ซึ่งงานวิจัย [20] ได้ให้ความเห็นว่าขนาดของประชากรไม่ใช่ปัจจัยสำคัญในความสำเร็จของ GA หากขนาดประชากรไม่เล็กเกินไป (เช่น < 40 โครโมโซม) อย่างไรก็ตามผลลัพธ์ที่ระบุถึงความสัมพันธ์ของขนาดประชากรในความสำเร็จของ GA ถูกกล่าวใน [21] งานวิจัย [22] ประยุกต์ใช้ GA ในการคัดเลือกตัวแปรที่มีประโยชน์และมีความสำคัญซึ่งตัวแปรเหล่านี้จะถูกส่งไปยังตัวจำแนกข้อมูล หลังจากนั้นจะทำการประเมินประสิทธิภาพตัวจำแนกข้อมูลเพื่อลดขนาดมิติของตัวแปรและปรับปรุงประสิทธิภาพของโมเดลจำแนกข้อมูลซึ่งตัวแปรที่ไม่เกี่ยวข้องและไม่จำเป็นมีผลต่อประสิทธิภาพของโมเดลในงานวิจัย [23] ประยุกต์ใช้ GA และ SVM เพื่อแก้ปัญหาเกี่ยวกับดวงดาวและอวกาศ โดยประยุกต์ใช้ GA เพื่อเลือกตัวแปรที่เกี่ยวข้องและค้นหาพารามิเตอร์ที่เหมาะสมของ SVM งานวิจัย [24] ประยุกต์ใช้ GA เพื่อหาจำนวน layer ที่เหมาะสมของโมเดล LSTM งานวิจัย [25] บทความนี้ประยุกต์ใช้ GA เพื่อกำจัดตัวแปรที่ซ้ำซ้อนและไม่สำคัญ และป้อนให้กับโมเดล BPNN เพื่อปรับปรุงประสิทธิภาพ

จากงานวิจัยที่เกี่ยวข้องข้างต้นชี้ให้เห็นว่าโรงไฟฟ้าพลังงานแสงอาทิตย์มีความผันผวนสูงซึ่งโรงไฟฟ้าที่มีอยู่เดิมในระบบไม่สามารถตอบสนองต่อความผันผวนของกำลังผลิตไฟฟ้าจากโรงไฟฟ้า

ดังกล่าวได้ทัน หากมีโรงไฟฟ้าพลังงานแสงอาทิตย์เข้ามาในระบบในสัดส่วนที่มากขึ้นจะมีแนวโน้มทำให้เกิดผลกระทบในการควบคุมคุณภาพไฟฟ้าของระบบมากยิ่งขึ้น โดยเฉพาะในอนาคตเมื่อระบบไมโครกริดขยายตัวไปในทุกพื้นที่ซึ่งโดยทั่วไประบบไมโครกริดจะเป็นระบบไฟฟ้าขนาดเล็ก มีเพียงโรงไฟฟ้าแบบกระจายตัวขนาดเล็กทำหน้าที่ควบคุมความถี่ไฟฟ้าของระบบทำให้ระบบมีความเฉื่อยต่ำ จึงจำเป็นที่จะต้องมีการออกแบบระบบพยากรณ์กำลังผลิตไฟฟ้าที่มีความแม่นยำซึ่งจะทำให้สามารถเตรียมกำลังผลิตสำรองเพื่อชดเชยกำลังไฟฟ้าที่ผันผวนให้แก่ระบบเป็นผลให้สามารถตอบสนองความต้องการกำลังผลิตไฟฟ้าได้ทันทำให้ระดับแรงดันและความถี่ไฟฟ้าอยู่ในเกณฑ์มาตรฐานสำหรับโรงไฟฟ้าพลังงานแสงอาทิตย์ 8 แห่ง ได้แก่ บริษัทพัฒนาพลังงานธรรมชาติจำกัด 55 MW, บริษัทพีซีพีจีจำกัด 30 MW, บริษัทอีเอโซล่านครสวรรค์จำกัด 90 MW, บริษัทเสริมสร้างพลังงานจำกัด 40 MW, บริษัทอีเอโซล่าปางจำกัด 90 MW, บริษัทเอสพีพีชิจากัด 41 MW, บริษัทอีเอโซล่าพิษณุโลกจำกัด 90 MW, และจุฬาลงกรณ์มหาวิทยาลัย 8 kW ซึ่งคิดเป็นกำลังผลิตรวม 436.008 MW วิทยานิพนธ์นี้จึงนำเสนอระบบการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ในระบบภูมิภาคและแบบแยกเดี่ยว (แต่ละโรงไฟฟ้า) ด้วยโครงข่ายประสาทเทียมโดยประยุกต์ใช้ 4 เทคนิค ได้แก่ 1) การประยุกต์ใช้อินพุตจากหลายโมเดลพยากรณ์อากาศเพื่อผสมผสานข้อดีของแต่ละโมเดลด้วยวิธีการเรียนรู้แบบรวมกลุ่ม (Ensemble Learning Method) เพื่อลดค่า Bias ของแต่ละโมเดล 2) การประยุกต์ใช้ GA ในการคัดเลือกตัวแปรเพื่อคัดเลือกตัวแปรที่เกี่ยวข้องและสำคัญต่อการพยากรณ์เนื่องจากตัวแปรที่ไม่เกี่ยวข้องจะทำให้ประสิทธิภาพของโมเดลแยกลง 3) ประยุกต์ใช้เคมีนคลัสเตอร์ริง (K-means clustering) เพื่อคัดเลือกค่าพยากรณ์อากาศรอบโรงไฟฟ้าเพื่อลดค่าความผันผวน 4) ประยุกต์ใช้ GA ปรับจูนไฮเปอร์พารามิเตอร์ของโมเดลเพื่อปรับปรุงประสิทธิภาพของโมเดล และ 5) ผสมผสานค่าพยากรณ์จากแต่ละโมเดลเข้าด้วยกันด้วย NN เพื่อลดความผันผวนของโมเดลอีกทั้งวิทยานิพนธ์นี้ได้ทำการศึกษาและวิเคราะห์ประสิทธิภาพของโมเดลแยกเดี่ยวโดยเปรียบเทียบกับวิจัย [1] และ [5]

2.2 ความรู้พื้นฐานที่เกี่ยวข้อง

วิทยานิพนธ์ฉบับนี้มีการประยุกต์ความรู้ในหลายสาขาวิชา ได้แก่ การรู้จำรูปแบบ (Pattern Recognition), การเรียนรู้แบบเสริมแรง (Reinforcement Learning, RL) และ ขั้นตอนวิธีเชิงวิวัฒนาการ (Evolution Algorithm) ซึ่งทั้งหมดเป็นส่วนหนึ่งของวิชาปัญญาประดิษฐ์ (Artificial Intelligence, AI) สำหรับวิธีการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์และการปรับปรุงประสิทธิภาพแบบจำลองแสดงดังรูปที่ 4



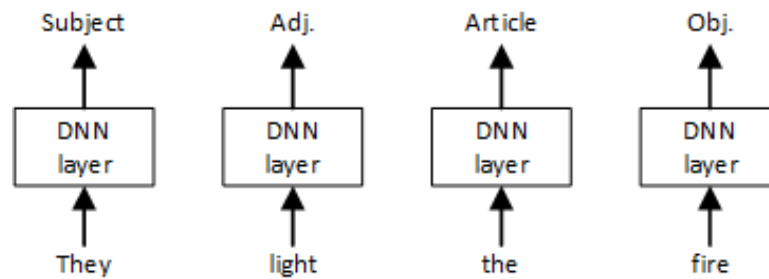
รูปที่ 4 วิธีการปรับปรุงประสิทธิภาพแบบจำลอง

2.2.1 การรู้จำรูปแบบ (Pattern Recognition)

Pattern Recognition คือการรู้จำรูปแบบต่างๆโดยอาศัยฐานข้อมูลและทฤษฎีของ Discriminative Modeling เช่น Deep Learning ซึ่งถูกนำไปใช้ในงานด้านต่างๆ เช่น bioinformatics, natural language processing และ computer vision เป็นต้น ซึ่งมีข้อดีในหลากหลายแง่มุมด้วยกัน เช่น การสกัดตัวแปร การอัพเดทแบบจำลอง และการถ่ายทอดการเรียนรู้

2.2.1.1 ปัญหาของ Fully Connected Layer (FCL)

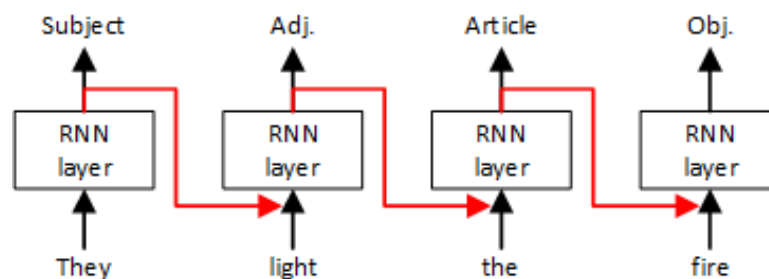
ตัวอย่างของปัญหาที่เกิดขึ้นใน FCL แสดงดัง รูปที่ 5 ซึ่งคือโจทย์ของการหาชนิดของคำ (Part of Speech) ของคำในประโยค เช่น ถ้าเราป้อน “They” ไปยังเน็ตเวิร์กธรรมดา (Dense Neural Network, DNN) เราได้ “Subject” ออกมาและที่เหลืออีกสามคำได้เป็น “Adjective”, “Article” และ “Object” ตามลำดับซึ่งเราจะเห็นได้ว่าคำว่า “light” ผิด เนื่องจากว่าปกติคำว่า “light” จะใช้ในรูปแบบของคำคุณศัพท์ที่มีความหมายว่า “เบา” หรือ “แสง” ดังนั้นการมองที่ละคำแยกจากกันจะให้ผลลัพธ์ที่ผิด เนื่องจาก DNN แต่ละเซลล์ไม่มีเกี่ยวข้องทางเวลาซึ่งจากปัญหานี้ทำให้เกิดแนวความคิดที่ว่า ถ้าเราสามารถจำได้ว่าคำก่อนหน้านี้เป็นคำชนิดใด เราจะสามารถบอกได้ว่าคำว่า “light” มันควรจะชนิดของคำเป็น “Verb” เนื่องจากว่ามันอยู่ถัดจาก “Subject” ด้วยเหตุนี้จึงมีคนพัฒนาเน็ตเวิร์กแบบวนซ้ำ (Recurrent Neural Network, RNN) ขึ้นมาเพื่อที่จะจดจำสิ่งที่เกิดขึ้นในสถานะ (State) ก่อนหน้าดังแสดงในรูปที่ 6



รูปที่ 5 ข้อเสียของเน็ตเวิร์ก FCL

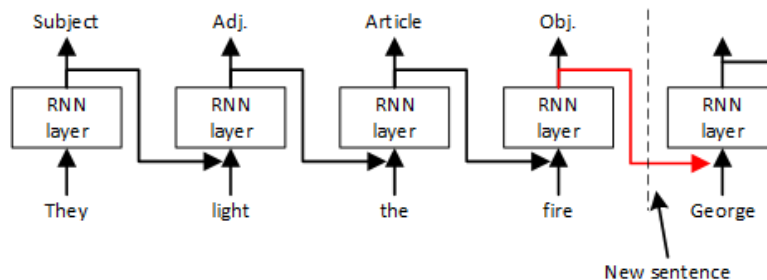
ที่มา : วิชา Pattern Recognition L10 CNN RNN, อ.ดร.เอกพล ช่วงสุวนิช

ซึ่งรูปที่ 6 แสดงลักษณะการต่อกันของเซลล์ RNN ซึ่งจะเห็นได้ว่าแต่ละเซลล์ RNN จะมี 2 อินพุต ได้แก่ สถานะก่อนหน้า (previous state) และ สถานะปัจจุบัน (current state) การมีอินพุตที่เป็นสถานะก่อนหน้าทำให้มันสามารถที่จะจดจำสิ่งที่เกิดขึ้นหรือสถานะก่อนหน้าได้ ผลลัพธ์ที่เกิดขึ้นก็คือมันสามารถบอกได้ว่าคำว่า “light” ควรจะมีชนิดของคำเป็น “Verb” เนื่องจากว่ามันสามารถจำได้ว่าสถานะก่อนหน้ามีชนิดของคำเป็น “Subject” แต่อย่างไรก็ตามการต่อกันของเซลล์ RNN ก็ยังไม่สามารถหาความสัมพันธ์ระหว่างสถานะที่อยู่ไกลกันมากได้ เนื่องจากการทำ Backpropagation ซึ่งเป็นการย้อนความผิดพลาดกลับมาเพื่อปรับค่าน้ำหนักของเส้นเชื่อมในแต่ละเลเยอร์จะถูกทำให้หายไป (Vanish) เนื่องจากการคูณค่าที่ใกล้ศูนย์หลายๆรอบในลักษณะของรีเคอร์ซีฟหรือที่คนส่วนใหญ่รู้จักกันในชื่อของปัญหา “Vanishing Gradient Problem”



รูปที่ 6 การต่อกันของแต่ละเซลล์ RNN

ที่มา : วิชา Pattern Recognition L10 CNN RNN, อ.ดร.เอกพล ช่วงสุวนิช



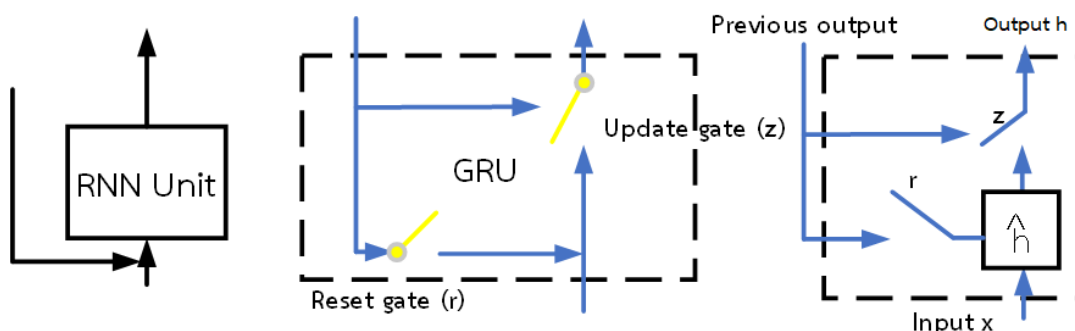
รูปที่ 7 ปัญหาการจำสถานะเก่าที่ไม่เกี่ยวข้องกับสถานะปัจจุบันของเน็ตเวิร์ก RNN

ที่มา : วิชา Pattern Recognition L10 CNN RNN, อ.ดร.เอกพล ช่วงสุนิช

รูปที่ 7 แสดงปัญหาการจำสิ่งที่ไม่จำเป็นที่เกิดขึ้นในเลเยอร์ของ RNN เราจะเห็นได้ว่าแม้ว่าแต่ละเซลล์ RNN จะสามารถที่จะจำสถานะก่อนหน้าได้ แต่ก็ยังคงมีปัญหาที่เกิดขึ้นตามมา นั่นคือการที่เซลล์ RNN มองเพียงแค่สถานะก่อนหน้าเพียงแค่ time-step เดียว จะทำให้เกิดปัญหาที่แสดงดังรูปที่ 7 นั่นคือเซลล์ RNN จะมองคำว่า “fire” มีชนิดของคำเป็น “Verb” ซึ่งจะทำให้เซลล์ RNN ถัดมามอง “George” ว่าควรมีสถานะเป็น “Object” ทั้งๆที่จริงๆแล้ว คำว่า “fire” มีชนิดของคำเป็น “object” เนื่องจากเป็นกรรมของ verb (light) และ “George” ควรจะมีชนิดของคำเป็น “Subject” เนื่องจากการขึ้นต้นประโยคใหม่ ดังนั้นเซลล์ RNN ควรที่จะลืมสถานะก่อนหน้าที่มันจำอยู่ เพื่อที่จะเริ่มต้นจำสิ่งใหม่ ด้วยปัญหาที่เกิดขึ้นนี้ทำให้มีการพัฒนาเซลล์ที่มีชื่อว่า Gated Recurrent unit (GRU) ขึ้นมาดังแสดงในรูปที่ 8 ซึ่งภายในเซลล์ GRU จะมีเกต (gate) เพิ่มขึ้นมา 2 เกตจากเซลล์ RNN (ไม่มี gate) ได้แก่ เกตลืม (forget gate) และ เกตอัปเดต (update gate)

1. เกตลืม (forget gate) คือเกตที่ทำหน้าที่ตัดสินใจว่าจะลืมสถานะก่อนหน้านี้ไปกี่เปอร์เซ็นต์ หรือพูดอีกนัยหนึ่งก็คือการลืมสิ่งที่จำมาก่อนหน้านี้มากหรือน้อยเท่าใด
2. เกตอัปเดต (update gate) คือเกตที่ทำหน้าที่ตัดสินใจว่าจะเลือกอัปเดตระหว่างสถานะก่อนหน้าและสถานะปัจจุบันในสัดส่วนเท่าไร หรือกล่าวอีกนัยหนึ่งก็คือการเลือกอัปเดตระหว่างสถานะก่อนหน้าและสถานะ (\hat{h}) ในสัดส่วนเท่าใด ดังแสดงใน รูปที่ 8 ทางขวา

ด้วยเกตทั้งสองนี้ทำให้ GRU เซลล์สามารถใช้ประโยชน์จากสถานะก่อนหน้าและปัจจุบันได้อย่างเหมาะสม กล่าวคือเซลล์สามารถตัดสินใจได้ว่าจะให้ความสำคัญกับการลืมสถานะก่อนหน้าหรือการจำสถานะปัจจุบันในสัดส่วนที่เหมาะสมอย่างไร



รูปที่ 8 โครงสร้างทางสถาปัตยกรรมของเซลล์ GRU เทียบกับเซลล์ RNN

ที่มา : วิชา Pattern Recognition L10 CNN RNN, อ.ดร.เอกพล ช่วงสุวนิช

จากรูปที่ 8 ทางซ้ายมือคือโครงสร้างทางสถาปัตยกรรมของเซลล์ RNN รูปกลางคือสถาปัตยกรรมของเซลล์ GRU แบบไม่แสดง \hat{h} และรูปขวาคือสถาปัตยกรรมของเซลล์ GRU แบบแสดง \hat{h}

จากรูปทางซ้ายมือเราจะเห็นได้ว่าเซลล์ RNN มี 2 อินพุต ได้แก่ สถานะก่อนหน้าและสถานะปัจจุบัน และ 1 เอาต์พุต ซึ่งเมื่อเปรียบเทียบกับเซลล์ GRU ในรูปกลางก็จะมี 2 อินพุตและ 1 เอาต์พุต เช่นเดียวกัน แต่มันจะแตกต่างกันตรงที่ภายในเซลล์ เซลล์ GRU จะมี 2 เกต ในขณะที่เซลล์ RNN ไม่มีเกต ซึ่งเกตในที่นี้หมายถึง Sigmoid Function (σ) ซึ่งเป็นฟังก์ชันที่จำกัดค่าให้อยู่ในช่วงจำนวนจริงระหว่าง 0 ถึง 1 ยกตัวอย่างเช่น ถ้าเกตมีค่าเป็น 1 เกตนี้จะยอมให้ค่าที่ผ่านเข้ามาผ่านไปได้อย่างเต็มที่ แต่ถ้าเกตมีค่าเป็น 0 มันจะไม่ยอมให้ค่าผ่านเข้าไปในเกต จากรูปกลางจะเห็นได้ว่า GRU ประกอบไปด้วย 2 เกต ได้แก่ reset Gate (r) และ update gate (z)

2.2.1.2 สมการ Recurrent Neural Network (RNN)

เน็ตเวิร์กแบบ RNN เป็นเน็ตเวิร์กที่เหมาะสมกับข้อมูลที่มีลักษณะเป็นอนุกรมเวลา (time series data) เนื่องจากเน็ตเวิร์กแบบ RNN มีชั้นซ่อนสำหรับจดจำสถานะตามสมการที่ (1) – (2) เมทริกซ์ค่าน้ำหนัก U , V , และ W ได้จากการฝึกสอนเน็ตเวิร์ก (Train network)

$$h_t = \sigma(x_t W + h_{t-1} U) \quad (1)$$

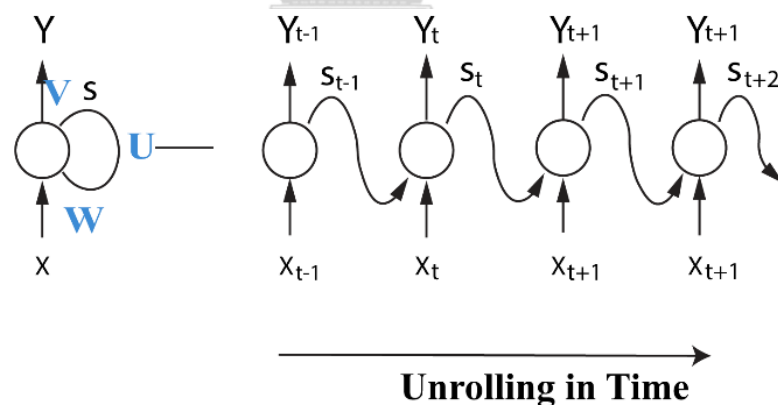
$$o_t = \sigma(h_t V) \quad (2)$$

โดยที่

h_t คือสถานะ ณ เวลา t	x_t คือเวกเตอร์อินพุต
o_t คือเวกเตอร์เอาต์พุต	σ คือฟังก์ชันกระตุ้น (Sigmoid Function)
U คือเมทริกซ์ค่าน้ำหนักของสถานะ	V คือเมทริกซ์ค่าน้ำหนักของอินพุต
W คือเมทริกซ์ค่าน้ำหนักของเอาต์พุต	U, V, W คือ <i>Linear transform</i> ด้วย <i>matrix multiply</i>

2.2.1.3 การคำนวณ Forward Propagation ของเน็ตเวิร์กรีเคอร์ซีฟ

จากรูปที่ 9 เอาต์พุตของสถานะปัจจุบัน (current state, S_t) ณ time step t เกิดจากการรวมกันระหว่าง อินพุตของสถานะปัจจุบัน (x_t) ที่คูณอยู่กับเมทริกซ์ค่าน้ำหนักของสถานะ (U) กับ เอาต์พุตของสถานะก่อนหน้า (previous state, S_{t-1}) ที่คูณอยู่กับเมทริกซ์ค่าน้ำหนักของเอาต์พุต (W) ผ่านฟังก์ชันไม่เชิงเส้น (nonlinearity function) ซึ่งในที่นี้คือฟังก์ชัน \tanh ตามสมการ (3) หลังจากนั้นในสมการ (4) เราจะผ่าน (S_t) ที่คูณด้วยเมทริกซ์ค่าน้ำหนักของอินพุต (V) เข้าไปในฟังก์ชัน softmax ได้ค่าการทำนาย (predicted value, \hat{y}_t) ณ time step t ออกมา (ฟังก์ชัน softmax คือฟังก์ชันที่คืนค่าความน่าจะเป็นของแต่ละคลาส (class) ออกมา ทุกคลาสรวมกันเป็น 1)



รูปที่ 9 ขั้นตอนวิธีการแพร่ไปข้างหน้าของเน็ตเวิร์ก RNN

ที่มา : Fundamentals Of Deep Learning For NLP, อ.ดร.พีรพล เวทีกุล

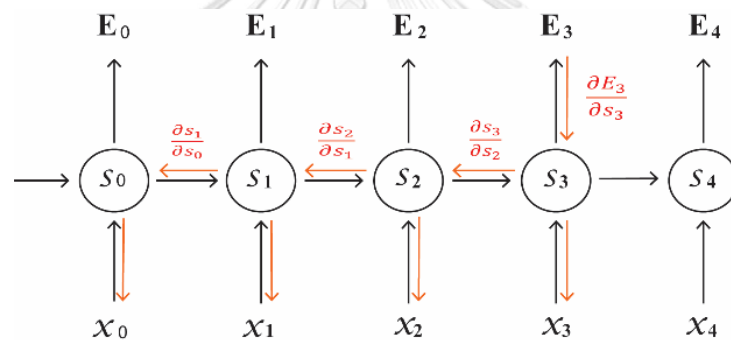
จากรูปที่ 9 ทางด้านซ้ายมือจะเป็นรูปแบบอย่างง่าย (simplify) ของเน็ตเวิร์ก RNN และรูปทางด้านขวามือเป็นการคลี่ (unrolling) รูปทางด้านซ้ายออกมาซึ่งจะแสดงเซลล์ในแต่ละ time-step

$$s_t = \tanh(Ux_t + Ws_{t-1}) \quad (3)$$

$$\hat{y}_t = \text{softmax}(Vs_t) \quad (4)$$

2.2.1.4 การคำนวณ Backpropagation ของเน็ตเวิร์กรีเคอร์ซีฟ

หลังจากผ่านการคำนวณไปด้านหน้า (Forward Propagation) ที่กล่าวไปก่อนหน้านี้ผลลัพธ์ที่ได้คือค่าความผิดพลาด (error) ที่เกิดจากผลต่างระหว่างค่าจริง (y_t) และค่าทำนาย (\hat{y}_t) ซึ่งค่าความผิดพลาดนี้จะถูกคำนวณย้อนกลับ (backpropagate) ไปที่เลเยอร์แรก ซึ่งในแต่ละเลเยอร์จะมีการคำนวณย้อนกลับแบบรีเคอร์ซีฟเพื่อคำนวณในแต่ละ time step ด้วยซึ่งเรียกว่าการคลี่ในเชิงของเวลา (Unrolling in Time) ซึ่งทำให้การคำนวณย้อนกลับของเน็ตเวิร์ก RNN ถูกเรียกว่า Backpropagation Through Time (BPTT) ซึ่งมีความแตกต่างจากเน็ตเวิร์กทั่วไป (fully connected layer, FCL) ที่ไม่มีการคำนวณย้อนกลับในลักษณะของ time-step



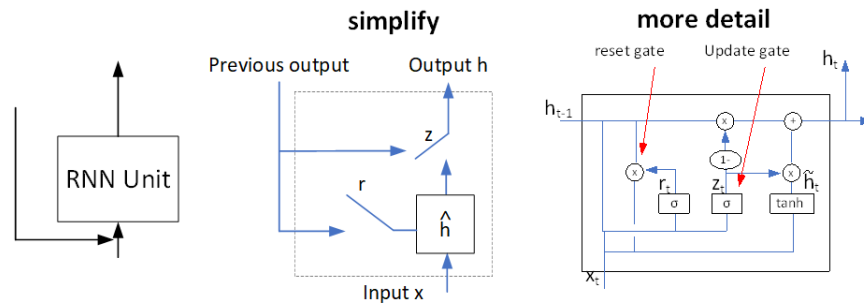
Backpropagation Through Time (BPTT)

รูปที่ 10 ขั้นตอนวิธีการแพร่ย้อนกลับของเน็ตเวิร์ก RNN

ที่มา : Fundamentals Of Deep Learning For NLP, อ.ดร.พีรพล เวทีกุล

2.2.1.5 สมการ Gated Recurrent Unit (GRU)

GRU เน็ตเวิร์กหรือที่เรียกว่า Gated Recurrent Neural Networks (GRNN) ถูกพัฒนาขึ้นเพื่อแก้ปัญหการจำระยะสั้นในเน็ตเวิร์กแบบ RNN ทำให้เน็ตเวิร์กแบบ GRU สามารถที่จะจดจำรูปแบบความสัมพันธ์ระหว่างตัวแปรอินพุตที่เข้ามาในเทอม (term) หรือ time-step ที่ยาวขึ้นได้ ซึ่งเซลล์ GRU ประกอบไปด้วย 2 เกต ได้แก่ reset gate (r) และ update gate (z) ดังรูปที่ 11



รูปที่ 11 โครงสร้างทางสถาปัตยกรรมของเซลล์ GRU เทียบกับเซลล์ RNN

ที่มา : วิชา Pattern Recognition L10 CNN RNN, อ.ดร.เอกพล ช่างสุวนิช

$$z_t^j = \text{sigmoid}^j(W_z x_t + U_z h_{t-1}) \quad (5)$$

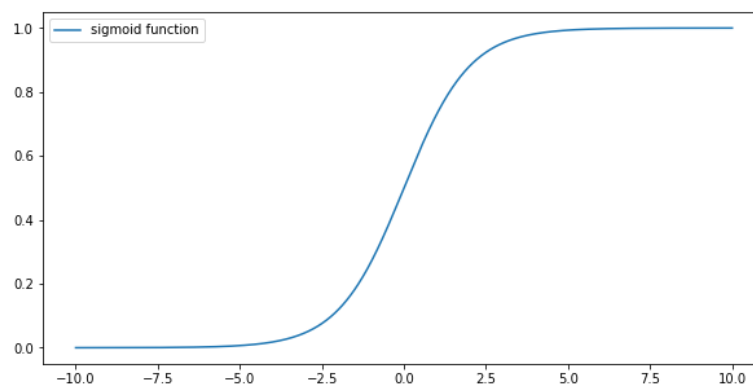
$$r_t^j = \text{sigmoid}(W_r x_t + U_r h_{t-1}) \quad (6)$$

$$\hat{h}_t^j = \tanh^j(W x_t + U(r_t \odot h_{t-1})) \quad (7)$$

$$h_t^j = (1 - z_t^j)h_{t-1}^j + z_t^j \hat{h}_t^j \quad (8)$$

ซึ่งสมการของเน็ตเวิร์กแบบ GRU แสดงตาม (5) - (8)

สมการ (5) คือการคำนวณอัปเดตเกต (z_t^j) โดยที่ j คือดัชนีของนิวรอน (neuron index) และ t ดัชนีของเวลา (time index) ซึ่ง (z_t^j) คำนวณจากการรวมระหว่าง อินพุตที่ timestep t (x_t) ที่ซึ่งคุณอยู่กับเมทริกซ์ค่าน้ำหนักเอาต์พุตของอัปเดตเกต (W_z) กับ เอาต์พุตของสถานะก่อนหน้า (h_{t-1}) ที่ซึ่งคุณอยู่กับเมทริกซ์ค่าน้ำหนักสถานะของอัปเดตเกต (U_z) หลังจากนั้นผ่านผลรวมนี้เข้าไปยังฟังก์ชันไม่เชิงเส้น (nonlinearity function) ที่ซึ่งคือฟังก์ชัน sigmoid ดังแสดงในรูปที่ 12 และมีสมการดัง (9) ฟังก์ชันนี้จะคืนค่าอยู่ในช่วงจำนวนจริงระหว่าง 0 ถึง 1 ผลลัพธ์ที่ได้ก็คือ z_t^j จะเป็นจำนวนจริงค่าหนึ่ง (scalar) คำนวณเช่นนี้ทุกนิวรอน

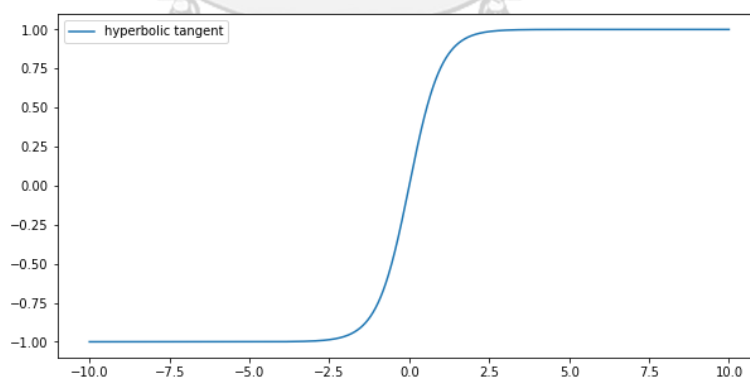


รูปที่ 12 ฟังก์ชันซิกมอยด์ (sigmoid)

$$\text{Sigmoid } \sigma(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

สมการ (6) คือการคำนวณรีเซตเกต (r_t^j) จากการรวมระหว่าง อินพุตที่ timestep t (x_t) ที่ซึ่งคุณอยู่กับเมทริกซ์ค่าน้ำหนักเอาต์พุตของรีเซตเกต (W_r) กับ เอาต์พุตของสถานะก่อนหน้า (h_{t-1}) ที่ซึ่งคุณอยู่กับเมทริกซ์ค่าน้ำหนักสถานะของรีเซตเกต (U_r) หลังจากนั้นผ่านผลรวมนี้เข้าไปยังฟังก์ชัน sigmoid ก็จะได้ r_t^j ออกมา

สมการ (7) คือการคำนวณเอาต์พุตที่ทำนายได้ที่ time-step ปัจจุบัน (\hat{h}_t^j) ซึ่งคำนวณจากผลรวมระหว่าง อินพุตที่ timestep t (x_t) ที่ซึ่งคุณอยู่กับเมทริกซ์ค่าน้ำหนักเอาต์พุต (W) กับ การทำ Element-wise product (\odot) ระหว่างรีเซตเกต (r_t) กับเอาต์พุตของสถานะก่อนหน้า (h_{t-1}) เพื่อตัดสินใจว่าจะเอา h_{t-1} มาเท่าใด ซึ่งผลลัพธ์จากการทำ Element-wise product จะคุณอยู่กับเมทริกซ์ค่าน้ำหนักสถานะของฮิปเตตเกต (U) หลังจากนั้นผ่านผลรวม ($Wx_t + U(r_t \odot h_{t-1})$) ไปยังฟังก์ชันไม่เชิงเส้น (nonlinearity function) ที่ซึ่งคือฟังก์ชัน tanh ดังแสดงในรูปที่ 13 และมีสมการดัง (10) เนื่องจากฟังก์ชันนี้จะคืนค่าอยู่ในช่วงจำนวนจริงระหว่าง -1 ถึง 1 ซึ่งจะช่วยป้องกันการเกิดปัญหา explosion gradient problem เนื่องจากฟังก์ชันนี้จะจำกัด (Bounds) ค่าให้อยู่ในช่วง -1 ถึง 1 ทำให้พอคำนวณ backpropagate แล้วค่าไม่ระเบิดเป็นค่าอนันต์ (infinity) ผลลัพธ์ที่ได้ก็คือ \hat{h}_t^j จะเป็นจำนวนจริงค่าหนึ่ง (scalar) ที่อยู่ในช่วง -1 ถึง 1



รูปที่ 13 ฟังก์ชัน tanh

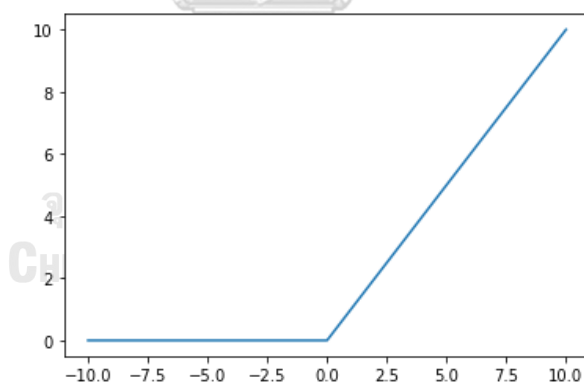
$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (10)$$

สมการ (8) คือการคำนวณเอาต์พุตที่ time-step ปัจจุบัน (h_t^j) ซึ่งคำนวณได้จากการประมาณค่าในช่วง (interpolation) ด้วยอ็อปเตดเกต (z_t^j) ระหว่างเทอม h_{t-1}^j และเทอม \hat{h}_t^j กล่าวอีกในหนึ่งก็คือเป็นการตัดสินใจว่าจะอ็อปเตดระหว่างเอาต์พุตของสถานะก่อนหน้า (h_{t-1}^j) และเอาต์พุตที่ทำนายได้ที่ time-step ปัจจุบัน (\hat{h}_t^j) ในสัดส่วนเท่าใด

ฟังก์ชัน Rectified Linear Unit (ReLU) มีสมการตาม (11) และแสดงดังรูปที่ 14

$$R(z) = \begin{cases} z, & z > 0 \\ 0, & z \leq 0 \end{cases} \quad (11)$$

จากรูปที่ 14 จะเห็นได้ว่าฟังก์ชัน ReLU จะส่งกลับค่าจากฟังก์ชัน $\text{Max}(x, 0)$ ยกตัวอย่างเช่น ถ้าค่าที่เข้ามามีค่าน้อยกว่า 0 จะคืนค่า 0 แต่ถ้ามากกว่า 0 จะคืนค่า x ที่เข้ามา ข้อดีของฟังก์ชันนี้คือมีค่าอนุพันธ์หรือความชันเป็นหนึ่งทำให้เกรเดียนต์ไม่หายหรือกล่าวอีกนัยหนึ่งก็คือไม่ถูกลดขนาดลงซึ่งช่วยลดปัญหา Vanishing Gradient Problem แต่อย่างไรก็ตามค่าอนุพันธ์ (Derivative) อีกค่าหนึ่งของมันเป็นศูนย์ดังนั้นมันก็มีโอกาสที่จะโยนเกรเดียนต์ (Gradient) นั้นทิ้งไปหมดได้ ยิ่งไปกว่านั้นฟังก์ชันนี้ยังช่วยให้เทรนโมเดลได้เร็วและลู่เข้า (Converge) เร็วขึ้น

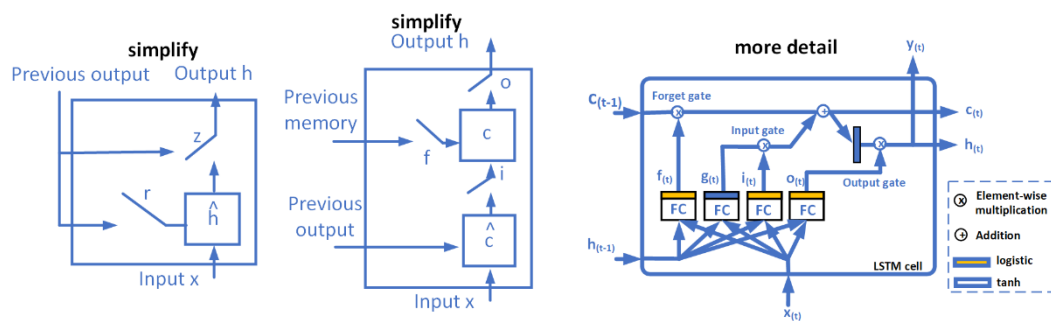


รูปที่ 14 ฟังก์ชันเส้นตรงที่ถูกปรับแก้ Rectified (Rectified Linear Unit ; ReLU)

2.2.1.6 สมการ Long Short-Term Memory (LSTM)

เน็ตเวิร์ก LSTM หรือที่เรียกว่า Long Short-Term Memory (LSTM) ถูกพัฒนาขึ้นเพื่อแก้ปัญหาการจำระยะสั้นในเน็ตเวิร์กแบบ RNN ทำให้เน็ตเวิร์กแบบ LSTM สามารถที่จะจดจำรูปแบบความสัมพันธ์ระหว่างตัวแปรอินพุตที่เข้ามาในเทอม (term) หรือ time-step ที่ยาวขึ้นได้ ซึ่ง LSTM ประกอบไปด้วย 3 เกต ได้แก่

1. อินพุตเกต (input gate, i) คือเกตที่ทำหน้าที่ตัดสินใจว่าจะอินพุตข้อมูลระหว่าง อินพุตที่สถานะปัจจุบัน (x_t) กับ เอาต์พุตของสถานะก่อนหน้า (h_{t-1}) ในสัดส่วนเท่าไร
2. เอาต์พุตเกต (Output gate, o) คือเกตที่ทำหน้าที่ตัดสินใจว่าจะเอาต์พุตข้อมูลข้อมูลในหน่วยความจำ (c_t^j) ออกไปมากเท่าใด
3. เกตลืม (forget gate, f) คือเกตที่ทำหน้าที่ตัดสินใจว่าจะลืมข้อมูลระหว่างอินพุตที่สถานะปัจจุบัน (x_t) กับ เอาต์พุตของสถานะก่อนหน้า (h_{t-1}) ในสัดส่วนเท่าไร และ 1 หน่วยความจำ (memory cell) ดังแสดงใน รูปที่ 15



รูปที่ 15 โครงสร้างของเซลล์ LSTM เทียบกับ GRU

ที่มา : วิชา Pattern Recognition L10 CNN RNN, อ.ดร.เอกพล ช่างสุวนิช

เน็ตเวิร์ก LSTM มีสมการตาม (12) – (17) ซึ่งสามารถอธิบายได้ดังนี้

$$i_t^j = F^j(W_i x_t + U_i h_{t-1} + V_i c_{t-1}) \quad (12)$$

$$o_t^j = F^j(W_o x_t + U_o h_{t-1} + V_o c_t) \quad (13)$$

$$f_t^j = F^j(W_f x_t + U_f h_{t-1} + V_f c_{t-1}) \quad (14)$$

$$\hat{c}_t^j = \tanh^j(W_c x_t + U_c h_{t-1}) \quad (15)$$

$$c_t^j = f_t^j c_{t-1}^j + i_t^j \hat{c}_t^j \quad (16)$$

$$h_t^j = o_t^j \tanh(c_t^j) \quad (17)$$

สมการ (12) คือการคำนวณอินพุตเกต (i_t^j) ซึ่งคำนวณได้จากผลรวมของ 3 เทอม ได้แก่

1. เทอม $W_i x_t$ ซึ่งเกิดจากการคูณกันระหว่างอินพุตที่สถานะปัจจุบัน (x_t) กับ เมทริกซ์ค่าน้ำหนักเอาต์พุตของอินพุตเกต (W_i)
2. เทอม $U_i h_{t-1}$ ซึ่งเกิดจากการคูณกันระหว่างเอาต์พุตของสถานะก่อนหน้า (h_{t-1}) กับ เมทริกซ์ค่าน้ำหนักสถานะของอินพุตเกต (U_i)

3. เทอม $V_i c_{t-1}$ ซึ่งเกิดจากการคูณกันระหว่างข้อมูลที่เก็บอยู่ในหน่วยความจำของ time step ก่อนหน้านี้นี้ (c_{t-1}) กับ เมทริกซ์ค่าน้ำหนักอินพุตของอินพุตเกต (V_i)

หลังจากนั้นผ่านผลรวมของทั้ง 3 เทอมนี้เข้าไปยังฟังก์ชันไม่เชิงเส้น (nonlinearity function) ซึ่งคือ ฟังก์ชัน sigmoid ก็จะได้ผลลัพธ์เป็นอินพุตเกต (i_t^j) ซึ่งเป็นค่าสเกลาร์ (scalar) ค่าหนึ่ง

สมการ (13) คือการคำนวณเอาต์พุตเกต (o_t^j) ซึ่งคำนวณได้จากผลรวมของ 3 เทอม ได้แก่

1. เทอม $W_o x_t$ ซึ่งเกิดจากการคูณกันระหว่างอินพุตที่สถานะปัจจุบัน (x_t) กับ เมทริกซ์ค่าน้ำหนักเอาต์พุตของเอาต์พุตเกต (W_o)
2. เทอม $U_o h_{t-1}$ ซึ่งเกิดจากการคูณกันระหว่างเอาต์พุตของสถานะก่อนหน้านี้นี้ (h_{t-1}) กับ เมทริกซ์ค่าน้ำหนักสถานะของเอาต์พุตเกต (U_o)
3. เทอม $V_o c_t$ ซึ่งเกิดจากการคูณกันระหว่างข้อมูลที่เก็บอยู่ในหน่วยความจำของ time step ก่อนหน้านี้นี้ (c_{t-1}) กับ เมทริกซ์ค่าน้ำหนักอินพุตของเอาต์พุตเกต (V_o)

หลังจากนั้นผ่านผลรวมของทั้ง 3 เทอมนี้เข้าไปยังฟังก์ชันไม่เชิงเส้น (nonlinearity function) ซึ่งคือ ฟังก์ชัน sigmoid ก็จะได้ผลลัพธ์เป็นเอาต์พุตเกต (o_t^j) ซึ่งเป็นค่าสเกลาร์ (scalar) ค่าหนึ่ง

สมการ (14) คือการคำนวณฟอร์เกทเกต (f_t^j) ซึ่งคำนวณได้จากผลรวมของ 3 เทอม ได้แก่

1. เทอม $W_f x_t$ ซึ่งเกิดจากการคูณกันระหว่างอินพุตที่สถานะปัจจุบัน (x_t) กับ เมทริกซ์ค่าน้ำหนักเอาต์พุตของฟอร์เกทเกต (W_f)
2. เทอม $U_f h_{t-1}$ ซึ่งเกิดจากการคูณกันระหว่างเอาต์พุตของสถานะก่อนหน้านี้นี้ (h_{t-1}) กับ เมทริกซ์ค่าน้ำหนักสถานะของฟอร์เกทเกต (U_f)
3. เทอม $V_f c_{t-1}$ ซึ่งเกิดจากการคูณกันระหว่างข้อมูลที่เก็บอยู่ในหน่วยความจำของ time step ก่อนหน้านี้นี้ (c_{t-1}) กับ เมทริกซ์ค่าน้ำหนักอินพุตของฟอร์เกทเกต (V_f)

หลังจากนั้นผ่านผลรวมของทั้ง 3 เทอมนี้เข้าไปยังฟังก์ชันไม่เชิงเส้น (nonlinearity function) ซึ่งคือ ฟังก์ชัน sigmoid ก็จะได้ผลลัพธ์เป็นฟอร์เกทเกต (f_t^j) ซึ่งเป็นค่าสเกลาร์ (scalar) ค่าหนึ่ง

สมการ (15) คือการคำนวณข้อมูลที่เก็บในหน่วยความจำที่คำนวณได้ที่สถานะปัจจุบัน (\hat{c}_t^j) ซึ่งเป็นผลรวมของ 2 เทอม ได้แก่

1. เทอม $W_c x_t$ เป็นการคูณกันระหว่างอินพุตที่สถานะปัจจุบัน (x_t) กับ เมทริกซ์ค่าน้ำหนักเอาต์พุตของหน่วยความจำ (W_c)

2. เทอม $U_c h_{t-1}$ เป็นการคูณกันระหว่างเอาต์พุตของสถานะก่อนหน้า (h_{t-1}) กับ เมทริกซ์ค่าน้ำหนักสถานะของหน่วยความจำ (U_c)

หลังจากนั้นผ่านผลรวมของทั้ง 2 เทอมนี้เข้าไปยังฟังก์ชันไม่เชิงเส้น (nonlinearity function) ซึ่งคือฟังก์ชัน tanh ก็จะได้ผลลัพธ์เป็นข้อมูลที่เก็บในหน่วยความจำที่คำนวณได้ที่สถานะปัจจุบัน (\hat{c}_t^j) ซึ่งเป็นค่าสเกลาร์ (scalar) ค่าหนึ่ง

สมการ (16) คือการคำนวณข้อมูลที่เก็บในหน่วยความจำที่สถานะปัจจุบัน (c_t^j) ซึ่งเป็นการรวมกันของ 2 เทอม ได้แก่

1. เทอม $f_t^j c_{t-1}^j$ คือการตัดสินใจว่าจะเอาข้อมูลที่เก็บอยู่ในหน่วยความจำใน time-step ก่อนหน้านี้นี้มากน้อยแค่ไหน หรือพูดได้อีกอย่างหนึ่งก็คือการตัดสินใจว่าจะลืมข้อมูลที่เก็บอยู่ในหน่วยความจำใน time-step ก่อนหน้านี้นี้มากขนาดไหน
2. เทอม $i_t^j \hat{c}_t^j$ เป็นการตัดสินใจว่าจะอินพุตข้อมูลที่เก็บในหน่วยความจำที่คำนวณได้ที่สถานะปัจจุบัน (\hat{c}_t^j) มากขนาดไหน

สมการ (17) คือการคำนวณเอาต์พุตที่สถานะปัจจุบัน (h_t^j) ซึ่งเกิดจากการผ่านข้อมูลที่เก็บในหน่วยความจำที่สถานะปัจจุบัน (c_t^j) ไปยังฟังก์ชัน tanh และตัดสินใจว่าจะเอาต์พุตมันออกไปมากน้อยแค่ไหน

2.2.1.7 การต่อ GRU เลเยอร์ (Gated Recurrent Unit (GRU) layer)

การต่อ GRU เลเยอร์จะมีการนำเอาต์พุตของสถานะก่อนหน้ามาเป็นอินพุตให้กับสถานะปัจจุบันจากรูปที่ 16 ที่สถานะเริ่มต้นซึ่งในตัวอย่างนี้คือ time-step t-1 จะมีการ initialize ค่าเอาต์พุตของสถานะก่อนหน้าขึ้นเนื่องจากเป็นสถานะเริ่มต้นจึงยังไม่มีสถานะก่อนหน้า

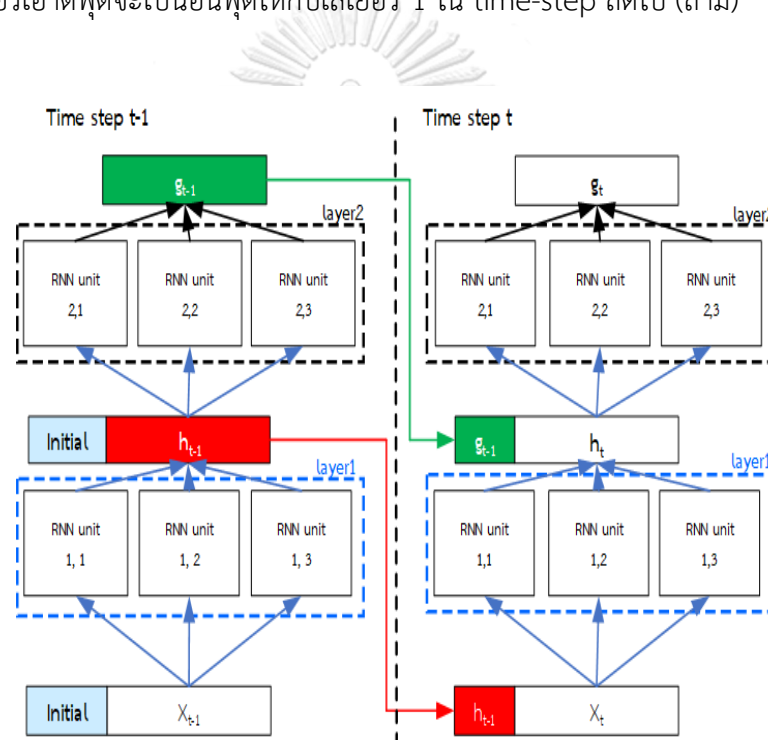
พิจารณา time-step t-1 ซึ่งจะแบ่งออกเป็น 4 เลเยอร์ด้วยกันคือ

1. เลเยอร์อินพุต จะประกอบไปด้วย 2 ส่วนด้วยกัน ได้แก่ เอาต์พุตของสถานะก่อนหน้า (initialize) และ อินพุตที่ time-step นี้ซึ่งคือ (x_{t-1})
2. เลเยอร์ 1 เป็นเลเยอร์ของเซลล์ GRU ซึ่งจะให้อาต์พุต (h_{t-1}) ซึ่งเป็นผลรวมของแต่ละเซลล์ GRU ในเลเยอร์นี้ และมีการ initialize ค่าเอาต์พุต (g_{t-1}) ของ time-step ก่อนหน้านี้ด้วย
3. เลเยอร์ 2 เป็นเลเยอร์ของเซลล์ GRU ซึ่งจะให้อาต์พุต (g_{t-1}) ซึ่งเป็นผลรวมของแต่ละเซลล์ GRU ในเลเยอร์นี้

4. เลเยอร์เอาต์พุต จะเป็นอินพุตให้กับเลเยอร์ 1 ใน time-step ถัดไป

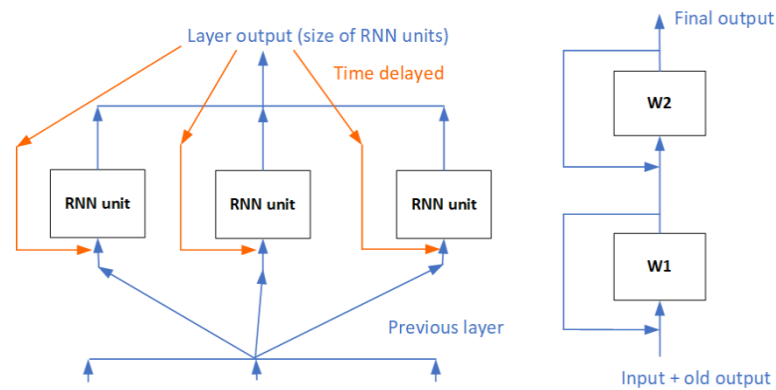
พิจารณา time-step t ซึ่งจะแบ่งออกเป็น 4 เลเยอร์ เช่นเดียวกัน ได้แก่

1. เลเยอร์อินพุต ประกอบไปด้วยเอาต์พุตของเลเยอร์ 2 ของ time step $t-1$ (g_{t-1}) และอินพุตที่ timestep ปัจจุบัน (x_t)
2. เลเยอร์ 1 เป็นเลเยอร์ของเซลล์ GRU
3. เลเยอร์ 2 อินพุตของเลเยอร์นี้ประกอบไปด้วยเอาต์พุตของเลเยอร์ที่ 2 ของ time step ก่อนหน้านี้ (g_{t-1}) และเอาต์พุตของเลเยอร์ 1 (h_t)
4. เลเยอร์เอาต์พุตจะเป็นอินพุตให้กับเลเยอร์ 1 ใน time-step ถัดไป (ถ้ามี)



รูปที่ 16 การต่อ GRU layer (แบบละเอียด Unrolling)

ที่มา : วิชา Pattern Recognition L10 CNN RNN, อ.ดร.เอกพล ช่างสุวนิช



รูปที่ 17 การต่อ GRU layer (แบบย่อ)

ที่มา : วิชา Pattern Recognition L10 CNN RNN, อ.ดร.เอกพล ช่วงสุวนิช

2.2.1.8 ข้อดีข้อเสียของเน็ตเวิร์กแบบวนกลับ

เน็ตเวิร์กแบบวนกลับแต่ละประเภทมีข้อดีและข้อเสียแตกต่างกันดังตารางที่ 1

ตารางที่ 1 ข้อดีและข้อเสียของเน็ตเวิร์กแบบวนกลับ

	RNN	GRU	LSTM
ข้อดี	สำหรับข้อมูลที่เป็นลำดับเวลา RNN ให้ประสิทธิภาพที่ดีกว่า DNN	GRU มีโครงสร้างซับซ้อนน้อยกว่า LSTM - พารามิเตอร์น้อยกว่าทำให้ฝึกสอนได้เร็วกว่าและมีประสิทธิภาพเทียบเท่า LSTM	สำหรับข้อมูลที่มีความสัมพันธ์ของสถานะยาว, LSTM ให้ประสิทธิภาพที่ดีกว่า GRU เนื่องจากมีหน่วยความจำภายในทำให้สามารถจำความสัมพันธ์ของสถานะที่ยาวได้
ข้อเสีย	RNN ไม่สามารถลืมนสถานะก่อนหน้าได้	สำหรับข้อมูลที่มีความสัมพันธ์ของสถานะยาว GRU จะให้ประสิทธิภาพที่แย่กว่า LSTM	ฝึกสอนแบบจำลองได้ช้าเนื่องจากมี 3 เกต 1 หน่วยความจำภายใน

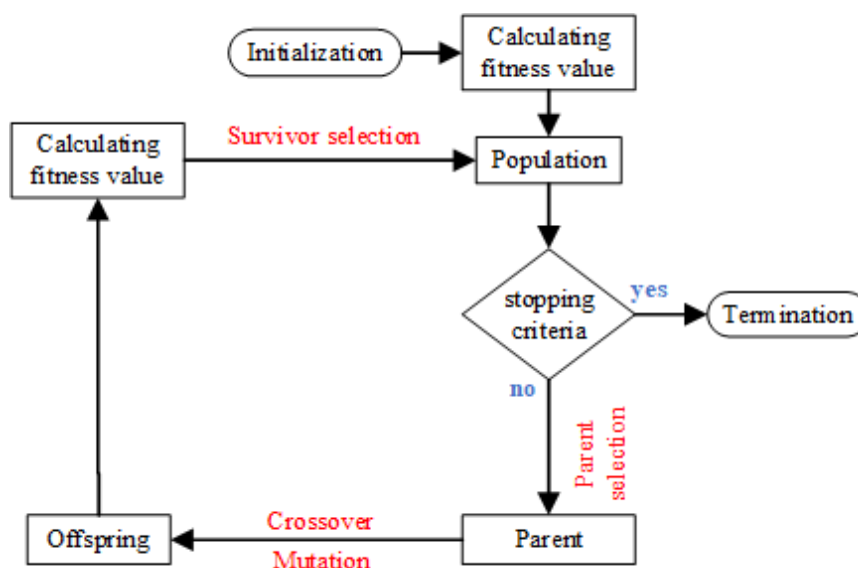
2.2.2 ขั้นตอนวิธีเชิงวิวัฒนาการ (Evolutionary Algorithm)

ขั้นตอนวิธีเชิงวิวัฒนาการเป็นระบบการแก้ปัญหาที่จำลองการวิวัฒนาการที่เกิดขึ้นในธรรมชาติ เช่น การสืบพันธุ์ การสลับสายพันธุ์ การกลายพันธุ์ และการคัดเลือกซึ่งเป็นการหาค่าเหมาะสมที่สุดแบบฮิวริสติก (heuristic) โดยขั้นตอนวิธีเชิงวิวัฒนาการนี้ถูกพัฒนาอย่างต่อเนื่องดังแสดงในตารางที่ 2 ซึ่งขั้นตอนวิธีเชิงวิวัฒนาการส่วนใหญ่จะมีรูปแบบขั้นตอนที่คล้ายกัน [26]

ตารางที่ 2 วิวัฒนาการของขั้นตอนวิธีเชิงวิวัฒนาการ

ปี ค.ศ	นักวิจัย	การนำเสนอ
1859	Charles Darwin	ทฤษฎีวิวัฒนาการ
1965, 1966	Fogel et al.	การโปรแกรมเชิงวิวัฒนาการ (Evolutionary programming) ซึ่งทำงานร่วมกับเครื่องจักรสถานะแบบจำกัด (Finite state machine)
1973 1975	Holland	ขั้นตอนวิธีเชิงพันธุกรรม (Genetic algorithm: GA) ซึ่งทำงานร่วมกับสายบิต (กลายเป็นที่รู้จักในวงกว้างจาก หนังสือของ Goldberg, 1989)
1965, 1971, 1974	Rechenberg, Schwefel	กลยุทธ์เชิงวิวัฒนาการ (Evolution strategies: ES) ที่ทำงานร่วมกับเวกเตอร์ของจำนวนจริง เพื่อหาค่าเหมาะสมที่สุดเชิงตัวเลข (Numerical optimization)
1990	Koza	การโปรแกรมเชิงพันธุกรรม (Genetic programming: GP) ซึ่งทำงานร่วมกับโครงสร้างต้นไม้
2001	O'Neill และ Ryan	เสนอวิวัฒนาการทางไวยากรณ์ (Grammatical evolution: GE) ซึ่งทำงานร่วมกับสายบิตและกระบวนการเชื่อมโยงด้วยไวยากรณ์ รูปแบบแบ็กคัส-เนอาร์ (Backus-Naur form: BNF)
2007	Eiben และ Smith	ขั้นตอนวิธีเชิงวิวัฒนาการ

จากรูปที่ 18 แสดงขั้นตอนการทำงานของขั้นตอนวิธีเชิงวิวัฒนาการ (Eiben และ Smith, 2007) เริ่มต้นด้วยการสร้างประชากร (population) รุ่นแรกขึ้นมาด้วยการสุ่มจากคำตอบของปัญหา หลังจากนั้นหาค่าของฟังก์ชันความเหมาะสม (fitness function) เพื่อประเมินคุณภาพของประชากรแต่ละตัว ต่อมาทำการตรวจสอบเงื่อนไขของการหยุดว่าเป็นจริงหรือไม่ ถ้าเป็นจริงให้หยุดการทำงาน ถ้าไม่จริงจะทำการคัดเลือกประชากรพ่อแม่ (parent selection) เพื่อสร้างประชากรลูก (offspring) ด้วยตัวดำเนินการแปรผัน (variation operators) ได้แก่ การสลับสายพันธุ (crossover) ซึ่งคือการผสมคู่ประชากรเข้าด้วยกันเพื่อให้ได้ผู้สืบทอด และการกลายพันธุ์ (mutation) ซึ่งคือการกลายพันธุ์ผู้สืบทอดถูกสร้างขึ้นจากขั้นตอนที่แล้ว หลังจากนั้นผู้สืบทอดที่ถูกสร้างทั้งหมดจะถูกคัดเลือกเพื่อความอยู่รอด (survival selection) สำหรับเป็นประชากรรุ่นต่อไป โดยกระบวนการวิวัฒนาการจะดำเนินไปอย่างต่อเนื่องเพื่อปรับปรุงคุณภาพของประชากรผ่านประชากรหลาย ๆ รุ่น จนกระทั่งเงื่อนไขการหยุดเป็นจริง



รูปที่ 18 แผนภาพการทำงานของขั้นตอนวิธีเชิงวิวัฒนาการ

ทั้งนี้ขั้นตอนวิธีเชิงวิวัฒนาการก็มีหลายอัลกอริทึมซึ่งวิธีการที่รู้จักกันมากคือ ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm)

2.2.2.1 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm; GA)

ขั้นตอนวิธีเชิงพันธุกรรมถูกนำเสนอครั้งแรกโดยคุณฮอลแลนด์ [27] ซึ่งคำตอบที่เป็นไปได้ของปัญหาต่างๆจะถูกแทนด้วยโครโมโซม (Chromosome) ด้วยการเข้ารหัสโครโมโซมซึ่งประชากรหนึ่งตัวอาจมีมากกว่าหนึ่งโครโมโซมได้และประชากรหลายๆตัวรวมกันเป็นกลุ่มประชากร (Population) และกลุ่มประชากรนี้จะมีการวิวัฒนาการอย่างต่อเนื่องซึ่งแต่ละรอบของการวิวัฒนาการจะเรียกว่ารุ่นประชากร (Generation) ซึ่งจะค่อยๆลู่เข้าหาคำตอบ

ขั้นตอนวิธีเชิงพันธุกรรมเป็นวิธีการแก้ไขปัญหาโดยใช้วิธีการหาค่าเหมาะที่สุดแบบเฟ้นสุ่ม (Stochastic optimization) ด้วยการค้นหาข้อมูลแบบฮิวริสติก (Heuristic Search) ซึ่งใช้ข้อมูลความรู้แล้วเพื่อช่วยในการหาคำตอบซึ่งถูกใช้อย่างกว้างขวางเพื่อหาคำตอบที่เกือบจะเหมาะสมที่สุด (near optimal solution) ในปริภูมิการค้นหา (parameters space) ขนาดใหญ่ [28]

ขั้นตอนวิธีเชิงพันธุกรรมประกอบไปด้วย 4 ขั้นตอน ได้แก่

1. การสุ่มประชากรเริ่มต้น (Initialization) ตามขนาดประชากร (Population size) ที่กำหนด

2. การคำนวณค่าความเหมาะสม (Fitness value) ของประชากรแต่ละตัวจากฟังก์ชันค่าความเหมาะสม (Fitness function) ซึ่งเป็นฟังก์ชันที่มีความแตกต่างกันตามชนิดของปัญหา

3. การคัดเลือกประชากร (Selection)

การคัดเลือกประชากร คือการเปรียบเทียบประชากรในแต่ละรุ่นแล้วคัดเลือกประชากรที่ดีในแต่ละรุ่นเพื่อทำหน้าที่เป็นพ่อแม่ในการให้กำเนิดประชากรรุ่นถัดไป ขณะเดียวกันก็กำจัดประชากรที่ไม่มีประสิทธิภาพทิ้งไป

3.1 Tournament Selection มีประสิทธิภาพดีกว่าการคัดเลือกแบบวงล้อรูเล็ต (Roulette wheel Selection) เนื่องจากกลุ่มเข้าหาคำตอบได้เร็วกว่า [29]

4. การดำเนินการทางพันธุกรรม คือการสร้างประชากรรุ่นใหม่ที่มีความหลากหลายมากขึ้น ซึ่งแต่ละครั้งที่ประชากรรุ่นใหม่ถูกสร้างจะเรียกว่า “รุ่น” (Generation) การสร้างประชากรด้วยวิธีการต่างๆจะช่วยป้องกันไม่ให้เกิดการค้นหาติดอยู่ที่คำตอบเหมาะสมท้องถิ่น (Local Optimal Solution)

4.1 การสลับสายพันธุ์ (Crossover)

คือการสร้างประชากรขึ้นมาใหม่จากประชากรที่มีอยู่ (การสร้างลูกจากการผสมของพ่อแม่) เพื่อให้เกิดความหลากหลายของประชากร เมื่อมีความหลากหลายของประชากรก็จะมีโอกาสที่จะได้ประชากรที่ดี ซึ่งมีด้วยกันหลายวิธี เช่น วิธีการสลับสายพันธุ์แบบจุดเดียว (One Point Crossover)

4.2 การกลายพันธุ์ (Mutation)

การกลายพันธุ์คือการเพิ่มความหลากหลายของประชากรหลังจากผ่านกระบวนการสลับสายพันธุ์ซึ่งเป็นการสร้างประชากรขึ้นมาใหม่ด้วยวิธีการ เช่น Randomly Swapping, Turning-Off Solution Bits เป็นต้น ซึ่งทั้งสองวิธีเป็น binary mutation

กระบวนการวิวัฒนาการจะเลือกประชากรที่มีคุณภาพในแต่ละรุ่นเพื่ออยู่รอดในรุ่นต่อไปและวนซ้ำจนกระทั่งถึงเงื่อนไขการหยุด (Stopping Criterion) ที่กำหนดไว้

2.2.2.2 การเข้ารหัสโครโมโซม

การเข้ารหัสโครโมโซมคือกระบวนการที่ทำหน้าที่สร้างตัวแทนของคำตอบหรือพารามิเตอร์ที่ใช้ในการค้นหา (search) มีหลากหลายรูปแบบของการแทนโครโมโซม (Representations) ดังต่อไปนี้

2.2.2.3 การแทนแบบเลขฐานสอง

การแทนแบบเลขฐานสอง (Binary Representation) แต่ละโครโมโซมจะถูกแทนด้วยสายบิตที่แต่ละตำแหน่งมีค่าไม่ 0 ก็ 1 เท่านั้น เช่น 10110101001 เป็นต้น วิธีในการกลายพันธุ์ยีนในโครโมโซมจะใช้วิธีการกลับบิต (0->1 หรือ 1->0) ซึ่งแต่ละบิตอิสระต่อกันในการกลับบิต สำหรับการรวมตัวแบ่งได้ 3 แบบ คือ

1. แบบจุดข้ามเดี่ยว คือการกำหนดจุดข้ามขึ้นมาหนึ่งจุดแล้วทำการสลับสายของบิต เช่น 00111|10 × 01101|00 -> 0011100,0110110 เป็นต้น
2. แบบจุดข้ามมากกว่าหนึ่ง คือการกำหนดจุดข้ามขึ้นมาตามจำนวนที่ต้องการ
3. แบบจุดข้ามสม่ำเสมอ คือการสุ่มสลับยีนกันของคู่ผสมในแต่ละตำแหน่งของยีนแยกกัน

2.2.2.4 การแทนแบบเลขจำนวนเต็ม

การแทนแบบเลขจำนวนเต็ม (Integer Representation) แต่ละโครโมโซมจะถูกแทนด้วยสายของเลขจำนวนเต็มในช่วงที่กำหนดวิธีที่ใช้ในการกลายพันธุ์ยีนในโครโมโซมมีอยู่ด้วยกัน 2 แบบ

1. การสุ่มใหม่ คือเมื่อตัวเลขในตำแหน่งใดๆของยีนถูกกำหนดให้กลายพันธุ์ตำแหน่งนั้นจะต้องทำการสุ่มตัวเลขขึ้นมาใหม่จากช่วงที่กำหนด
2. การค่อย ๆ เปลี่ยน คือเมื่อตัวเลขในตำแหน่งใดๆของยีนถูกกำหนดให้กลายพันธุ์ตัวเลขในตำแหน่งนั้นๆจะต้องถูกบวกเข้าไปด้วยค่าที่สุ่มขึ้นมาด้วยการแจกแจงที่สมมาตรรอบ 0 และมีค่าน้อย

สำหรับการรวมตัวของการแทนค่าแบบนี้สามารถทำได้ด้วยวิธีการเดียวกับแทนค่าแบบเลขฐานสอง

2.2.2.5 การแทนแบบเลขจำนวนจริง

การแทนแบบเลขจำนวนจริง (Real number representation) แต่ละโครโมโซมจะถูกแทนด้วยสายของเลขจำนวนจริงในช่วงที่กำหนด วิธีที่ใช้ในการกลายพันธุ์คล้ายกับวิธีการแทนค่าแบบเลขจำนวนเต็ม

1. การกลายพันธุ์อย่างสม่ำเสมอ คือค่าในตำแหน่งที่ถูกกำหนดให้กลายพันธุ์จะถูกสุ่มใหม่จากช่วงที่กำหนด
2. การกลายพันธุ์อย่างไม่สม่ำเสมอด้วยการแจกแจงคองที่ คือการบวกค่าที่สุ่มจากการแจกแจงที่กำหนดเข้าไปในตัวเลขตำแหน่งที่ถูกพิจารณาให้กลายพันธุ์ซึ่งนิยมใช้การแจกแจงแบบเกาส์เซียนและโคชี

สำหรับการรวมตัวของ การแทนค่าแบบนี้ทำได้เช่นเดียวกับแบบเลขฐานสองซึ่งเรียกว่า “การรวมแบบไม่ต่อเนื่อง” หรืออาจรวมเชิงคณิตศาสตร์แทนก็ได้ซึ่งมีลักษณะคล้ายกันเพียงแต่ในส่วนของยีนที่จะทำการสลับกันระหว่างคู่ผสมจะใช้ค่าระหว่างคู่แทนที่จะสลับกันตรง ๆ

2.2.2.6 การแทนแบบเรียงสับเปลี่ยน

การแทนแบบเรียงสับเปลี่ยน (Permutation representation) แต่ละโครโมโซมจะถูกแทนด้วยสายของอันดับซึ่งเป็นเลขจำนวนเต็มซึ่งจะพิจารณารูปแบบการเรียงมากกว่าค่าในแต่ละตำแหน่ง ตัวอย่างของปัญหาที่มักใช้การแทนแบบนี้ เช่น ปัญหาการเดินทางของพนักงานขาย (Traveling salesman problem) [30]

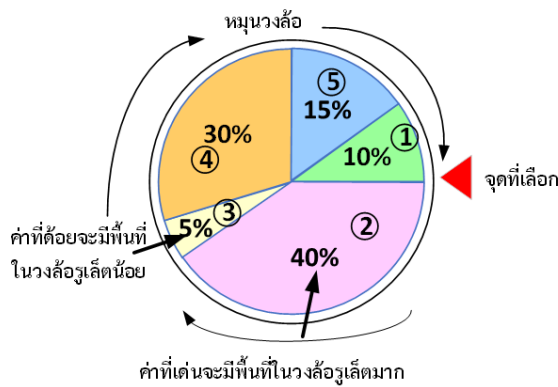
2.2.2.7 การคัดเลือกประชากรพ่อแม่ (Parent Selection)

จากการบวนการการเข้ารหัสโครโมโซมจะเห็นว่าการคัดเลือกโครโมโซมพ่อแม่เป็นกระบวนการที่สำคัญเพื่อให้ได้โครโมโซมลูกที่มีคุณภาพซึ่งวิธีการเลือกคู่ผสมมีดังนี้

1. Fitness Proportional
2. Selection Ranking
3. Selection Tournament Selection
4. Roulette wheel algorithm
5. Stochastic Universal Sampling

2.2.2.8 การคัดเลือกแบบวงล้อรูเล็ต (Roulette Wheel Selection)

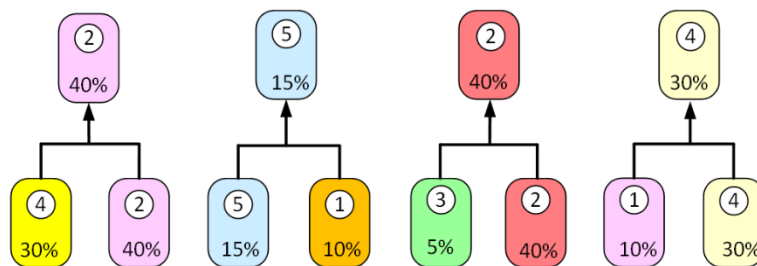
เป็นวิธีการที่เลียนแบบมาจากการเล่นรูเล็ตดังแสดงในรูปที่ 19 ซึ่งโครโมโซมที่มีค่าความเหมาะสม (fitness value) มากจะมีโอกาสถูกคัดเลือกมากกว่าโครโมโซมที่มีค่าความเหมาะสมน้อย โดยเริ่มจาก ความกว้างของช่องแต่ละช่องของวงล้อรูเล็ตได้จากค่าความเหมาะสมของสมาชิกแต่ละตัว หลังจากนั้นกำหนดตัวชี้ตำแหน่งจุดคงที่ (Fixed Point) ต่อมาหมุนวงล้อรูเล็ต เมื่อวงล้อหยุด หมุนจะเลือกโครโมโซมในตำแหน่งที่ถูกชี้ ทำเช่นนี้จนได้ประชากรครบในหนึ่งรุ่น แต่อย่างไรก็ตามวิธีนี้มีความลำเอียง (Bias) ค่อนข้างมากเนื่องจากโครโมโซมที่มีค่าความเหมาะสมที่สูงจะมีโอกาสถูกเลือกซ้ำหลายครั้งทำให้ประชากรในรุ่นถัดไปมีลักษณะของประชากรที่ถูกเลือกซ้ำๆ จำนวนมาก



รูปที่ 19 การคัดเลือกแบบวงล้อรูเล็กลง

2.2.2.9 การคัดเลือกแบบการแข่งขัน (Tournament Selection)

การคัดเลือกแบบการแข่งขันเป็นวิธีการคัดเลือกโครโมโซมพ่อแม่ที่ดีที่สุดวิธีนี้คล้ายกับการแข่งขันกีฬาที่แสดงในรูปที่ 20 กล่าวคือเริ่มต้นสุ่มแบ่งกลุ่มโครโมโซมหลังจากนั้นคัดเลือกโครโมโซมที่ดีในแต่ละกลุ่มเพื่อเป็นโครโมโซมพ่อแม่ในการให้กำเนิดประชากรรุ่นถัดไป ซึ่งขั้นตอนโดยละเอียดคือขั้นแรกทำการสุ่มเลือกโครโมโซมสำหรับจัดการแข่งขันซึ่งสามารถกำหนดค่า K ได้ (K คือ Tournament Size หรือจำนวนโครโมโซมที่ชนะการแข่งขันและถูกคัดเลือก) ยิ่งไปกว่านั้นเราสามารถเลือกได้ว่าเราจะเอาโครโมโซมที่ชนะออกจากการแข่งขันหรืออยู่ต่อได้ (เพื่อสามารถถูกเลือกได้อีก) การปรับขนาด K เหมือนปรับความเข้มข้นของการแข่งขันให้เหมาะสม เช่น ถ้า K มีขนาดใหญ่ (การแข่งขันขนาดใหญ่) แล้ว โครโมโซมที่ไม่ดีมาก (ค่าความเหมาะสมน้อย) ก็จะมีโอกาสชนะน้อย ในทางกลับกันถ้า K เล็ก โครโมโซมที่ไม่ดีมากก็ยังมีโอกาสชนะ จะเห็นได้ว่า การเลือกค่า K ที่เหมาะสมจะช่วยจัดการปัญหาความเหลื่อมล้ำของค่าความเหมาะสมของสมโครโมโซมให้หมดไป นอกจากนี้วิธียังช่วยลดปัญหาเรื่องความลำเอียง (Bias) ได้



รูปที่ 20 การคัดเลือกแบบการแข่งขัน

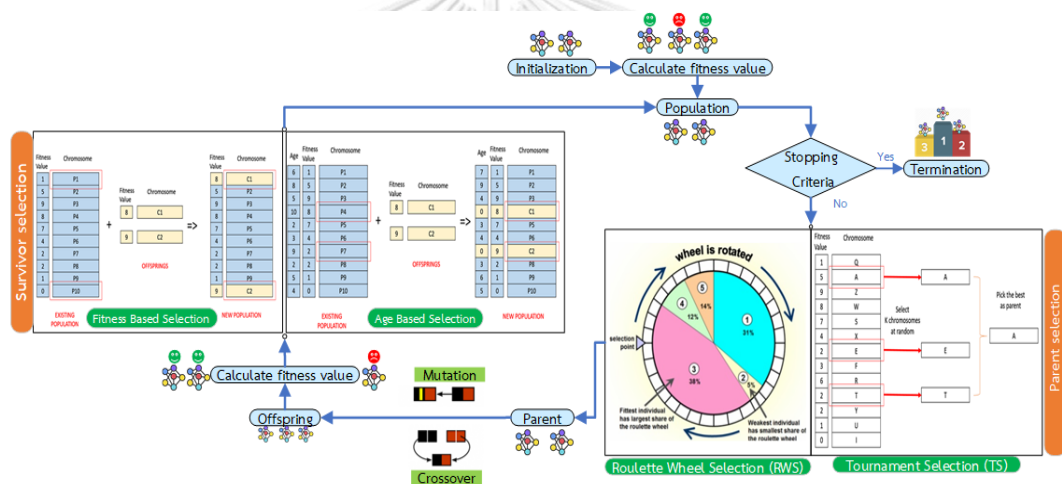
2.2.2.10 การคัดเลือกผู้รอด (Survivor Selection)

หลังจากผ่านกระบวนการการสลับสายพันธุ์และการกลายพันธุ์จะเกิดประชากรผู้สืบทอดกลุ่มใหม่ขึ้นดังนั้นจึงต้องมีการเลือกประชากรที่อยู่รอดและกำจัดประชากรส่วนเกินออก โดยวิธีการคัดเลือกผู้รอดมีดังนี้

1. Age-based replacement คือการคัดเลือกประชากรผู้รอดโดยการแทนที่ประชากรที่มีอายุมากด้วยประชากรลูกที่มีค่าความเหมาะสมสูง
2. Fitness-based replacement คือการคัดเลือกประชากรผู้รอดโดยการแทนที่ประชากรที่มีค่าความเหมาะสมน้อยด้วยประชากรลูกที่มีค่าความเหมาะสมสูง

2.1 Replace worst

2.2 Elitism



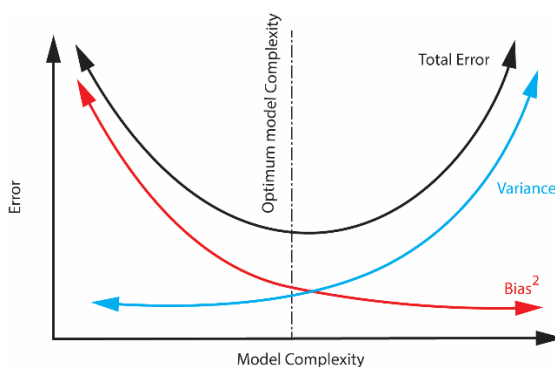
รูปที่ 21 กระบวนการคัดเลือกประชากรผู้รอด

ตารางที่ 3 ข้อดีและข้อเสียของวิธีการคัดเลือกประชากรพ่อแม่

Parent Selection	Selection Tournament	Roulette Wheel
ข้อดีและข้อเสีย	Low Bias (Good)	High Bias (Bad)
	High variance (Good)	Low variance (Bad)

2.2.3 การเรียนรู้แบบรวมกลุ่ม (Ensemble Learning Method)

คือการรวมโมเดลการเรียนรู้ที่หลากหลาย แตกต่าง และอิสระต่อกันเข้าด้วยกันเพื่อเพิ่มประสิทธิภาพของโมเดลและต้องพิจารณา Bias-Variance trade-off ของ error ตามรูปที่ 22 เพื่อให้สอดคล้องกับการนำไปใช้งาน



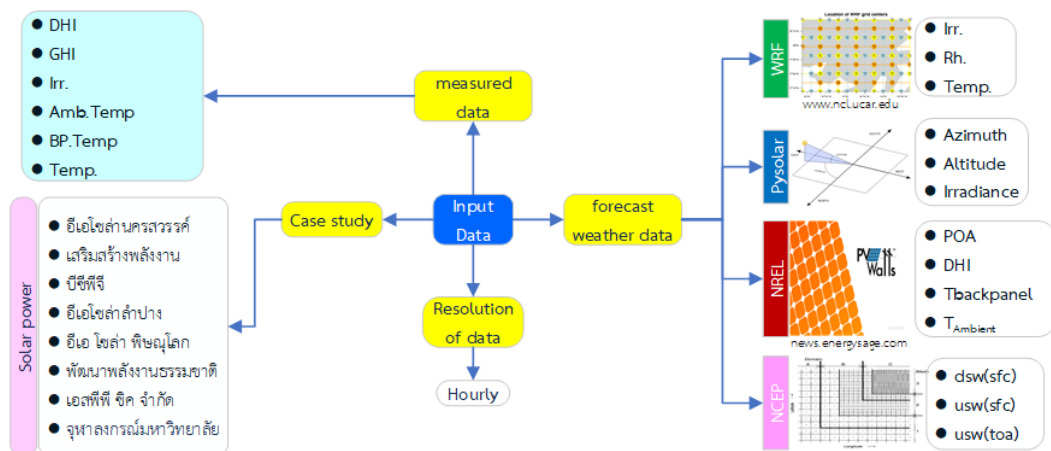
รูปที่ 22 Bias-Variance trade-off

Ensemble Learning มีหลากหลายเทคนิคได้แก่ Bagging, boosting และ Voting เป็นต้น

1. Bootstrap Aggregating (Bagging) คือสุ่มแบ่งข้อมูลฝึกสอนออกเป็นหลายชุดแต่สร้างโมเดลด้วยอัลกอริทึมเดียวกันทั้งหมด เช่น Decision Trees, Random Forest และ Extra Trees เป็นต้น
2. Boosting คือทำการทดสอบกับข้อมูลชุดเดียวกัน โดยทำการทดสอบแบบ iteration (วนปรับปรุงค่าน้ำหนักหลายๆรอบ) ข้อดีคือช่วยปรับลด bias error ของโมเดลได้ดี เช่น AdaBoost และ Stochastic Gradient Boosting
3. Voting คือใช้ข้อมูลฝึกสอนชุดเดียวกันแต่สร้างโมเดลด้วยเทคนิคต่างกัน

2.2.4 แบบจำลองพยากรณ์อากาศ

เพื่อให้สามารถเข้าใจระบบพยากรณ์กำลังผลิตไฟฟ้าได้อย่างเหมาะสมจำเป็นที่จะต้องทราบข้อมูลเบื้องต้น ข้อมูลพยากรณ์มี 3 ประเภท ได้แก่ ข้อมูลกำลังผลิตไฟฟ้า ข้อมูลพยากรณ์อากาศ และข้อมูลสภาพอากาศที่วัดได้จากอุปกรณ์ตรวจวัดดังแสดงรูปที่ 23 สำหรับข้อมูลพยากรณ์อากาศเราใช้สามแบบจำลองสภาพอากาศที่มีให้ใช้งานฟรีได้แก่ Pysolar, PVWatts และ NCEP Forecast แบบจำลองทั้งสามให้ค่าตัวแปรต่างๆดังแสดงในตารางที่ 9 และรูปที่ 48 ซึ่งประกอบไปด้วยตัวแปรต่าง ๆ ดังนี้ Diffuse Horizontal Irradiance (I_{DHI}), Global Horizontal Irradiance (I_{GHI}), ambient temperature at the array height ($T_{Ambient}$), Back Panel Temperature ($T_{BackPanel}$), Downward shortwave Radiation at Surface (dswsfc), Upward shortwave Radiation at Surface (uswsfc), และ Upward shortwave Radiation at the Top of the Atmosphere (uswtoa)



รูปที่ 23 ภาพรวมอินพุตประเภทต่างๆของแบบจำลอง

2.2.4.1 Pysolar

Pysolar เป็นโมเดลที่ให้ค่าความเข้มแสงอาทิตย์ มุมอาซิมุท (Azimuth) และ มุมอัลติจูด (Altitude) ของดวงอาทิตย์ในทุก ๆ ตำแหน่งบนโลก

2.2.4.2 National Renewable Energy Laboratory (NREL)

PWatts เป็นแพลตฟอร์มที่พัฒนาขึ้นโดย U.S. Department of Energy's National Renewable Energy Laboratory (NREL) ที่ให้ค่าความเข้มแสงอาทิตย์ในพื้นที่โรงไฟฟ้าพลังงานแสงอาทิตย์ตามการระบุละติจูดและลองจิจูด

2.2.4.3 National Centers for Environmental Prediction (NCEP)

NCEP forecast ที่พัฒนาขึ้นโดย National Centers for Environmental Prediction (NCEP) เป็นโมเดลที่ให้ค่าพยากรณ์อากาศเริ่มพยากรณ์ 4 ครั้งต่อวันที่เวลา 0000, 0600, 1200, และ 1800 UTC โมเดลนี้มีสี่ความละเอียดให้เลือกใช้ได้แก่ 0.2, 0.5, 1.0, และ 2.5 Degree Horizontal Resolutions โดยให้ค่าพยากรณ์ทุกๆชั่วโมง [31]

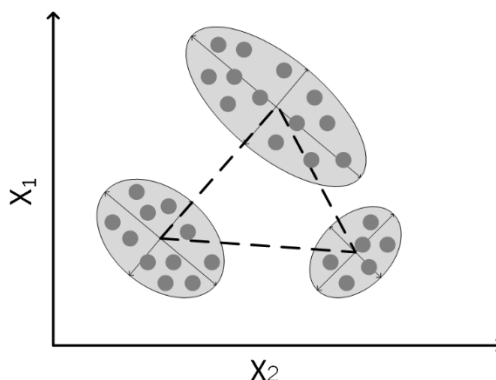
2.2.4.4 Measured Data (ค่าวัดจริง)

ข้อมูลกำลังผลิตไฟฟ้าที่วัดได้ทั้ง 7 โรงในประเทศไทยสามารถดาวน์โหลดได้จากฐานข้อมูลโรงไฟฟ้าพลังงานทดแทนของคณะกรรมการกำกับกิจการพลังงาน (กกพ.) เป็นข้อมูลแบบรายชั่วโมง ดังแสดงในรูปที่ 49 โดยตำแหน่งที่ตั้งและกำลังผลิตติดตั้งแสดงในตารางที่ 12

2.2.5 การจัดกลุ่มข้อมูล

การจัดกลุ่มคือการเรียนรู้แบบไม่มีผู้สอนซึ่งพยายามที่จะเรียนรู้โครงสร้างที่ซ่อนอยู่ภายในข้อมูล และจัดกลุ่มข้อมูลที่มีลักษณะคล้ายคลึงกันให้อยู่กลุ่ม (Cluster) เดียวกัน โดยที่ข้อมูลแต่ละตัว

นั้นไม่มีป้าย (label) กำกับ การจัดกลุ่มนั้นมีเป้าหมายคือ การสร้างกลุ่มที่มีระดับความคล้ายคลึงกันของตัวอย่างในแต่ละกลุ่มสูงและระดับความคล้ายคลึงกันของตัวอย่างระหว่างกลุ่มต่ำหรืออีกนัยหนึ่งคือ การจัดกลุ่มโดยให้มีความแปรผันภายในกลุ่ม (Within-cluster Variation) มีค่าน้อย และความแปรผันระหว่างกลุ่ม (Between-cluster Variation) มีค่ามาก ดังรูปที่



รูปที่ 24 การจัดกลุ่มโดยให้มีความแปรผันภายในกลุ่มมีค่าน้อยและระหว่างกลุ่มมีค่ามาก

2.2.6 ตัวชี้วัดพื้นฐาน (Metrics) สำหรับการวัดผลแบบจำลอง

อัลกอริทึมการเรียนรู้ของเครื่องมีตัวชี้วัดที่ใช้วัดความถูกต้องของแบบจำลองหลายตัว แบ่งเป็นสองกลุ่มใหญ่ ได้แก่

1. ปัญหาการถดถอย (Regression problem) ซึ่งมีตัวชี้วัด เช่น ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error; MAE), ค่าเฉลี่ยความผิดพลาดกำลังสอง (Mean Squared Error; MSE), RMSE, R-squared (R2) และ MBE
2. ปัญหาการจำแนกประเภทข้อมูล (Classification problem) ซึ่งมีตัวชี้วัด เช่น Accuracy, Precision, Recall, F1-Score, F-Beta Score, AUC

ข้อดีของตัวชี้วัดแต่ละตัวแสดงดังตารางที่ 4 และการคำนวณค่าความผิดพลาดหาได้จาก

ค่ากำลังเฉลี่ย $\bar{P}_7 = E[P_7]$ เมื่อ P_7 คือค่ากำลังที่วัดได้ตอน 7:00 น.

ค่าไบแอสเฉลี่ยหาได้จาก $MBE_7(Watt) = E[\hat{P}_7 - P_7]$ เมื่อ \hat{P}_7 คือค่ากำลังที่พยากรณ์ได้จากแบบจำลอง ซึ่งเรากำหนดให้ MBE มีนิยามเป็น

$$MBE_7(\%) \triangleq \frac{MBE_7(Watt)}{\bar{P}_7} * 100 \text{ ซึ่งนิยามนี้จะทำให้เราทราบว่าแบบจำลองเรา}$$

ให้ค่าไบแอสเทียบกับค่าเฉลี่ยเป็นอย่างไรต่อมาเป็นค่าที่บอกกำลัง RMS ของ Error

$$RMSE_7(Watt) = \sqrt{E[\hat{P}_7 - P_7]^2} \text{ และ } RMSE_7(\%) \triangleq \frac{RMSE_7(Watt)}{iCap} * 100$$

ค่านิยามที่ใช้ทั่วไปนี้อาจพิจารณาอะไรไม่ได้มากเนื่องจากกำลังผลิตติดตั้ง (iCapacity) เยอะมาก ถ้าให้

$$RMSE_7^*(\%) \triangleq \frac{RMSE_7(Watt)}{\sqrt{E[P_7^2]}} * 100$$

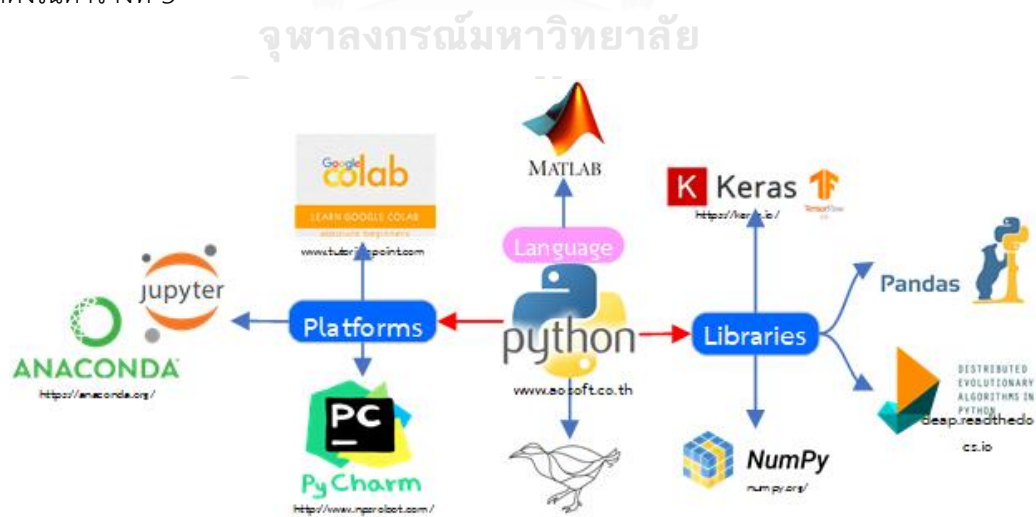
ซึ่งนิยามที่กำหนดนี้จะสื่อความหมายได้ดีกว่าเพราะ $\sqrt{E[P_7^2]}$ คือค่ากำลัง RMS ของ P_7

ตารางที่ 4 ข้อดีของตัวชี้วัดแต่ละประเภท

ตัวชี้วัด	ข้อดี
MSE	MSE จะดีกว่า MAE เวลาใช้อัลกอริทึมประเภท Gradient Descent Algorithm เนื่องจากว่าตอนหาอนุพันธ์ (Derivative) การหาอนุพันธ์ของสมการฟังก์ชันสูญเสีย (Loss function, Cost function) ด้วยวิธีการทางแคลคูลัสจะทำได้ง่ายกว่า
RMSE	- RMSE เป็นรากที่สองของ MSE เพื่อให้ได้ค่า Loss ที่มีหน่วยเดียวกับตัวแปร y เนื่องจากว่าการยกกำลังสองค่าความผิดพลาด ก่อนหาค่าเฉลี่ยทำให้หน่วยมันเปลี่ยนไปจากเดิม - เมื่อเปรียบเทียบทั้ง 3 ตัวชี้วัด (MAE, MSE และ RMSE) RMSE จะแปลงผลง่ายที่สุด ยกตัวอย่างเช่น ถ้า RMSE เท่ากับ 3.56 แปลว่าโดยเฉลี่ยโมเดลทำนาย y ผิดไป +/- 3.56 จุด

2.2.7 ภาษาและแพลตฟอร์มสำหรับการพัฒนาแบบจำลอง

ภาษาและแพลตฟอร์มสำหรับการพัฒนาแบบจำลองแต่ละประเภทมีข้อดีข้อเสียแตกต่างกันดังแสดงในตารางที่ 5



รูปที่ 25 ภาษาและแพลตฟอร์มสำหรับการพัฒนาแบบจำลอง

ตารางที่ 5 ข้อดีและข้อเสียของแพลตฟอร์ม

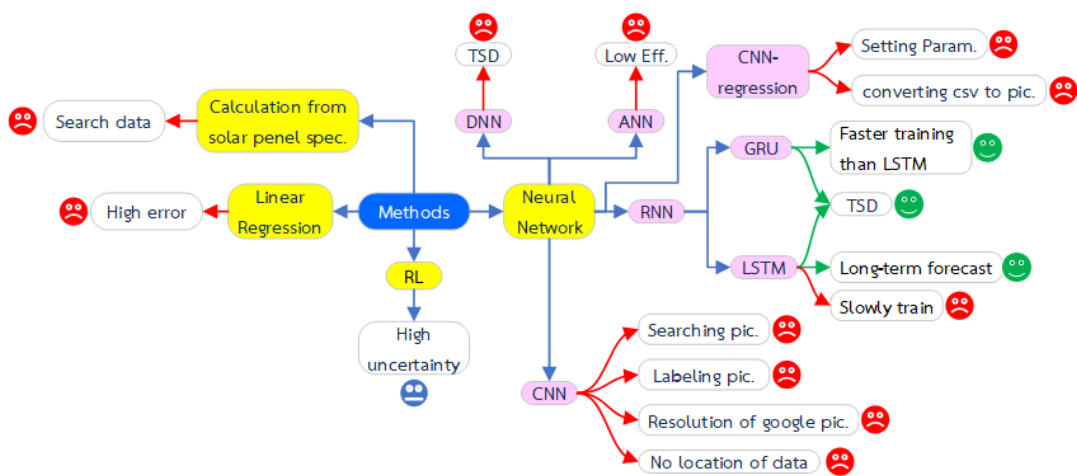
แพลตฟอร์ม	ข้อดี	ข้อเสีย
Goggle Colab	ไลบรารีส่วนใหญ่ถูก ติดตั้งมาให้แล้ว	อยากต่อการตรวจสอบตำแหน่งผิดพลาดของโค้ด
Pycharm	ง่ายต่อการตรวจสอบ ตำแหน่งผิดพลาดของโค้ด	บางไลบรารีต้องติดตั้งเอง
Weka	ง่ายต่อการดูประสิทธิภาพคร่าวๆ	ไม่สามารถปรับตัวแปรได้ในระดับที่ละเอียดมาก



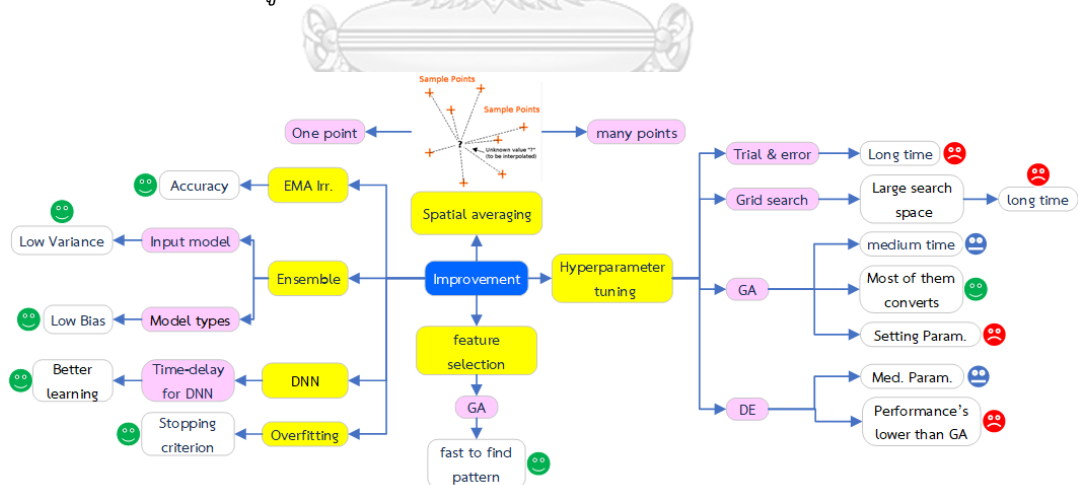
บทที่ 3

วิธีการดำเนินการวิจัย

วิธีการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์มีหลากหลายวิธีด้วยกันซึ่งแต่ละวิธีนั้นมีข้อดีและข้อเสียแตกต่างกันดังแสดงในรูปที่ 26 และวิธีการปรับปรุงประสิทธิภาพแบบจำลองแต่ละวิธีนั้นมีข้อดีและข้อเสียแตกต่างกันดังแสดงในรูปที่ 27 ซึ่งวิธีการปรับปรุงประสิทธิภาพแบบจำลองถูกนำมาประยุกต์ใช้ในวิทยานิพนธ์นี้ในหลายด้านดังแสดงในตารางที่ 6



รูปที่ 26 ข้อดีและข้อเสียของแต่ละวิธีการพยากรณ์

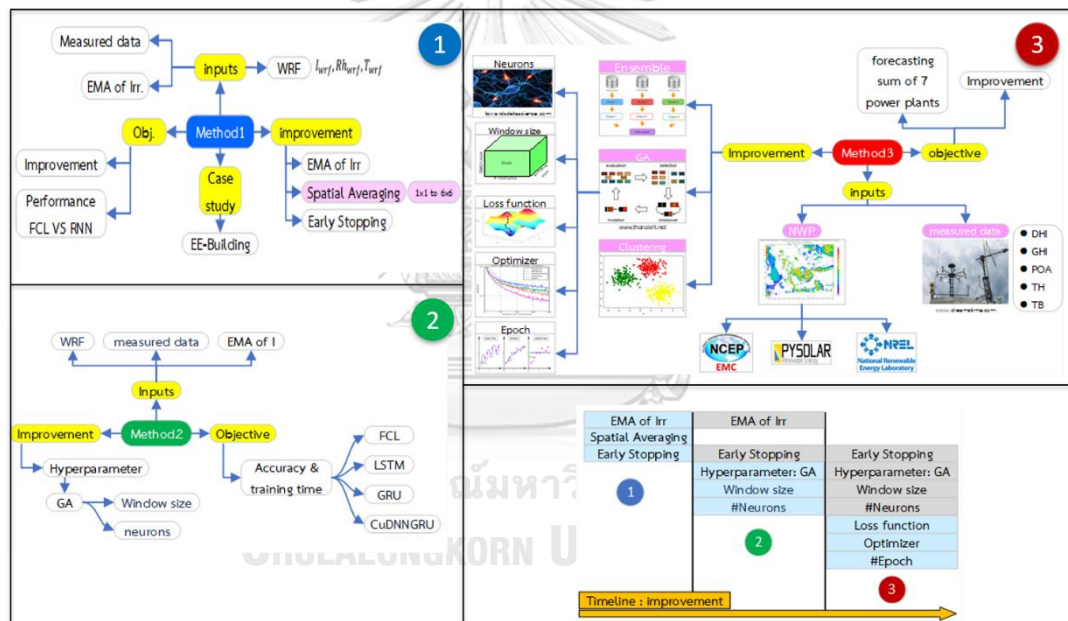


รูปที่ 27 ข้อดีและข้อเสียของแต่ละวิธีการปรับปรุงประสิทธิภาพแบบจำลอง

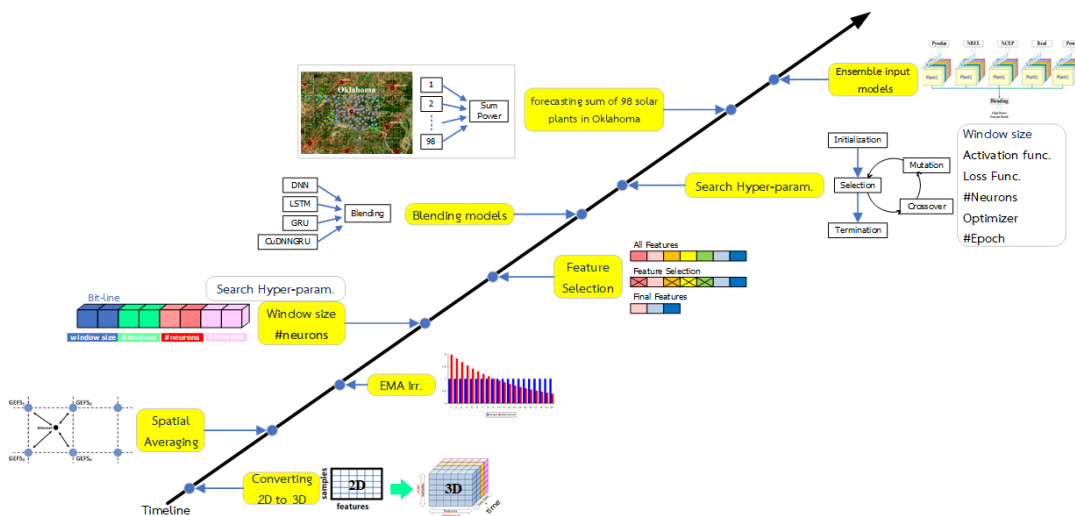
ตารางที่ 6 การประยุกต์ใช้วิธีการต่างๆในการปรับปรุงประสิทธิภาพ

Methods	Hyperparameter tuning using GA	Feature Selection -GA	Ensemble Input models	Ensemble Type models
Objective	<ul style="list-style-type: none"> Window size #Neurons Activation Func. Loss Func. Optimizer #Epoch 	<ul style="list-style-type: none"> Irr Rh Temp. 	<ul style="list-style-type: none"> WRF Pysolar NCEP Measured data 	<ul style="list-style-type: none"> DNN GRU LSTM CuDNNGRU

ซึ่งภาพรวมของทั้งสามวิธีสำหรับการพยากรณ์และการปรับปรุงประสิทธิภาพที่ทำการศึกษา แสดงดังรูปที่ 28 ในแง่มุมมองของข้อมูลอินพุต วัตถุประสงค์ กรณศึกษา และการปรับปรุงประสิทธิภาพ แบบจำลอง และไทม์ไลน์ของการปรับปรุงประสิทธิภาพแบบจำลองแสดงในรูปที่ 29



รูปที่ 28 ภาพรวมของวิธีการพยากรณ์ทั้งสามวิธีการ



รูปที่ 29 ไทม์ไลน์ของการปรับปรุงประสิทธิภาพแบบจำลอง

3.1 การพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ด้วยแบบจำลองการพยากรณ์อากาศเชิงตัวเลข (Very Short-Term SPF using data from NWP model)

3.1.1 ภาพรวมวิธีการดำเนินการวิจัย

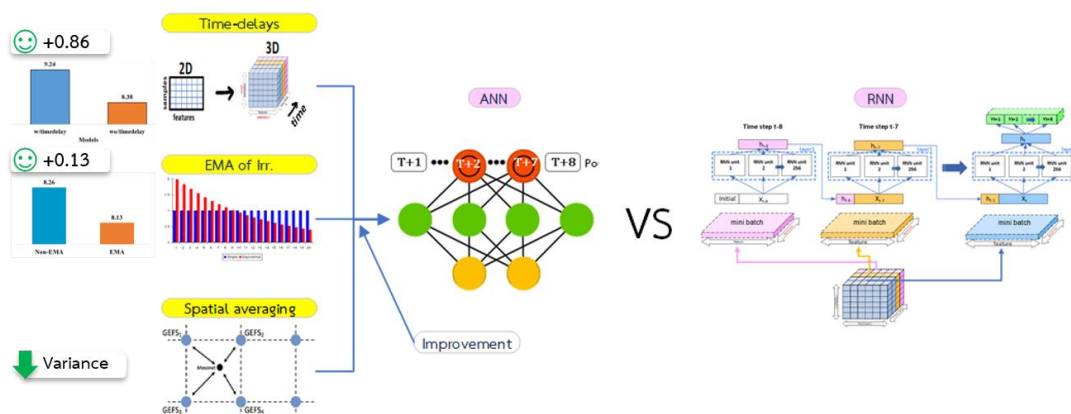
การศึกษาการปรับปรุงประสิทธิภาพของการพยากรณ์กำลังผลิตไฟฟ้าจากแบบจำลองการพยากรณ์อากาศเชิงตัวเลขแสดงดังรูปที่ 30 ซึ่งประกอบไปด้วย 4 วิธีการปรับปรุงประสิทธิภาพ ได้แก่

1. การประยุกต์ใช้ค่าวัดจริง (measured data)
2. การประยุกต์ใช้ค่าเฉลี่ยเคลื่อนที่แบบฟังก์ชันเอกซ์โพเนนเชียลของความเข้มแสงอาทิตย์ที่วัดได้ (Exponential Moving Average (EMA) of Measured Irradiance)
3. การเฉลี่ยค่าพยากรณ์อากาศเชิงพื้นที่ (spatial averaging) ของพื้นที่ขนาด $1 \times 1, 2 \times 2, 3 \times 3, 4 \times 4, 5 \times 5$ และ 6×6 จุดกริด (grid points)
4. การเพิ่มมิติของเวลา (Time-delays) ให้กับแบบจำลอง FCL

การศึกษาทำการเปรียบเทียบประสิทธิภาพระหว่างเน็ตเวิร์ก FCL และ RNN ซึ่งเป็นการพยากรณ์ไปล่วงหน้า 4 ชั่วโมงและดัชนีวัดประสิทธิภาพ (Performance Index) ที่ใช้ในการวัดผล ได้แก่ RMSE และ ค่าคลาดเคลื่อนจากความเอนเอียงเฉลี่ย (Mean Bias Error; MBE)

$$RMSE = \frac{100\%}{install\ cap} \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{P}(t) - P(t))^2} \quad (18)$$

$$MBE = \frac{1}{N} \sum_{t=1}^N (\hat{P}(t) - P(t)) \quad (19)$$



รูปที่ 30 ภาพรวมของวิธีการพยากรณ์และการปรับปรุงประสิทธิภาพวิธีที่หนึ่ง

3.1.2 ข้อมูลอินพุต

การพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์บนอาคารบูรณาการบิโชนทัศน์ติดตั้งแสดงในรูปแบบที่ 31 ซึ่งมีขนาดกำลังผลิตติดตั้ง 8 kw อินพุตของเน็ตเวิร์กนี้ประกอบไปด้วย 7 ตัวแปรซึ่งเป็นข้อมูลตั้งแต่วันที่ 1 มกราคม 2017 ถึง 31 ตุลาคม 2018 และความละเอียดของข้อมูลเป็นข้อมูลรายครึ่งชั่วโมง ดังต่อไปนี้

1. Irradiance forecast (I_{wrf}) คือความเข้มแสงอาทิตย์จากแบบจำลองพยากรณ์อากาศเชิงตัวเลข WRF (ในหน่วย W/m^2)
2. Relative humidity forecast (Rh_{wrf}) คือความชื้นสัมพัทธ์จากแบบจำลองพยากรณ์อากาศเชิงตัวเลข WRF (ในหน่วย %)
3. Temperature forecast (T_{wrf}) คืออุณหภูมิจากแบบจำลองพยากรณ์อากาศเชิงตัวเลข WRF (ในหน่วย $^{\circ}C$)
4. Exponential Moving Average (EMA) of Measured Irradiance คือค่าเฉลี่ยเคลื่อนที่แบบฟังก์ชันเอกซ์โพเนนเชียลของความเข้มแสงอาทิตย์ที่วัดได้ (ในหน่วย W/m^2) โดยมีจำนวน time-step ของการมอยย้อนกลับคือ 9 time-step
5. Measured Irradiance (I_{mea}) คือความเข้มแสงอาทิตย์ที่วัดได้ (ในหน่วย W/m^2)
6. Measured Temperature (T_{mea}) คืออุณหภูมิที่วัดได้ (ในหน่วย $^{\circ}C$)
7. Measured Power (P_{mea}) คือกำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ที่วัดได้ (ในหน่วย MW)

ฟังก์ชัน EMA มีสมการตาม (20) เป็นฟังก์ชันที่สร้างข้อมูลลำดับเวลา (time series data) โดยฟังก์ชันนี้ให้ความสำคัญกับข้อมูลในปัจจุบันมากกว่าอดีตเนื่องจากข้อมูลในอดีตได้ถูกรีเคอร์ซีฟลดค่าลงไปเรื่อยๆตาม (20)

$$I_{EMA} = \begin{cases} I_t, & t = 0 \\ \alpha \cdot I_t + (1 - \alpha) \cdot I_{t-1}, & t > 1 \end{cases} \quad (20)$$

โดยที่ α เป็นค่าคงที่ (constant) ซึ่งการศึกษานี้กำหนดให้มีค่าเท่ากับ 0.9, t คือเวลา, I คือค่าความเข้มแสงอาทิตย์

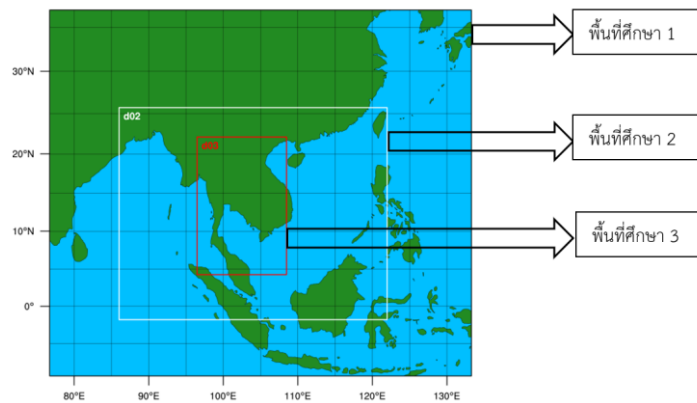


รูปที่ 31 อาคารของแผงโซลาร์เซลล์ที่ติดตั้งบนอาคารบูรณาการอดบิมทสันต์

The Weather Research and Forecasting Model (WRF) เป็นแบบจำลองที่พัฒนาโดย *National Center for Atmospheric Research (NCAR)* ประเทศสหรัฐอเมริกาซึ่งสามารถนำมาประยุกต์ใช้ศึกษาปรากฏการณ์ทางบรรยากาศในความละเอียดระดับ 10 เมตร จนถึง 1000 กิโลเมตร *WRF* เป็นแบบจำลองแบบสาธารณะ (community) ที่เปิดให้บุคคลทั่วโลกสามารถดาวน์โหลดใช้ได้ผ่านเว็บไซต์ www.mmm.ucar.edu/wrf โดยมีการนำไปศึกษาวิจัยในหลาย ๆ ภูมิภาคทั่วโลก เช่น สหรัฐอเมริกา ยุโรป เอเชีย ออสเตรเลีย เป็นต้น

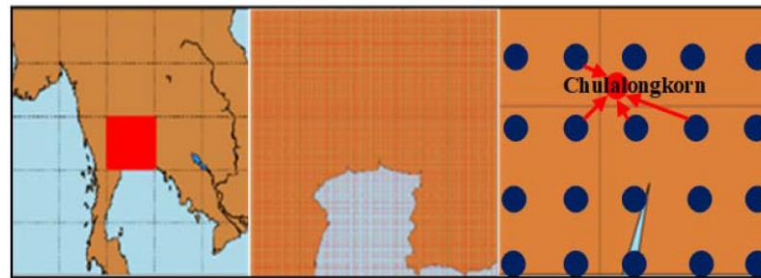
สำหรับกรมอุตุนิยมวิทยาของประเทศไทยระยะเวลาและความละเอียดของข้อมูลการพยากรณ์อากาศที่มีให้บริการประกอบไปด้วย 3 โดเมนด้วยกันดังแสดงดังรูปที่ 32

1. พยากรณ์ล่วงหน้า 10 วัน โดยมีรายละเอียดเชิงพื้นที่ 18 กิโลเมตร
2. พยากรณ์ล่วงหน้า 72 ชั่วโมง โดยมีรายละเอียดเชิงพื้นที่ 6 กิโลเมตร
3. พยากรณ์ล่วงหน้า 48 ชั่วโมง โดยมีรายละเอียดเชิงพื้นที่ 2 กิโลเมตร



รูปที่ 32 ตัวอย่างโดเมนของแบบจำลอง WRF ในพื้นที่ประเทศไทย
(ภาพจากสำนักงานสารสนเทศทรัพยากรน้ำและการเกษตร)

สำหรับแบบจำลองพยากรณ์อากาศ WRF ที่ใช้ในการศึกษานี้ประกอบไปด้วยตัวแปรต่างๆ ดังต่อไปนี้ อุณหภูมิ (Temperature) ความเข้มแสง (Irradiance) และ ความชื้นสัมพัทธ์ (Relative humidity) แสดงดังรูปที่ 34

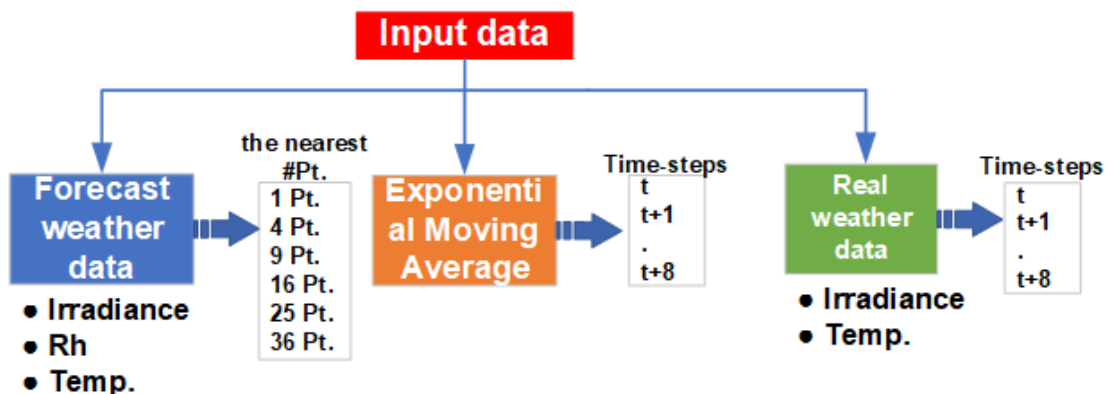


รูปที่ 33 โดเมนของแบบจำลอง WRF ซึ่งครอบคลุมกรณีศึกษา

CHULALONGKORN UNIVERSITY

ข้อมูลค่าวัดจริงจากอุปกรณ์ตรวจวัดเป็นข้อมูลที่ได้จากโครงการระบบบริหารจัดการพลังงานในอาคารจุฬาลงกรณ์มหาวิทยาลัย (Chulalongkorn University Building Energy Management System; CUBEMS) ซึ่งเป็นระบบที่เก็บและรวบรวมข้อมูลจากอุปกรณ์ตรวจวัดต่างๆ ที่ติดตั้งที่คณะวิศวกรรมศาสตร์จุฬาลงกรณ์มหาวิทยาลัย เช่น Irradiance (I_{mea}), Temperature (T_{mea}), Relative Humidity (Rh_{mea}), UV index (UV_{mea}) ซึ่งมีความละเอียดเป็นรายนาที่

ข้อมูลอินพุตสามารถแบ่งออกได้ 3 ประเภท ได้แก่ ข้อมูลค่าวัดจริงจากอุปกรณ์ตรวจวัด ข้อมูลสภาพอากาศพยากรณ์และข้อมูลค่าเฉลี่ยเคลื่อนที่แบบฟังก์ชันเอกซ์โพเนนเชียลของความเข้มแสงอาทิตย์ที่วัดได้ซึ่งแสดงดังรูปที่ 34



รูปที่ 34 ข้อมูลอินพุต 3 ประเภท
(ข้อมูลพยากรณ์อากาศ ข้อมูลค่าเฉลี่ยเคลื่อนที่ และ ข้อมูลค่าวัดจริง)

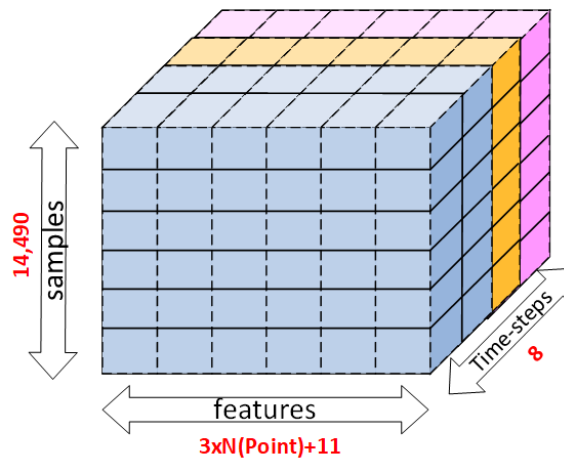
จำนวนตัวอย่าง (sample) ทั้งหมดเท่ากับ 14,490 (100%) ตัวอย่างซึ่งแบ่งออกเป็น 3 ชุดข้อมูล ได้แก่

1. ชุดข้อมูลฝึกสอน (Training set) แบ่งมา 70% หรือ 10,143 ตัวอย่าง (ตัวอย่างที่ 1 ถึง 10143)
2. ชุดข้อมูลตรวจสอบ (validation set) แบ่งมา 20% หรือ 2,898 ตัวอย่าง (ตัวอย่างที่ 10143 ถึง 13041)
3. ชุดข้อมูลทดสอบ (Testing set) แบ่งมา 10% หรือ 1,449 ตัวอย่าง (ตัวอย่างที่ 13041 ถึง 14490)

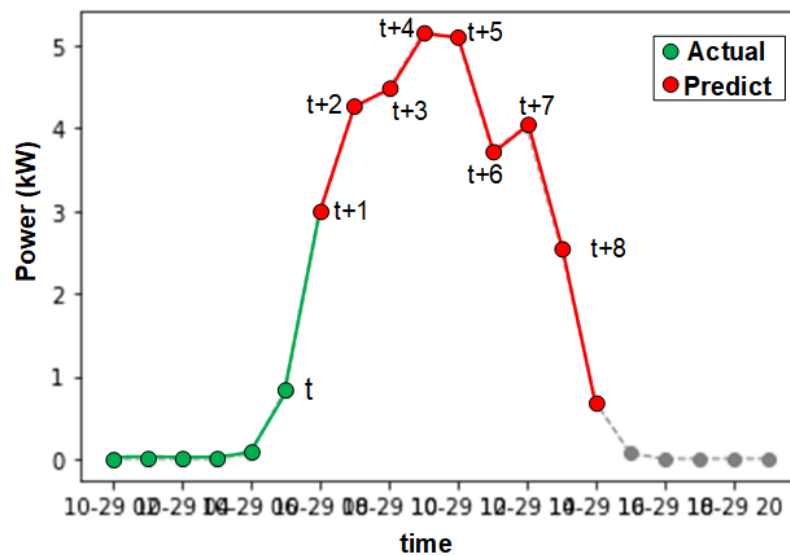
3.1.3 พารามิเตอร์ของเน็ตเวิร์ก

สำหรับส่วนนี้จะอธิบายถึงพารามิเตอร์ต่างๆของเน็ตเวิร์ก ได้แก่

1. Time step (t) หรือ Window size หรือมิติความลึก (ด้านยาว) ถ้ามองเป็นลูกบาศก์ดังรูปที่ 35 สำหรับการเรียนรู้กำหนดให้มีค่าเท่ากับ 8 ซึ่งจะเห็นได้ว่าแต่ละตัวอย่าง (แกนตั้ง) แต่ละตัวแปร (แกนนอน) จะถูกมองย้อนกลับไป 8 time-steps (แกนลึก)
2. HORIZON จำนวน time-step ที่ต้องการพยากรณ์ไปข้างหน้าสำหรับการเรียนรู้ กำหนดให้มีค่าเป็น 8 การพยากรณ์ไปล่วงหน้า 8 time-steps ($t + 1$ ถึง $t + 8$) หรือ 4 ชั่วโมงล่วงหน้าเนื่องจากแต่ละจุดห่างกันครึ่งชั่วโมง ดังแสดงในรูปที่ 36



รูปที่ 35 ลักษณะของชุดข้อมูลฝึกสอนในมิติต่างๆ



รูปที่ 36 การพยากรณ์ไปล่วงหน้า 8 HORIZON (8 time-steps)

3. พารามิเตอร์ภายในของเน็ตเวิร์กแสดงในตารางที่ 7 ได้แก่
 - 3.1 จำนวนนิวรอน (neurons) ซึ่งในที่นี้เป็นเซลล์นิวรอนธรรมดาซึ่งในไลบรารีของ keras เป็นที่รู้จักกันในชื่อของ *Dense layer* หรือ *Fully Connected Layer*
 - 3.2 ฟังก์ชันกระตุ้น (activation) สำหรับการศึกษานี้ใช้
4. Epoch คือจำนวนรอบในการฝึกสอนซึ่งแต่ละรอบอัลกอริทึมจะวนผ่านตัวอย่างทั้งหมด
5. Batch Size คือการปรับค่าน้ำหนักทุกๆชุดข้อมูลเล็กๆ (Batch) กล่าวอีกนัยหนึ่งก็คือเมื่อฝึกสอนครบทุกตัวอย่างในชุดข้อมูลเล็กๆนั้นแล้วเน็ตเวิร์กจะทำการปรับค่าน้ำหนักของมัน

6. ฟังก์ชันสูญเสียคือการคำนวณความผิดพลาดว่าค่าพยากรณ์ที่แบบจำลองทำนายออกมาต่างจากค่าจริงอยู่เท่าไรแล้วหาค่าเฉลี่ยฟังก์ชันเป้าหมายที่จะถูกทำให้มีค่าความผิดพลาดน้อยลงด้วยตัวออปติไมเซอร์ (Optimizer) หรือค่าความดีของแบบจำลองที่เราพยายามปรับเพื่อให้ได้แบบจำลองที่ดีที่สุด ซึ่งจำนวนพารามิเตอร์ในแต่ละเลเยอร์แสดงในรูปที่ 37

ตารางที่ 7 พารามิเตอร์ต่างๆของเน็ตเวิร์กสำหรับกรณีศึกษา

Parameter	Layer		
	Input	Hidden	Out
neurons	$3 \times N(\text{point}) + 11$	256	8
Activation	-	ReLU	-
Epoch: batch size	100: 32		
Optimizer: loss	Adam: mae		

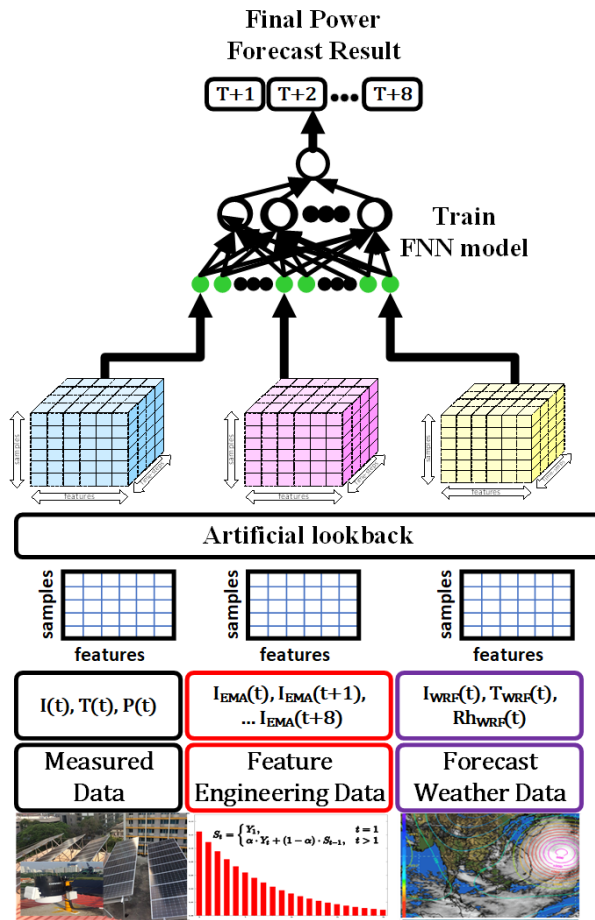
Model: "sequential_1"

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 256)	49408
dense_2 (Dense)	(None, 8)	2056
Total params: 51,464		
Trainable params: 51,464		
Non-trainable params: 0		

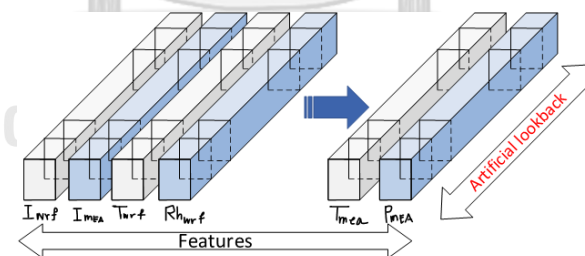
รูปที่ 37 จำนวนพารามิเตอร์ของเน็ตเวิร์กในแต่ละเลเยอร์สำหรับการกรณีศึกษา

3.1.4 โครงสร้างเน็ตเวิร์ก FCL และ RNN

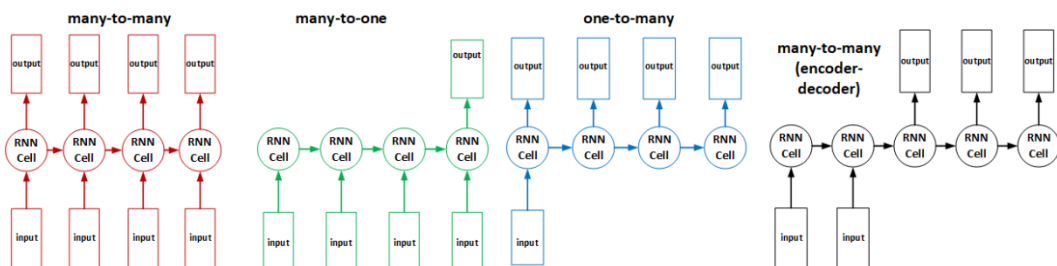
จากรูปที่ 38 แสดงโครงสร้างเน็ตเวิร์ก Fully Connected Layer (FCL) และด้านล่างของรูปจะเป็นส่วนของอินพุตประกอบไปด้วย 3 ตัวแปร ตามที่กล่าวใน 3.1.2 ถัดขึ้นมาจะเป็นส่วนของข้อมูลการมองย้อนกลับเทียม (Artificial lookback data) ดังแสดงในรูปที่ 39 และรูปที่ 41 แสดงโครงสร้างเน็ตเวิร์ก RNN ซึ่งมีโครงสร้างแบบต่างๆ



รูปที่ 38 ภาพรวมของเน็ตเวิร์ก FCL และข้อมูลอินพุต



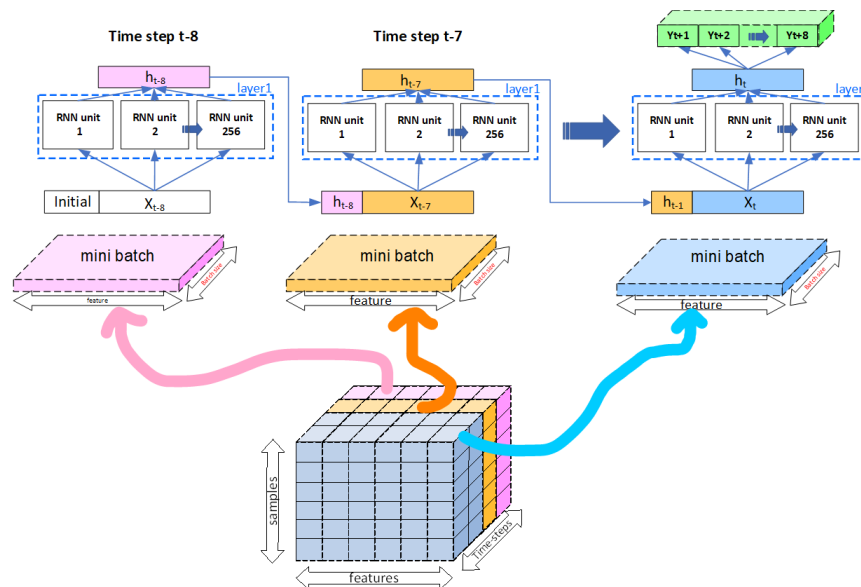
รูปที่ 39 ข้อมูลการมองย้อนกลับเทียม (artificial lookback data)



รูปที่ 40 การต่อเน็ตเวิร์กในลักษณะต่างๆของ RNN

3.1.5 เทคนิคการฝึกสอน

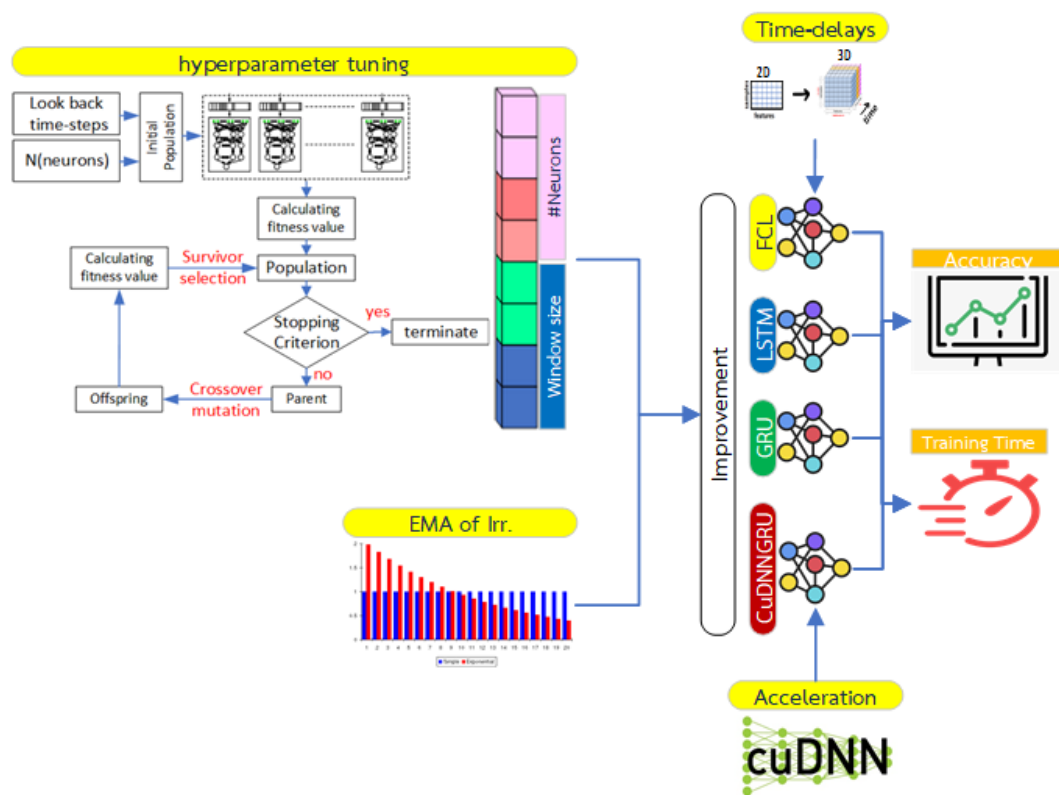
มีการใช้เทคนิคที่ชื่อว่า *Early stopping* ซึ่งเป็นเทคนิคการสร้างความสัมพันธ์ระหว่างค่าความเอนเอียง (Bias) และค่าความแปรปรวน (Variation) เพื่อไม่ให้แบบจำลองอิงกับข้อมูลฝึกสอนมากเกินไปหรือที่เรียกว่า *Overfitting* นั่นคือการสังเกตค่า *Validation loss* ในทุกๆ epoch โดยมีเงื่อนไขว่าถ้า *Validation loss* ไม่ดีขึ้น 10 Epoch ติดต่อกันให้หยุดการฝึกสอน



รูปที่ 41 การต่อเน็ตเวิร์ก RNN สำหรับกรณีศึกษา

3.2 การพยากรณ์กำลังผลิตไฟฟ้าระยะสั้นโดยใช้โครงข่ายประสาทเทียมและขั้นตอนวิธีเชิงพันธุกรรม (Very Short-Term SPF using GA Based Neural Network)

เน็ตเวิร์กที่ใช้กันปัจจุบันมีหลากหลายประเภท ไม่ว่าจะเป็น FCL, LSTM, และ GRU หรือแม้กระทั่ง CuDNNGRU ที่มีการเรียกใช้การ์ดแสดงผล (GPU) เพื่อช่วยในการประมวลผล เน็ตเวิร์กแต่ละประเภทมีความแม่นยำและระยะเวลาที่ใช้ในการฝึกสอนแตกต่างกัน ดังนั้นในการศึกษานี้เราจะทำการศึกษาว่าแบบจำลองใดให้ค่าความผิดพลาดน้อยและในขณะเดียวกันก็ใช้ระยะเวลาในการฝึกสอนน้อยตามไปด้วยเช่นกันดังแสดงในรูปที่ 42



รูปที่ 42 ภาพรวมการพยากรณ์และวิธีการปรับปรุงประสิทธิภาพด้วยวิธีที่สอง

3.2.1 ภาพรวมของวิธีการดำเนินการวิจัย

วิธีการดำเนินการวิจัยคือการค้นหาค่าที่เหมาะสมของไฮเปอร์พารามิเตอร์ของเน็ตเวิร์กดังแสดงในรูปที่ 43 โดยแสดงการเปรียบเทียบประสิทธิภาพของเน็ตเวิร์ก 4 ประเภท ได้แก่

1. เน็ตเวิร์กเชื่อมต่อแบบทั่วถึง (fully connected layer; FCL) หรือที่รู้จักในการเขียนโค้ดโปรแกรม Keras ว่า Dense layer
2. หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory; LSTM)

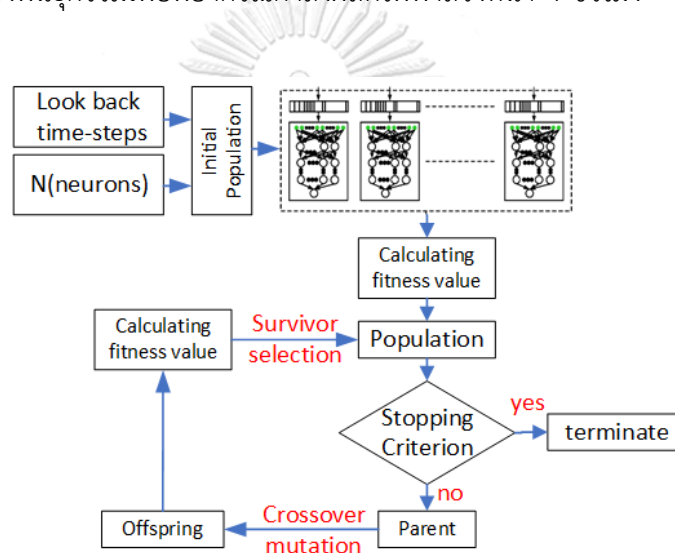
3. Gated Recurrent Unit (GRU)
4. Cuda Deep Neural Network Gated Recurrent Unit (CuDNNGRU)

CuDNNGRU เป็นไลบรารีที่มีชื่อว่า *Deep Learning Library* ที่ถูกพัฒนาโดย *NVIDIA* ซึ่งเป็นไลบรารีที่ทำหน้าที่เพิ่มความเร็วของ GRU ซึ่งโดยทั่วไปถูกเรียกว่า Fast GRU

ไฮเปอร์พารามิเตอร์ที่ทำการค้นหาทั้งหมดประกอบไปด้วย 2 พารามิเตอร์ ได้แก่

1. จำนวนสถานะเวลาในการมอยย้อนกลับ (Lookback time-steps) หรือเรียกอีกอย่างว่าขนาดของหน้าต่างที่มอยย้อนกลับ (window size) หรือความลึกของมิติเวลา
2. จำนวนนิวรอนในแต่ละชั้นซ่อน (the number of neurons in each hidden layers)

โดยใช้ขั้นตอนวิธีเชิงพันธุกรรมเพื่อพยากรณ์กำลังผลิตไฟฟ้าล่วงหน้า 4 ชั่วโมง



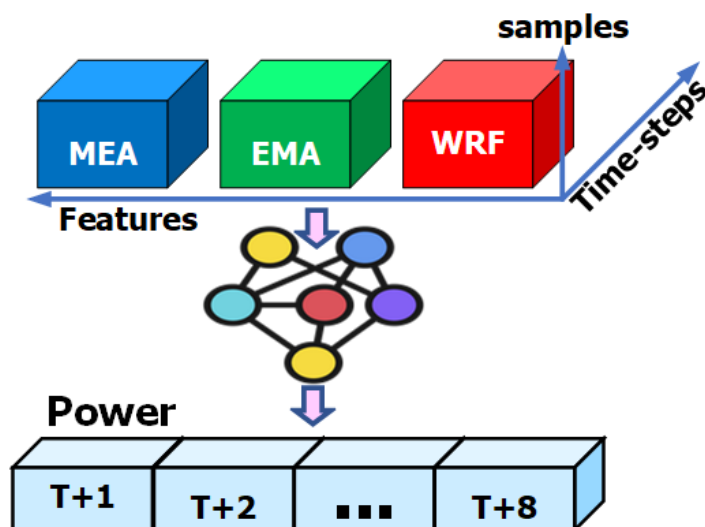
รูปที่ 43 ขั้นตอนการค้นหาไฮเปอร์พารามิเตอร์ของเน็ตเวิร์ก

CHULALONGKORN UNIVERSITY

3.2.2 ข้อมูลอินพุต

อินพุตของเน็ตเวิร์กเป็นข้อมูลที่มีความละเอียดรายครึ่งชั่วโมงแบ่งออกเป็น 3 ประเภทดังแสดงในรูปที่ 44

1. ข้อมูลพยากรณ์อากาศจากแบบจำลอง WRF ได้แก่ $I_{wrf}(t)$, $Rh_{wrf}(t)$, $T_{wrf}(t)$
2. ข้อมูลค่าวัดจริงจากอุปกรณ์ตรวจวัด ได้แก่ $P_{mea}(t)$, $I_{mea}(t)$, $T_{mea}(t)$
3. ค่าจากการทำ feature engineering คือ $I_{ema}(t)$, $I_{ema}(t + 1)$, ..., $I_{ema}(t + 8)$



รูปที่ 44 อินพุตของเน็ตเวิร์กในลักษณะสามมิติ

3.2.3 พารามิเตอร์ต่างๆที่ใช้ในการวิจัย

ในกระบวนการขั้นตอนวิธีเชิงพันธุกรรมพารามิเตอร์ต่างๆประกอบไปด้วย

1. การกำหนดประชากรรุ่นแรก (initialize) จะทำการสุ่มโดยใช้ Bernoulli Distribution ในส่วนของ การสลับสายพันธุ การกลายพันธุ และการเลือกตำแหน่งของ Roulette Wheel จะถูกสุ่มโดยใช้ Bernoulli Distribution เช่นเดียวกัน ขนาดของประชากร (population size) สำหรับกรณีศึกษานี้กำหนดให้มีค่าเป็น 5 และจำนวนรุ่นประชากร (generation) ก็กำหนดให้มีค่าเป็น 5 เช่นเดียวกัน ในขณะที่ความยาวของยีนกำหนดให้มีค่าเป็น 28 ซึ่งพารามิเตอร์เหล่านี้จะมีค่าแตกต่างกันออกไปขึ้นอยู่กับขนาดและชนิดของปัญหา
2. ฟังก์ชันค่าความเหมาะสมจะถูกคำนวณเพื่อวัดค่าความดีของแต่ละโครโมโซมตาม (21)

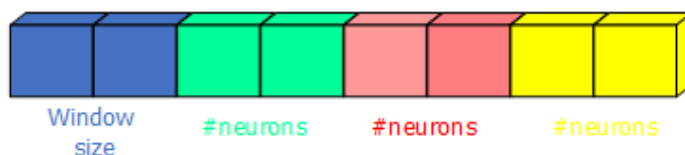
$$Fitness Function = \frac{1}{1 + NRMSE} \quad (21)$$

โดยที่ NRMSE คือความคลาดเคลื่อนกำลังสองเฉลี่ยแบบนอร์มัลไลซ์ (Normalized Root Mean Squared Error)

3. การเข้ารหัสโครโมโซม

ใช้รูปแบบไบนารีอาเรย์ (Binary Array) คือเป็นสายของบิตเรียงต่อกันดังแสดงในรูปที่ 45 โดยที่แต่ละตำแหน่งเป็นบิต 0 หรือ 1 ไบนารีอาเรย์นี้เป็นตัวแทนของประชากรหรือ

คำตอบที่ต้องการหาหลังจากผ่านขั้นตอนพันธุกรรมและถอดรหัสแล้วจะได้เป็นเลขฐานสิบ (decimal) ซึ่งเป็นค่าของ *Window size* และจำนวนนิวรอนในแต่ละเลเยอร์



รูปที่ 45 โครงสร้างโครงโมไซมสำหรับกรณีศึกษา

4. การคัดเลือก (selection)

การคัดเลือกใช้วิธีการที่เรียกว่าการคัดเลือกแบบวงล้อรูเล็ต (Roulette wheel selection) เพื่อคัดเลือกประชากรที่มีค่าความเหมาะสมสูงแต่อาจทำให้ติดจุดสูงสุดปลอม (local optimum) แก้โดยการเลือก k ให้เหมาะสมและเพิ่มการสุ่มตอนทำการกลายพันธุ์

5. การสลับสายพันธุ์ (crossover) เพื่อเพิ่มความหลากหลายของประชากร

6. การกลายพันธุ์ (mutation) ใช้วิธีการ *Randomly Swapping* เพื่อเพิ่มการสุ่มในการค้นหาคำตอบเพื่อไม่ให้ติด local optimum

ไลบรารีของ GA ที่ใช้สำหรับกรณีศึกษานี้คือ *Distributed Evolutionary Algorithms in Python (DEAP)* ซึ่งเป็น *python package* สำหรับเฟรมเวิร์กการคำนวณเชิงวิวัฒนาการ (*evolutionary computation framework*) ซึ่งส่วนใหญ่มักถูกใช้เชิงของการพัฒนาต้นแบบและทดสอบแนวความคิดอย่างรวดเร็ว [32]

จุฬาลงกรณ์มหาวิทยาลัย

3.2.4 โครงสร้างเน็ตเวิร์ก

1. Fully Connected Layer (FCL)

FCL เป็นเน็ตเวิร์กที่ใช้ในการรู้จำรูปแบบทั่วไปแต่อย่างไรก็ตามมันก็มีข้อจำกัดเนื่องจากมันไม่สามารถที่จะจดจำสถานะก่อนหน้าได้ ข้อจำกัดนี้เป็นปัญหาหลักสำหรับข้อมูลที่มีลักษณะเป็นลำดับเวลา (time series) ยิ่งไปกว่านั้นด้วยข้อมูลอนุกรมเวลาชุดเดียวกัน เน็ตเวิร์ก FCL ไม่สามารถเรียนรู้จากข้อมูลนี้ได้ดีเท่ากับเน็ตเวิร์กรีเคอร์ซีฟ (RNN) แม้ว่าเราจะเพิ่มความลึกของเน็ตเวิร์กก็ตาม เนื่องจากว่ามันไม่มีความสามารถในการจดจำสถานะก่อนหน้าเป็นผลให้อินพุตมีลักษณะคงที่ (fixed-size input) ในขณะที่เน็ตเวิร์กรีเคอร์ซีฟสามารถที่จะปรับปริมาณการเรียนรู้ในแต่ละรอบ (cycle) ของอินพุตได้ทำให้มันสามารถที่จะเรียนรู้ได้มากกว่า เพื่อที่จะประยุกต์ใช้เน็ตเวิร์ก FCL กับข้อมูลที่มีลำดับเวลาเราจะเพิ่มความต่อเนื่องทางเวลาในลักษณะเดียวกับที่เน็ตเวิร์กรีเคอร์ซีฟทำ นั่นคือการสร้างเวกเตอร์การมอยย้อนกลับเทียม (artificial look back) ดังแสดงในรูปที่ 44 ยิ่งไปกว่านั้นการ

เพิ่มตัวแปรทางเวลา เช่น hour-of-day และ day-of-year ความลึก และความใหญ่ของเน็ตเวิร์ก จะช่วยให้เน็ตเวิร์กทำงานกับข้อมูลที่เป็นลำดับเวลาได้ดีขึ้น

อย่างไรก็ตามสำหรับข้อมูลลำดับเวลาที่สั้นๆ (น้อยกว่า 100 time-steps) ไม่จำเป็นต้องสร้างตัวแปรทางเวลาขึ้นมาเนื่องจากลำดับที่สั้นเน็ตเวิร์กทั้งสองจะมีประสิทธิภาพพอกัน ยิ่งไปกว่านั้นมันขึ้นอยู่กับปัจจัยอื่นๆด้วย เช่น โครงสร้างสถาปัตยกรรมของเน็ตเวิร์ก เวลาที่ใช้ในการฝึกสอน การกำหนดค่าน้ำหนักเริ่มต้นและพารามิเตอร์ต่างๆ

ดังนั้นจึงมีการประยุกต์ใช้ GA เพื่อช่วยในการหาโครงสร้างที่เหมาะสมของเน็ตเวิร์ก โดยโครงสร้างของเน็ตเวิร์ก FCL ที่ได้จาก GA แสดงในตารางที่ 8

ตารางที่ 8 โครงสร้างของเน็ตเวิร์กประเภทต่างๆที่ได้จากขั้นตอนวิธีเชิงพันธุกรรม

Model	T	I	H1	H2	H3	O	RMSE (%)
FCL	33	23	70	116	130	8	8.88
GRU	41	$23 \times T$	72	61	251	8	7.83
LSTM	62	$23 \times T$	250	24	443	8	7.92
CuDNNGRU	35	$23 \times T$	46	30	15	8	7.87

โดยที่ T คือขนาดความกว้างของเวลา (window size) , $H\#$ คือจำนวนนิวรอนในชั้นซ่อน, I คือขนาดของอินพุต (input shape), และ O คือขนาดของเอาต์พุต (output shape)

2. ความแตกต่างระหว่างเน็ตเวิร์ก GRU และ LSTM

สำหรับข้อมูลที่เป็นลำดับเวลาเน็ตเวิร์กกรีเคอร์ซีฟมีประสิทธิภาพดีกว่าเน็ตเวิร์ก FCL ตามที่กล่าวในข้อ 1 ดังนั้นในหัวข้อนี้จะความแตกต่างระหว่างเน็ตเวิร์ก GRU และ LSTM ซึ่งสามารถยกตัวอย่างได้ดังต่อไปนี้

1. GRU มีประสิทธิภาพการคำนวณที่ดีกว่า LSTM เนื่องจากมันไม่มีเซลล์หน่วยความจำ (memory cell) ทำให้มันมีโครงสร้างที่ซับซ้อนน้อยกว่า พารามิเตอร์น้อยกว่า กล่าวคือ GRU มี 2 เกต (Reset gates (r) และ update gates (z)) ในขณะที่ LSTM มี 3 เกต (input gate (i), output gate (o) และ forget gate (f))
2. GRU สามารถฝึกสอนได้เร็วกว่า LSTM
3. GRU ต้องการข้อมูลน้อยกว่าเพื่อ generalize เนื่องจากมันมีพารามิเตอร์น้อยกว่า (U และ W เล็กกว่า) ทำให้การคำนวณง่ายกว่า

4. LSTM อาจให้ประสิทธิภาพที่ดีกว่า ถ้าปริมาณของข้อมูลมากระดับหนึ่ง

ต่อมาจะกล่าวถึงข้อได้เปรียบของ LSTM เหนือ GRU สำหรับงานที่ต้องการแบบจำลองที่สามารถดำเนินการกับข้อมูลที่เป็นลำดับเวลายาวได้ดี (long-distance relations) ต้องการความสัมพันธ์ของข้อมูลในอดีตที่ยาวขึ้น มองกลับไปในอดีตมากขึ้น แล้ว LSTM จะมีประสิทธิภาพดีกว่า GRU เนื่องจาก LSTM ถูกออกแบบมาให้สามารถจดจำลำดับความสัมพันธ์ของข้อมูลที่ยาวได้ อย่างไรก็ตามนั่นคือสาเหตุที่มันมีโครงสร้างที่ซับซ้อน (มีพารามิเตอร์จำนวนมาก) ทำให้ LSTM ใช้เวลาในการฝึกสอนที่นานกว่าและใช้ memory bandwidth ในการคำนวณค่อนข้างมาก

3.2.5 พารามิเตอร์ที่ใช้ในการฝึกสอน

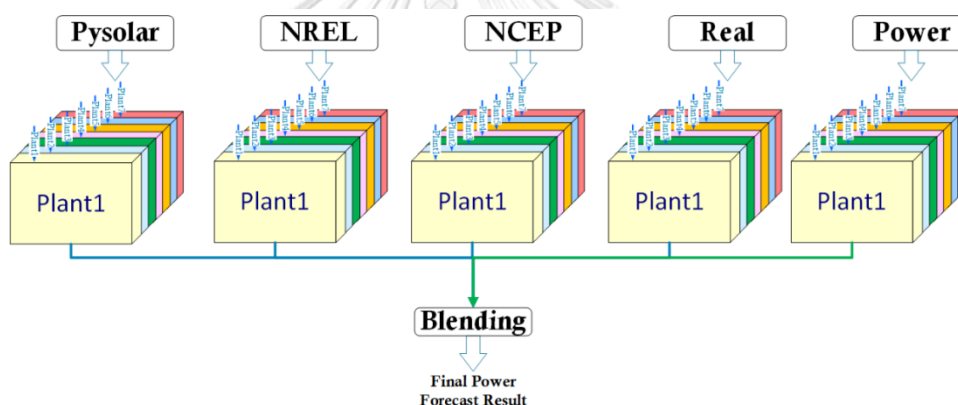
จำนวนของตัวอย่างต่อ Mini-batch หรือที่เรียกว่า batch size คือ 32 และทดสอบทั้งหมด 50 epochs ซึ่งคือจำนวนรอบที่มากที่สุดที่อัลกอริทึมที่ใช้ในการฝึกสอนวนผ่านทุกๆตัวอย่าง

1. ฟังก์ชันกระตุ้นคือฟังก์ชันเส้นตรงที่ถูกปรับแก้ Rectified
2. ออปติไมเซอร์คือ Root Mean Square Prop (RMSProp) [33]
3. ฟังก์ชันสูญเสียคือค่าเฉลี่ยความผิดพลาดกำลังสอง
4. การหยุดก่อน (early Stopping ; ถ้า validation loss ไม่ดีขึ้น 5 epoch ติดต่อกันให้หยุดการฝึกสอน)

3.3 การพยากรณ์กำลังผลิตไฟฟ้ารายภาคโดยใช้นิเวรอนเน็ตเวิร์ก ร่วมกับการปรับจูนไฮเปอร์พารามิเตอร์กรณีโรงไฟฟ้า 7 แห่ง (Neural Network based on GA and Ensemble for RSPF)

กรณีศึกษาเป็นการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ด้วยนิเวรอนเน็ตเวิร์กซึ่งแบ่งออกเป็น 4 ขั้นตอนตามรูปที่ 46 และรูปที่ 47 ได้แก่

1. ขั้นตอนการเตรียมข้อมูล (แก้ไขและลบข้อมูลที่ไม่ถูกต้องออกจากชุดข้อมูลฝึกสอน)
 - 1.1 การจัดกลุ่มข้อมูล
2. ขั้นตอนการฝึกสอนแบบจำลองด้วยนิเวรอนเน็ตเวิร์ก
 - 2.1 การค้นหาไฮเปอร์พารามิเตอร์แบบจำลองด้วย GA
3. ขั้นตอนการรวมกำลังการผลิตจากโรงไฟฟ้าทั้ง 7 แห่ง
4. ขั้นตอนการรวมกำลังผลิตรวมจากแต่ละแบบจำลอง



รูปที่ 46 ภาพรวมขั้นตอนการพยากรณ์รวม 7 โรงไฟฟ้า

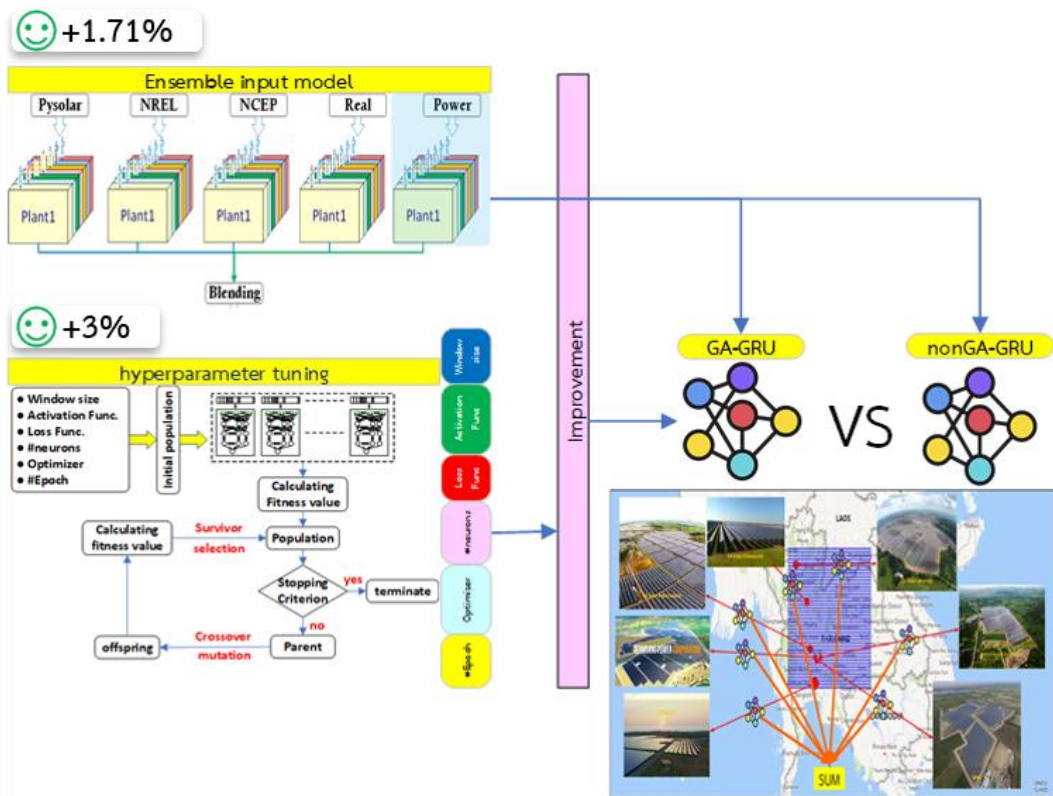
3.3.1 ข้อมูลอินพุต

3.3.1.1 ข้อมูลจากแบบจำลองพยากรณ์อากาศ

แบบจำลองพยากรณ์อากาศ 3 ประเภทถูกนำมาใช้เป็นข้อมูลอินพุตให้กับแบบจำลองพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ซึ่งแบบจำลองทั้ง 3 แบบจำลองดังแสดงในรูปที่ 48 ได้แก่

1. แบบจำลอง NCEP
2. แบบจำลอง Pysolar
3. แบบจำลอง NREL

และเอาต์พุตของแบบจำลองทั้ง 3 แสดงดังตารางที่ 9 ซึ่งสามารถดาวน์โหลดได้จากตารางที่ 10



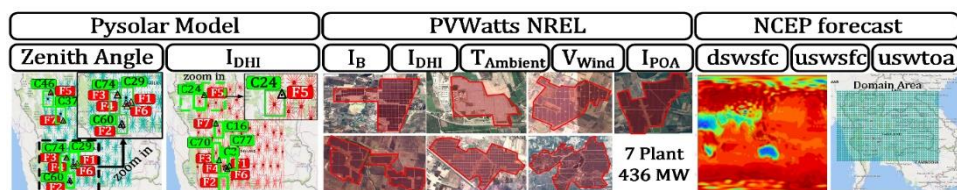
รูปที่ 47 ภาพรวมการพยากรณ์และการปรับปรุงประสิทธิภาพแบบจำลองด้วยวิธีที่สาม

ตารางที่ 9 ตัวแปรของแบบจำลองพยากรณ์อากาศทั้ง 3

Model	Variables
Pysolar	θ_z, DHI
PVWatts	$POA, DHI, T_{BackPanel}, T_{Ambient}$
NCEP	$dsw(sfc), usw(sfc), usw(toa)$

ตารางที่ 10 รายการแหล่งข้อมูลแบบจำลองพยากรณ์อากาศที่ให้บริการฟรี

Data source	
Model	Link
NCEP	https://cfs.ncep.noaa.gov/
Pysolar	https://pysolar.readthedocs.io/en/latest/
PVWatts	https://pvwatts.nrel.gov/



รูปที่ 48 แบบจำลองพยากรณ์อากาศประเภทต่างๆ

3.3.1.2 ข้อมูลค่าวัดจากอุปกรณ์ตรวจวัด

ข้อมูลค่าวัดเป็นค่าวัดของตัวแปรสภาพอากาศที่วัดได้บริเวณโรงไฟฟ้าซึ่งถูกนำมาใช้เป็นข้อมูลอินพุตให้กับแบบจำลองพยากรณ์กำลังผลิตไฟฟ้าซึ่งประกอบไปด้วย 5 ตัวแปร ได้แก่

1. Diffuse Horizontal Irradiance (DHI)
2. Global Horizontal Irradiance (GHI)
3. Plan of Array (POA)
4. Ambient temperature at the array height (T_H)
5. Back Panel Temperature (T_B)

หรือสามารถดาวน์โหลดได้จาก <https://www.wunderground.com> และตัวแปรข้อมูลค่าวัดจริงของแต่ละโรงไฟฟ้าแสดงในตารางที่ 11

ตารางที่ 11 ตัวแปรต่างๆที่ได้จากอุปกรณ์ตรวจวัดที่ถูกวัดในแต่ละโรงไฟฟ้า

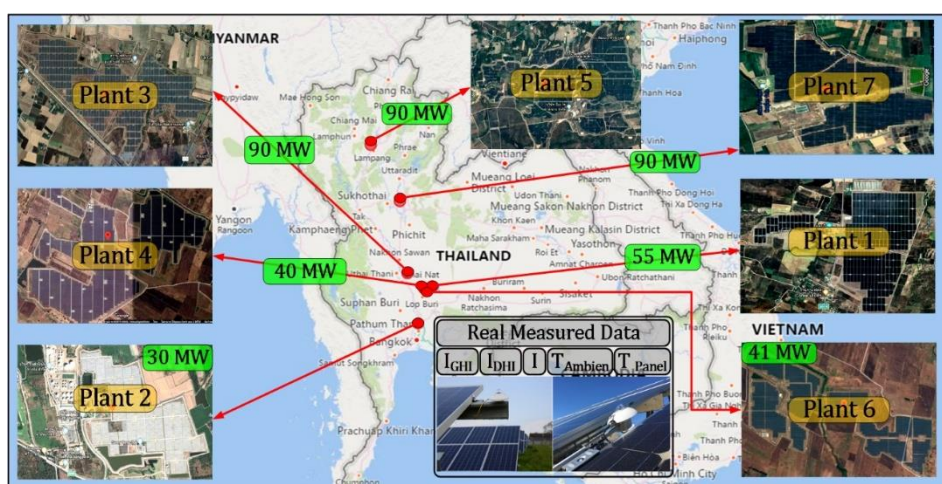
Variables	Plant						
	1	2	3	4	5	6	7
DHI	✓			✓		✓	
GHI			✓				✓
Irradiation		✓					
Amb.Temp	✓		✓	✓		✓	
BP.Temp	✓		✓	✓		✓	
MW	✓	✓	✓	✓	✓	✓	✓

3.3.1.3 ข้อมูลกำลังผลิตไฟฟ้า

ข้อมูลกำลังผลิตไฟฟ้าเป็นข้อมูลเป้าหมาย (label) ที่ต้องการพยากรณ์ซึ่งมีหน่วยเป็น MW แทนด้วยตัวแปร P ซึ่งมาจากอุปกรณ์ตรวจวัดหรือสามารถดาวน์โหลดได้จากเว็บไซต์ <http://control.egat.co.th/renewpattern> สำหรับในส่วนของการกำลังผลิตไฟฟ้าและตำแหน่งที่ตั้งของแต่ละโรงไฟฟ้าแสดงดังตารางที่ 12 ซึ่งโรงไฟฟ้าพลังงานแสงอาทิตย์สำหรับกรณีศึกษาที่นำมาใช้เพื่อพยากรณ์กำลังผลิตรวมประกอบไปด้วย 7 โรงไฟฟ้า

ตารางที่ 12 กำลังผลิตติดตั้งและตำแหน่งที่ตั้งของโรงไฟฟ้าทั้ง 7 โรง

Name	Location		Capacity (MW)	
	Latitude	longitude	Installed	Contracted
Plant 1	15.053	100.889	73	55
Plant 2	14.169	100.551	34.2	30
Plant 3	15.367	100.298	126.13	90
Plant 4	15.012	100.694	52	40
Plant 5	18.392	99.419	128.4	90
Plant 6	15.000	100.715	41	41
Plant 7	17.071	100.116	133.92	90
Total Production Capacity (MW)			588.64	436

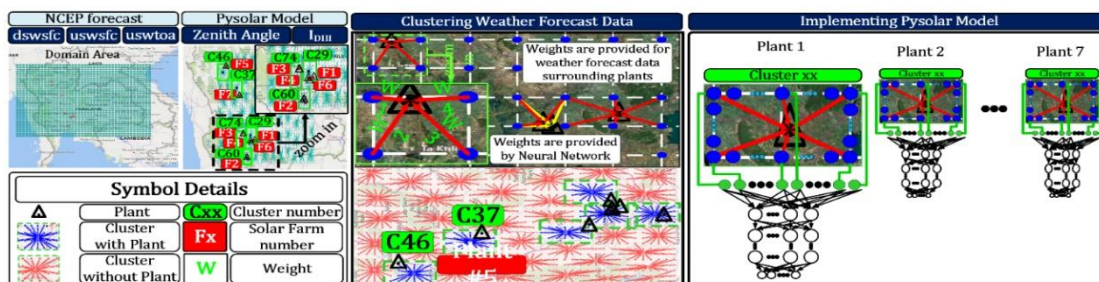


รูปที่ 49 กำลังผลิตไฟฟ้าของแต่ละโรงไฟฟ้าในแต่ละพื้นที่

3.3.1.4 การจัดกลุ่มข้อมูล

แบบจำลอง Pysolar และ NCEP ให้ค่าพยากรณ์ในลักษณะจุดกริด (Grid Point) อ้างอิงกับตำแหน่งละติจูดและลองจิจูด การพยากรณ์โดยทั่วไปจะใช้ค่าพยากรณ์อากาศ ณ ตำแหน่งที่ใกล้ที่สุดเพียงจุดเดียวซึ่งจะทำให้แบบจำลองมีค่าความเอนเอียง (bias) ที่ค่อนข้างสูงดังนั้นเพื่อที่จะลดค่าความเอนเอียงที่เกิดขึ้นนี้ การกรณศึกษาจึงใช้ค่าพยากรณ์อากาศในลักษณะเชิงพื้นที่โดยมีวิธีการในการจัดกลุ่มข้อมูลดังต่อไปนี้

1. คัดเลือกตัวแปรสำหรับการจัดกลุ่มข้อมูล
2. จัดกลุ่มข้อมูลด้วยเคมีน (k-Means Clustering) โดยกำหนดค่า K เป็น 78 ซึ่งกลุ่มข้อมูลของโรงไฟฟ้าแต่ละแห่งแสดงดังตารางที่ 13
3. เมื่อแบ่งกลุ่มเสร็จแล้วให้นำสมาชิกแต่ละตัวในกลุ่มเข้าไปยังเน็ตเวิร์กเพื่อหาค่าน้ำหนักที่เหมาะสมสำหรับสมาชิกแต่ละตัวในกลุ่มดังแสดงในรูปที่ 50



รูปที่ 50 การจัดกลุ่มข้อมูลด้วยเคมีนสำหรับกรณีศึกษา

ตารางที่ 13 หมายกลุ่มข้อมูลของโรงไฟฟ้าแต่ละโรง

Plants	Cluster Number				
	Pysolar		NCEP		
	DHI	θ_z	dswsf	uswsf	Uswtoa
Plant1	77	29	46	2	10
Plant2	58	60	75	19	40
Plant3	70	74	46	2	66
Plant4	3	29	46	2	10
Plant5	24	46	52	44	64
Plant6	3	29	46	2	10
Plant7	16	37	3	50	33

3.3.2 พารามิเตอร์ของเน็ตเวิร์ก

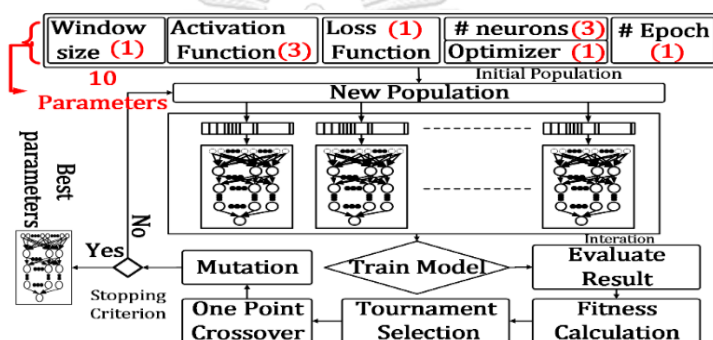
การค้นหาไฮเปอร์พารามิเตอร์ของเน็ตเวิร์ก โดยทั่วไปโครงสร้างเน็ตเวิร์กจะถูกออกแบบผ่านการลองผิดลองถูกซึ่งใช้เวลานาน ปัจจุบันคอมพิวเตอร์มีความสามารถในการคำนวณสูงขึ้นทำให้การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) และขั้นตอนวิธีเชิงวิวัฒนาการ (Evolutionary Algorithms) ถูกนำมาใช้อย่างแพร่หลายมากขึ้นเพื่อค้นหาพารามิเตอร์ที่เหมาะสมสำหรับกรณีศึกษา นี้เราประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรม (GA) ในการค้นหาไฮเปอร์พารามิเตอร์ดังแสดงในรูปที่ 51 ซึ่งยีนในโครโมโซมแสดงถึงพารามิเตอร์ของแบบจำลอง 10 ตัวได้แก่

1. วินโดวไซส์ (window size)
2. จำนวนนิวรอนในชั้นซ่อน (neuron) (*3)
3. ฟังก์ชันกระตุ้น (activation function) (*3)
4. ฟังก์ชันสูญเสีย
5. ออปติไมเซอร์ (optimizer)
6. จำนวนอีพอส (epoch)

ดังนั้นแต่ละโครโมโซมแสดงถึงแบบจำลองที่มีไฮเปอร์พารามิเตอร์ที่แตกต่างกันออกไปเนื่องจากขั้นตอนการสุ่มเริ่มต้น โดยพารามิเตอร์ของแบบจำลองที่ได้จาก GA แสดงในตารางที่ 14

ตารางที่ 14 โครงสร้างของเน็ตเวิร์กที่ได้จาก GA

Parameters	RNN with GA	RNN without GA
Window size	32	8
Number of neurons	278-303-142	512-512-512
Activation Function	Relu - Relu - Relu	Relu - Relu - Relu
Loss Function	MAE	MSE
Optimizer	Adam	RMSprop
Number of Epochs	27	30

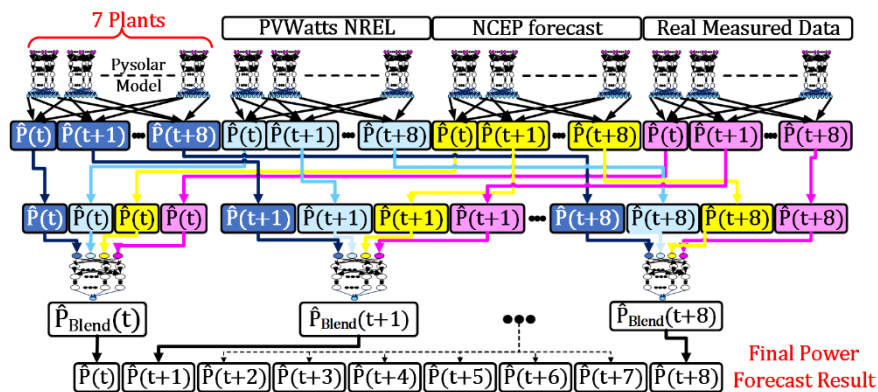


รูปที่ 51 การค้นหาไฮเปอร์พารามิเตอร์ของเน็ตเวิร์กด้วยขั้นตอนวิธีเชิงพันธุกรรม

3.3.3 ขั้นตอนการรวมกำลังผลิตรวม (ผลรวม 7 โมง) ของแต่ละแบบจำลอง

ขั้นตอนนี้เป็นการรวมกำลังผลิตรวมจากแต่ละแบบจำลองด้วยวิธีการเรียนรู้แบบรวมกลุ่ม (Ensemble Learning) เนื่องจากทั้ง 3 แบบจำลองพยากรณ์อากาศและ 1 ข้อมูลค่าวัดจริงมีจุดเด่นในการพยากรณ์ที่ต่างกัน บางแบบจำลองมีจุดเด่นในการพยากรณ์ที่กรอบเวลาสั้นๆ เช่น 15 นาที ถึง 3 ชั่วโมง บางแบบจำลองมีจุดเด่นในการพยากรณ์ที่กรอบเวลา 1 วัน ดังนั้นการการเรียนรู้แบบรวมกลุ่มจะช่วยลดค่าความเอนเอียง (Bias) ที่มีอยู่ในแต่ละแบบจำลองลง

เนื่องจากแต่ละแบบจำลองให้เอาต์พุตออกมาล่วงหน้า 8 ชั่วโมง หรือ 8 time-steps แนวคิดคือเราจะนำค่า ณ time-step เดียวกันจากแต่ละแบบจำลองมาเข้าเน็ตเวิร์กเพื่อหาค่าน้ำหนักที่เหมาะสมของแต่ละแบบจำลองดังแสดงในรูปที่ 52



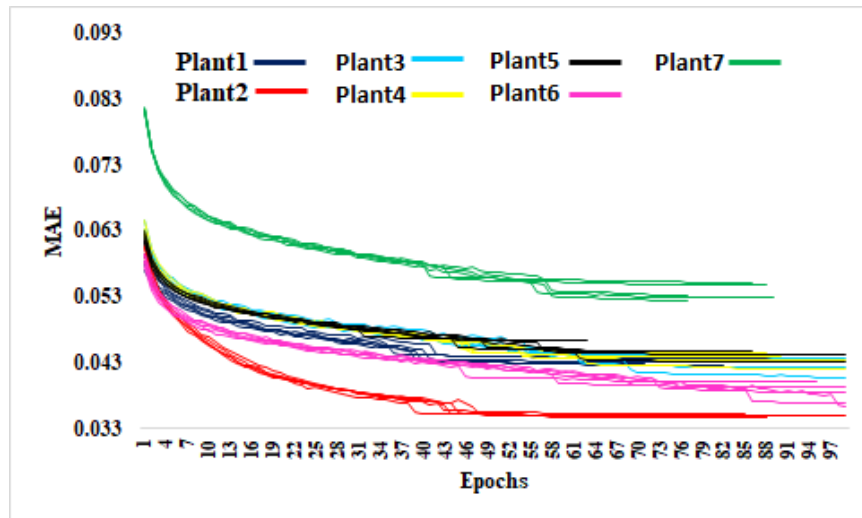
รูปที่ 52 ขั้นตอนการผสมผสานค่าพยากรณ์จากแต่ละแบบจำลอง

3.3.4 ขั้นตอนการวัดประสิทธิภาพ

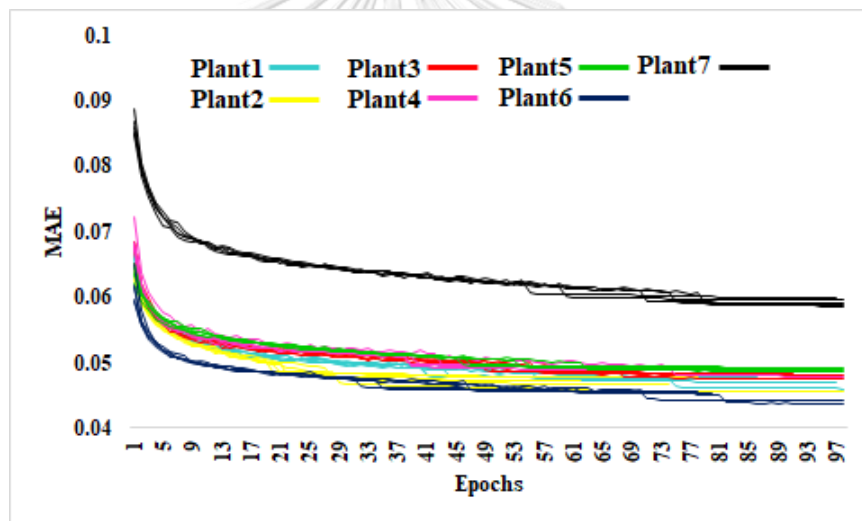
ข้อมูลที่ใช้สำหรับการประเมินผลของกรณีศึกษาเป็นข้อมูลตั้งแต่วันที่ 1 มกราคมถึงวันที่ 31 ธันวาคม 2017 โดยแบ่งเป็นข้อมูลที่ใช้ในการฝึกสอน (Training set) 70% ข้อมูลที่ใช้ในการตรวจสอบความถูกต้อง (validation set) 20% และ ข้อมูลที่ใช้ในการทดสอบ (testing set) 10% โดยดัชนีประสิทธิภาพที่ใช้ในการวัดผลคือค่า RMSE ดังแสดงใน (22)

$$RMSE = 100\% \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{P}(t) - \bar{P}(t))^2 / (Installed\ Capacity)} \quad (22)$$

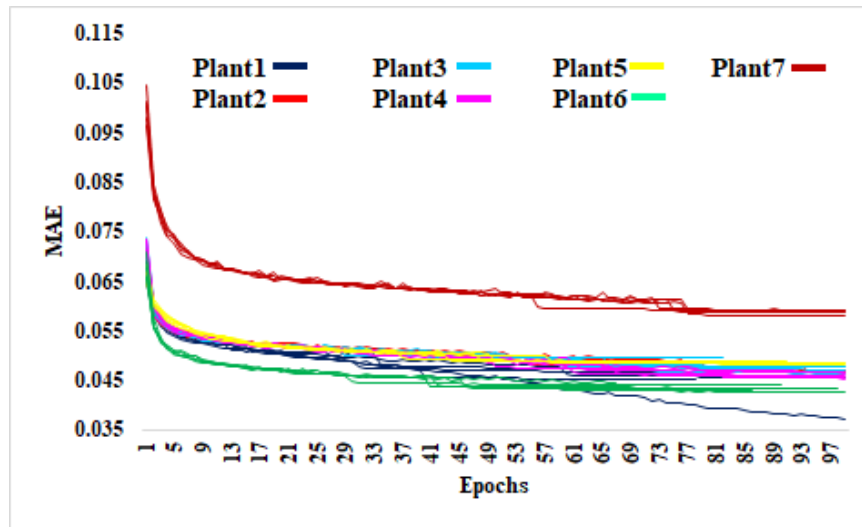
ในระหว่างการฝึกสอนมีการทดสอบประสิทธิภาพของแบบจำลองด้วยชุดข้อมูล validation การฝึกสอนจะหยุดลงเมื่อ validation loss ไม่ดีขึ้น 10 epochs ติดต่อกันซึ่งประสิทธิภาพของแบบจำลองพยากรณ์กำลังผลิตไฟฟ้าหรือค่า MAE ระหว่างการฝึกสอนในแต่ละแบบจำลองพยากรณ์อากาศแสดงดัง รูปที่ 53 รูปที่ 54 รูปที่ 55 และ รูปที่ 56



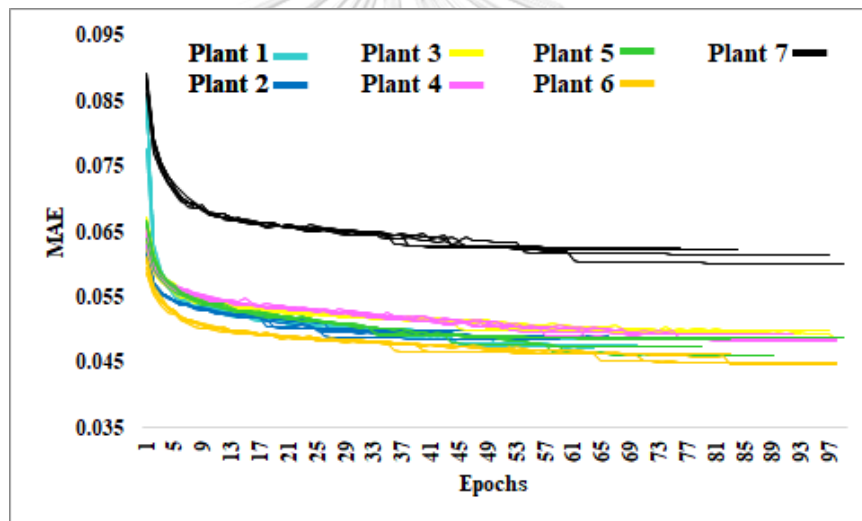
รูปที่ 53 ค่า MAE ของแบบจำลอง NCEP ในแต่ละโรงไฟฟ้า



รูปที่ 54 ค่า MAE ของแบบจำลอง NREL ในแต่ละโรงไฟฟ้า



รูปที่ 55 ค่า MAE ของแบบจำลอง Pysolar ในแต่ละโรงไฟฟ้า



รูปที่ 56 ค่า MAE ของแบบจำลอง measured data ในแต่ละโรงไฟฟ้า

บทที่ 4

ผลการวิจัย

4.1 แบบจำลอง Single Solar Power Forecast (SSPF)

แบบจำลอง SSPF เป็นแบบจำลองสำหรับการพยากรณ์กำลังผลิตไฟฟ้าโรงเดียว

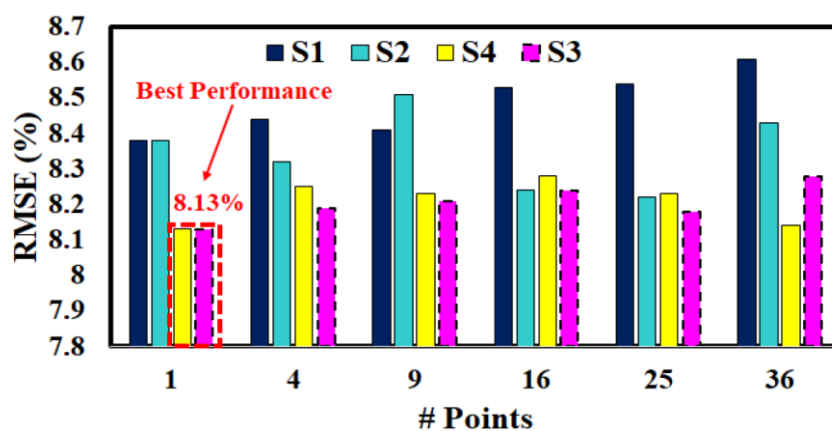
4.1.1 การพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ด้วยแบบจำลองการพยากรณ์อากาศเชิงตัวเลข (Very Short-Term SPF using data from NWP model)

4.1.1.1 การทดสอบบนเน็ตเวิร์ก FCL

กรณีศึกษานี้ชุดตัวแปรที่ใช้ทดสอบแบ่งออกเป็น 4 ชุดดังแสดงในตารางที่ 15 และทำการทดสอบทั้งหมด 6 ขนาดพื้นที่ ได้แก่ 1x1, 2x2, 3x3, 4x4, 5x5 และ 6x6 ผลลัพธ์ของการศึกษาแสดงดังรูปที่ 57 ซึ่งสามารถสรุปได้ว่าข้อมูลชุดที่ 3 ด้วยขนาดพื้นที่ 1x1 ให้ค่าความผิดพลาด (RMSE) น้อยที่สุดอยู่ที่ 8.13% ในขณะที่ชุดข้อมูลที่ 1 ด้วยขนาดพื้นที่ 1x1 เช่นเดียวกัน ให้ค่าความผิดพลาดอยู่ที่ 8.38% ซึ่งจะเห็นได้ว่าข้อมูลชุดที่ 3 มีค่าความผิดพลาดน้อยกว่าชุดที่ 1 อยู่ 0.25%

ตารางที่ 15 ชุดข้อมูล 4 ชุดที่ใช้ในการทดสอบประสิทธิภาพ

ชุดข้อมูล			
1	2	3	4
I_{wrf}	\bar{I}_{wrf}	$I_{wrf}, Rh_{wrf}, T_{wrf},$	$\bar{I}_{wrf}, \bar{Rh}_{wrf}, \bar{T}_{wrf},$
Rh_{wrf}	\bar{Rh}_{wrf}	$I_{ema},$	$I_{ema},$
T_{wrf}	\bar{T}_{wrf}	$I_{mea}, T_{mea}, P_{mea}$	$I_{mea}, T_{mea}, P_{mea}$

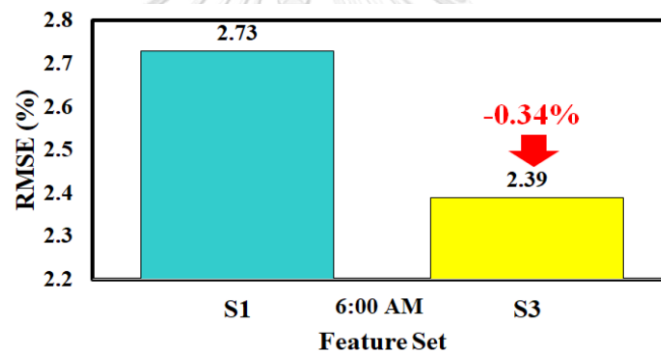


รูปที่ 57 เปรียบเทียบ RMSE ของ 4 ชุดข้อมูลในแต่ละขนาดพื้นที่

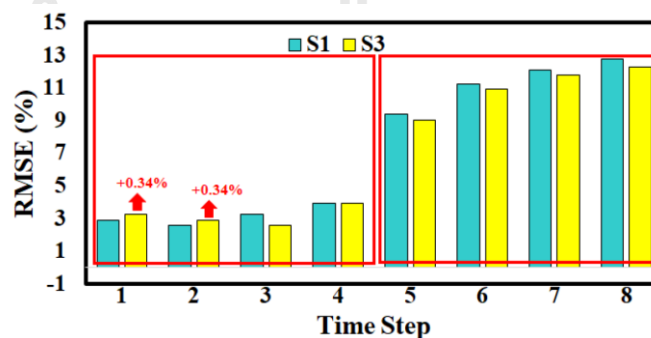
และจากรูปที่ 58 แสดงให้เห็นว่าข้อมูลชุดที่ 1 ให้ค่าความผิดพลาดที่ตอน 6 โมงเช้าอยู่ที่ 2.73% ซึ่งสูงกว่าข้อมูลชุดที่ 3 อยู่ 0.34% สำหรับพื้นที่ขนาด 1x1 กริดพอยด์ และจากรูปที่ 59 แสดงให้เห็นว่าข้อมูลชุดที่ 3 มีค่าความผิดพลาดสูงกว่าข้อมูลชุดที่ 1 ในสถานะเวลา (time-step) $t + 1$ และ $t + 2$ นอกนั้นมีค่าต่ำกว่าทั้งหมด

จากตารางที่ 16 ในสถานะเวลา $t + 3$ และ $t + 8$ แสดงให้เห็นว่าข้อมูลชุดที่ 3 ให้ค่าความผิดพลาดลดลงจากข้อมูลชุดที่ 1 อยู่ 0.68% และ 0.51% ตามลำดับในขณะที่สถานะเวลา $t + 1$ และ $t + 2$ ข้อมูลชุดที่ 3 ให้ค่าความผิดพลาดสูงกว่าข้อมูลชุดที่ 1 อยู่ 0.34% ทั้งสองสถานะเวลายิ่งไปกว่านั้นมันแสดงให้เห็นว่าค่าความผิดพลาดจะเพิ่มขึ้นอย่างเห็นได้ชัดในสถานะเวลาดังแต่ $t + 5$ เป็นต้นไปจนถึง $t + 8$

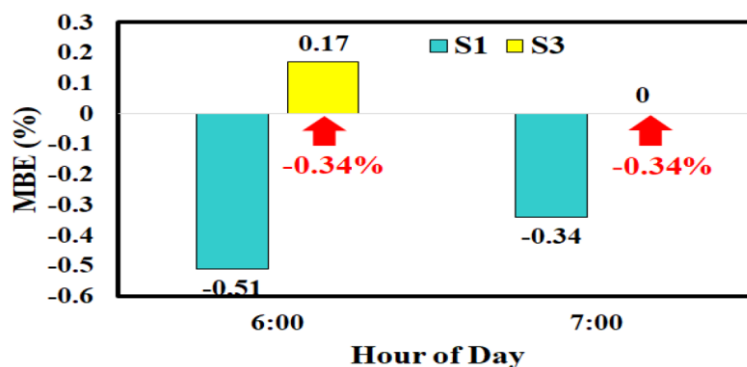
จากรูปที่ 60 จะเห็นได้ว่าค่าค่าคลาดเคลื่อนจากความเอนเอียงเฉลี่ย ณ เวลา 6 โมงเช้าและ 7 โมงเช้าของข้อมูลชุดที่ 3 มีค่าน้อยกว่าข้อมูลชุดที่ 1 อยู่ 0.34%



รูปที่ 58 เปรียบเทียบ RMSE ตอน 6 โมงเช้าระหว่างชุดข้อมูล 1 และ 3 (พื้นที่ขนาด 1x1)



รูปที่ 59 เปรียบเทียบ RMSE ระหว่างชุดข้อมูลที่ 1 และ 3 ในแต่ละ time-steps อนาคต



รูปที่ 60 เปรียบเทียบ MBE ระหว่างข้อมูลชุดที่ 1 และ 3 สำหรับขนาดพื้นที่ 1x1

ตารางที่ 16 การเปรียบเทียบ RMSE ระหว่างชุดข้อมูลที่ 1 และ 3 ในแต่ละ time-steps

ชุด	ความผิดพลาด (RMSE) ในแต่ละ time-steps							
	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8
1	2.9	2.56	3.24	3.92	9.37	11.24	12.1	12.78
3	3.24	2.9	2.56	3.92	9.03	10.9	11.75	12.27

จากผลการศึกษาพบว่า FCL เป็นแบบจำลองทั่วไปซึ่งถูกใช้เพื่อรู้จำรูปแบบต่างๆที่ไม่มีความซับซ้อนมากนัก เช่น ปัญหา Regression ทั่วไป เนื่องจากมันไม่มีความสามารถในการจดจำสถานะก่อนหน้าเหมือนเน็ตเวิร์ก RNN, GRU, หรือ LSTM ทำให้มันไม่สามารถทำงานกับข้อมูลที่มีลักษณะเป็นลำดับอนุกรมได้ดี แต่อย่างไรก็ตามเราสามารถปรับปรุงประสิทธิภาพของเน็ตเวิร์ก FCL ให้มีประสิทธิภาพเทียบเท่ากับเน็ตเวิร์กแบบรีเคอร์ซีฟได้ ยกตัวอย่างเช่น

1. การสร้างมิติเวลาให้กับข้อมูลอินพุตที่เข้ามา (มิติที่ 3 หรือมิติความลึก) ซึ่งคือสถานะเวลาย้อนกลับเทียบ (artificial lookback หรือ window size) ดังแสดงในรูปที่ 35
2. การเพิ่มตัวแปรทางเวลา เช่น นาทีของชั่วโมง(minute-of-hour), ชั่วโมงของวัน (hour-of-day) หรือ วันของปี (day-of-year) เป็นต้น
3. เพิ่มขนาดและความลึกของเน็ตเวิร์กเพื่อให้มันสามารถจดจำรูปแบบได้มากขึ้น

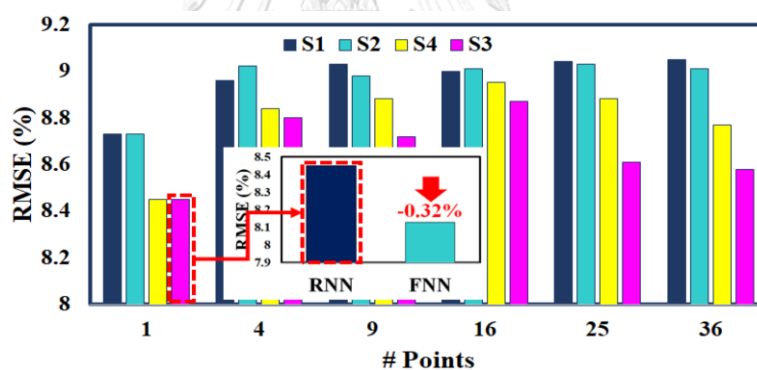
แต่สำหรับลำดับอนุกรมที่สั้น (น้อยกว่า 10 time-steps) เราอาจไม่จำเป็นต้องทำ 3 วิธีที่กล่าวมา เนื่องจากในลำดับอนุกรมที่สั้นเน็ตเวิร์ก FCL สามารถทำงานได้ดีพอๆกับเน็ตเวิร์กรีเคอร์ซีฟแต่อย่างไรก็ตามมันก็ขึ้นอยู่กับสถาปัตยกรรมของเน็ตเวิร์กด้วย เช่น เวลาที่ใช้ในการฝึกสอน การกำหนดค่าเริ่มต้น (initialize) ของค่าน้ำหนักและพารามิเตอร์อื่นๆ

4.1.1.2 การทดสอบบนเน็ตเวิร์กเรคิเคอร์ซีฟชนิด RNN

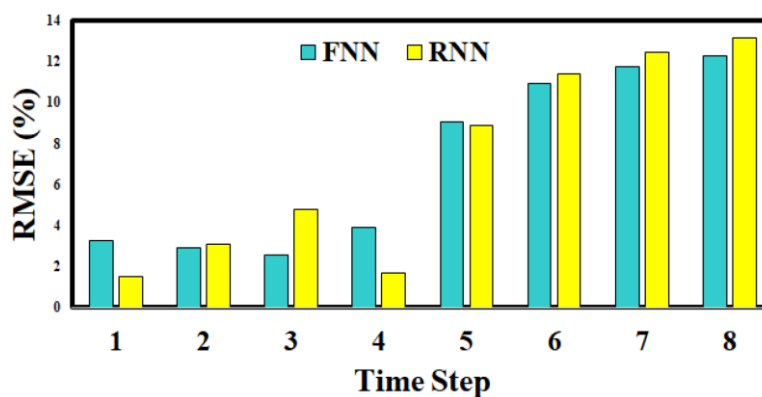
จากการศึกษาเน็ตเวิร์กเรคิเคอร์ซีฟชนิด RNN พบว่าข้อมูลชุดที่ 3 ให้ค่าความผิดพลาดอยู่ที่ 8.45% ซึ่งน้อยที่สุดเมื่อเทียบกับชุดข้อมูลอื่นๆ สำหรับพื้นที่ขนาด 1x1 ดังแสดงในรูปที่ 61 ซึ่งสูงกว่าค่าความผิดพลาดที่ได้จากเน็ตเวิร์ก FCL อยู่ 0.32%

จากรูปที่ 62 แสดงให้เห็นว่าแนวโน้มของค่าความผิดพลาดในแต่ละสถานะเวลา (time-steps) $t + 1$ ถึง $t + 8$ ของเน็ตเวิร์กเรคิเคอร์ซีฟ (RNN) มีลักษณะคล้ายกับเน็ตเวิร์ก FCL กล่าวคือค่าความผิดพลาดมีแนวโน้มเพิ่มสูงขึ้นใน time-steps ที่ไกลออกไป

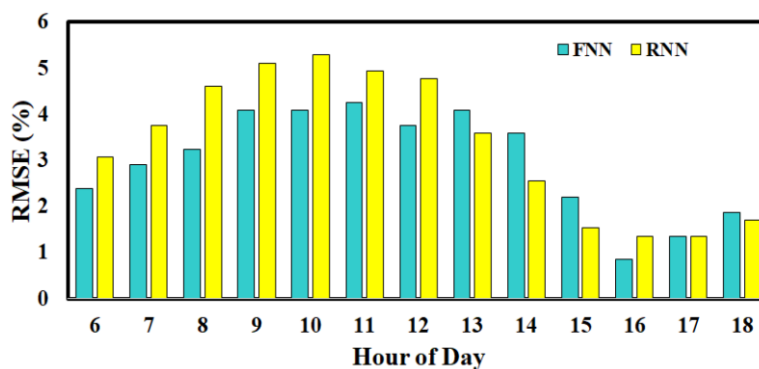
การเปรียบเทียบระหว่างเน็ตเวิร์กเรคิเคอร์ซีฟและเน็ตเวิร์ก FCL สำหรับข้อมูลชุดที่ 3 แสดงให้เห็นว่าค่าความผิดพลาดในพื้นที่ขนาด 1x1 ของเน็ตเวิร์ก FCL มีค่าน้อยกว่าเน็ตเวิร์กเรคิเคอร์ซีฟใน time-steps ที่ 2, 3 และ 5 ถึง 8 ยิ่งไปกว่านั้นทั้งสองเน็ตเวิร์กยังมีแนวโน้มของค่าความผิดพลาดที่คล้ายคลึงกัน จากรูปที่ 63 ค่าความผิดพลาดในพื้นที่ขนาด 1x1 ในช่วงเวลา 13:00-15:00 และ 18:00 pm ของเน็ตเวิร์ก FCL สูงกว่าเน็ตเวิร์กเรคิเคอร์ซีฟอยู่ 0.51, 1.02, 0.68, และ 0.17 เปอร์เซ็นต์ตามลำดับ



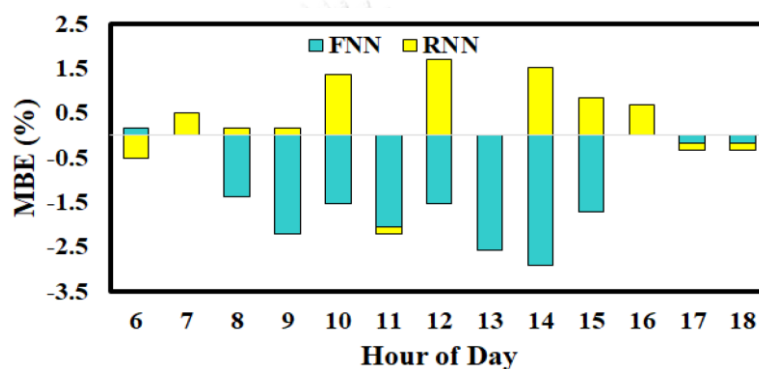
รูปที่ 61 เปรียบเทียบ RMSE (%) ของแต่ละชุดข้อมูลและแต่ละขนาดพื้นที่ระหว่าง RNN และ FCL



รูปที่ 62 RMSE (%) ของข้อมูลชุดที่ 3 สำหรับพื้นที่ขนาด 1x1 ในแต่ละ time-steps



รูปที่ 63 การเปรียบเทียบ RMSE ของเน็ตเวิร์กทั้งสอง (1x1) ในแต่ละช่วงเวลาของวัน

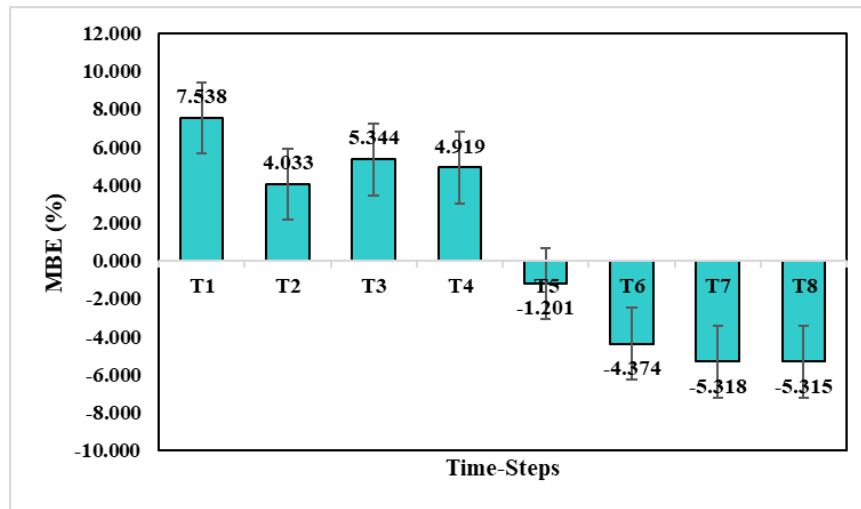


รูปที่ 64 การเปรียบเทียบ MBE (%) สำหรับ 1x1 ของข้อมูลชุดที่ 3 ตั้งแต่เวลา 6:00 - 18:00

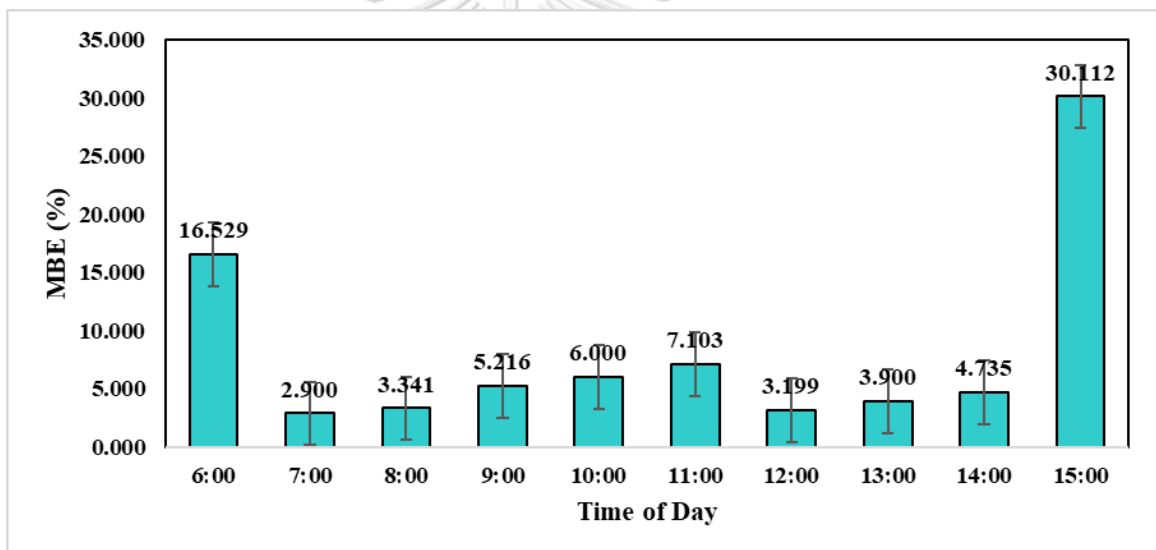
จากรูปที่ 64 แสดงให้เห็นว่าเน็ตเวิร์ก FCL ให้ค่าการพยากรณ์ที่ต่ำกว่าค่าจริงในขณะที่เน็ตเวิร์กเคอร์ซีฟให้ค่าการพยากรณ์ที่สูงกว่าค่าจริง

จากรูปที่ 69 และรูปที่ 71 เน็ตเวิร์กเคอร์ซีฟซึ่งข้อมูลอินพุตประกอบไปด้วย ข้อมูลพยากรณ์อากาศ ข้อมูลค่าวัดจริง และ ค่าเฉลี่ยเคลื่อนที่แบบฟังก์ชันเอกซ์โพเนนเชียลของความเข้มแสงอาทิตย์ที่วัดได้ ให้ค่าความผิดพลาดอยู่ที่ 8.45% ในขณะที่การใช้เน็ตเวิร์กเคอร์ซีฟด้วยค่าพยากรณ์อากาศเพียงอย่างเดียวจะให้ค่าความผิดพลาดอยู่ที่ 8.73% ซึ่งสูงกว่าการพยากรณ์อากาศด้วย ข้อมูลพยากรณ์อากาศ ข้อมูลค่าวัดจริง และ ค่าเฉลี่ยเคลื่อนที่แบบฟังก์ชันเอกซ์โพเนนเชียลของความเข้มแสงอาทิตย์ที่วัดได้อยู่เพียงแค่ 0.28%

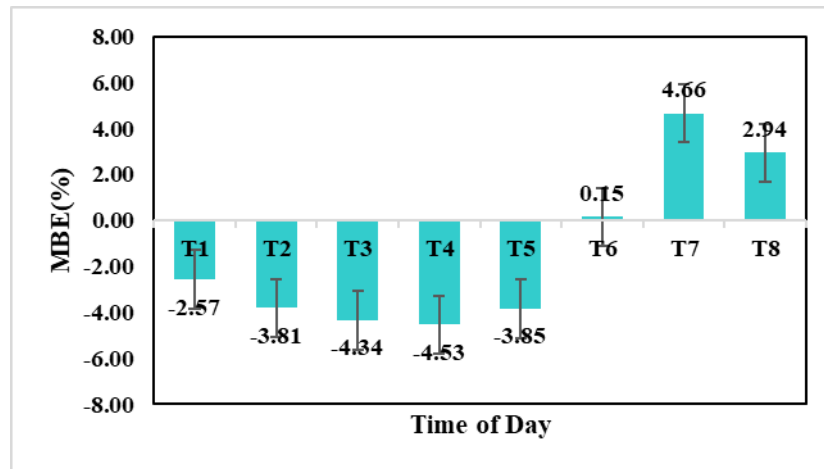
ค่า MBE(%) ในแต่ละ time-steps และแต่ละช่วงเวลาของวันของแบบจำลอง FCL แสดงดังรูปที่ 65 และรูปที่ 66 ตามลำดับ ในขณะที่แบบจำลอง RNN ค่า MBE(%) ในแต่ละ time-steps และแต่ละช่วงเวลาของวันแสดงดัง รูปที่ 67 และรูปที่ 68 ตามลำดับ



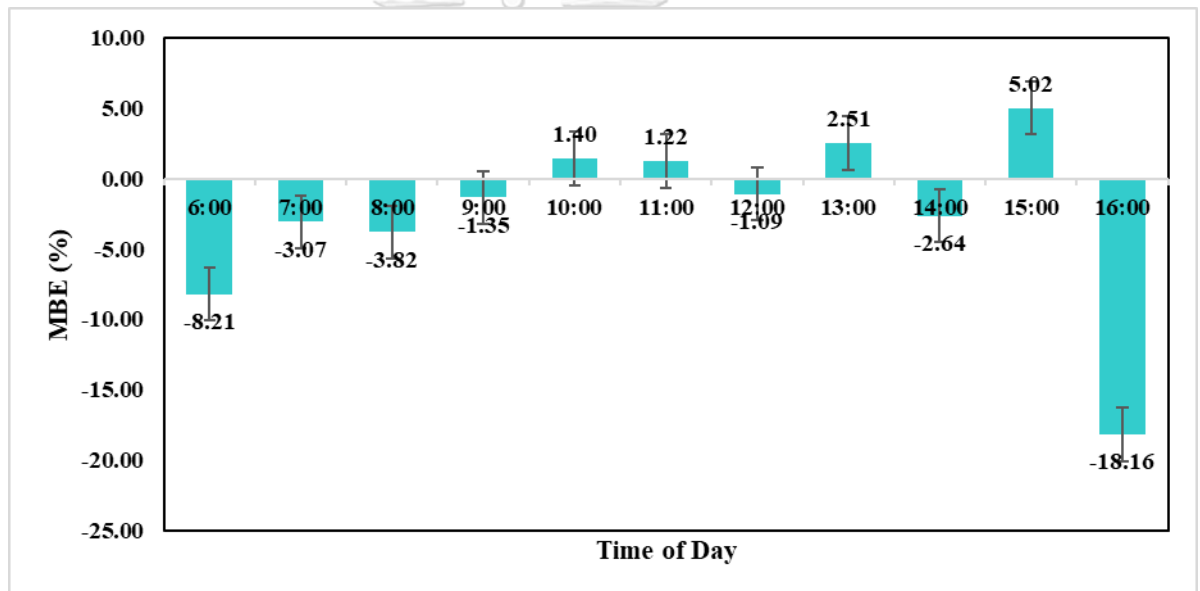
รูปที่ 65 ค่า MBE (%) ของแบบจำลอง FCL ในแต่ละ time-steps



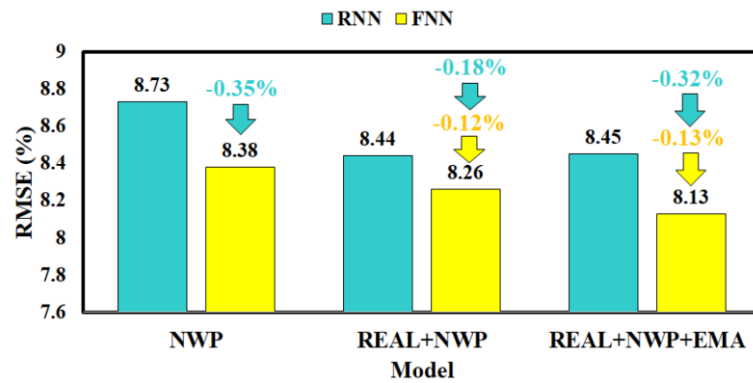
รูปที่ 66 ค่า MBE (%) ของแบบจำลอง FCL ในแต่ละช่วงเวลาของวัน



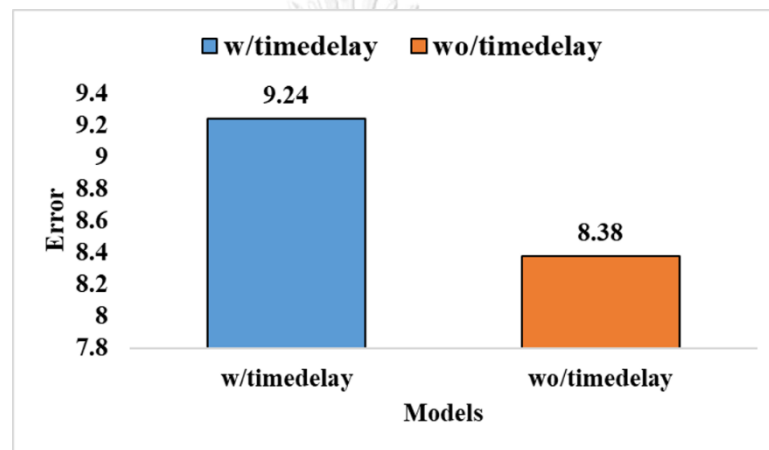
รูปที่ 67 ค่า MBE (%) ของแบบจำลอง RNN ในแต่ละ time-steps



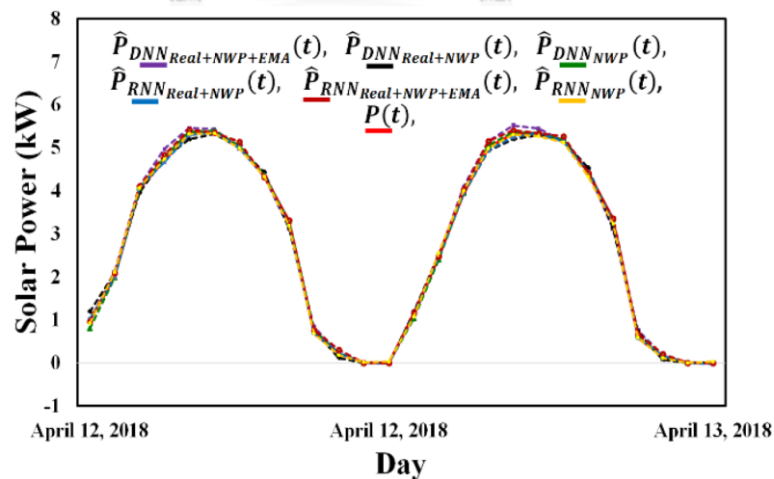
รูปที่ 68 ค่า MBE (%) ของแบบจำลอง RNN ในแต่ละช่วงเวลาของวัน



รูปที่ 69 การเปรียบเทียบการปรับปรุงประสิทธิภาพของเน็ตเวิร์กทั้งสอง



รูปที่ 70 เปรียบเทียบ RMSE ของแบบจำลอง FCL ที่มีและไม่มี time-delay



รูปที่ 71 การเปรียบเทียบกำลังผลิตไฟฟ้าที่เป็นค่าจริงและค่าพยากรณ์ของแต่ละเน็ตเวิร์ก

4.1.2 การพยากรณ์กำลังผลิตไฟฟ้าระยะสั้นโดยใช้นิวรอนเน็ตเวิร์กและขั้นตอนวิธีเชิงพันธุกรรม (Very Short-Term SPF using GA Based Neural Network)

แบบจำลองนี้เป็นแบบจำลองสำหรับการพยากรณ์กำลังผลิตไฟฟ้าแบบโรงไฟฟ้าเดี่ยว เช่นเดียวกันกับหัวข้อที่ผ่านมา

1. เน็ตเวิร์ก GA-FCL

ในมุมมองของความเร็ว เน็ตเวิร์ก FCL สามารถฝึกสอนได้เร็วกว่าเน็ตเวิร์กกรีเคอร์ซีฟเนื่องจากมีพารามิเตอร์ที่น้อยกว่าในขณะที่ในแง่ของประสิทธิภาพเน็ตเวิร์ก FCL ให้ผลลัพธ์ที่แม่นยำกว่าเน็ตเวิร์กกรีเคอร์ซีฟซึ่งในทางทฤษฎีแล้วเน็ตเวิร์กกรีเคอร์ซีฟสามารถทำงานได้ดีกับข้อมูลที่เป็นลำดับเวลาในขณะที่เน็ตเวิร์ก FCL เหมาะกับงานอื่นมากกว่า

จากการทดลองกรณีศึกษาพบว่าค่าความผิดพลาดของเน็ตเวิร์ก GA-FCL อยู่ที่ 8.88% ดังแสดงในรูปที่ 72 ซึ่งค่าความผิดพลาดนี้มีแนวโน้มเพิ่มสูงขึ้นใน time-steps ที่ไกลออกไป และถ้าพิจารณาค่าความผิดพลาดในแต่ละช่วงเวลาของวันตามรูปที่ 73 แสดงให้เห็นได้ว่าค่าความผิดพลาดมีค่าสูงสุดที่ตอน 10 โมงเช้าของแต่ละวัน

2. เน็ตเวิร์ก GA-GRU

เน็ตเวิร์ก GA-GRU ให้ค่าความผิดพลาดอยู่ที่ 7.83% ดังแสดงในรูปที่ 72 ซึ่งแนวโน้มของค่าความผิดพลาดในแต่ละ time-steps คล้ายกับของ GA-FCL แต่ค่าความผิดพลาดในช่วงเวลาก่อนเที่ยงมีค่าต่ำกว่าเน็ตเวิร์ก GA-FCL อย่างเห็นได้ชัด จากรูปที่ 73 เราจะเห็นได้ว่าเน็ตเวิร์ก GA-GRU มีค่าความผิดพลาดต่ำกว่าความเป็นจริง (under prediction) ณ ตอน 6 โมงเช้าและสูงกว่าความเป็นจริง ตอนบ่ายโมง (over prediction) ในแต่ละวัน

3. เน็ตเวิร์ก GA-LSTM

เน็ตเวิร์ก GA-LSTM ให้ค่าความผิดพลาดอยู่ที่ 7.92% จากรูปที่ 74 จะเห็นได้ชัดเจนว่าเน็ตเวิร์กมีความแม่นยำในตอนเช้าโดยเฉพาะอย่างยิ่งที่ตอน 7 โมงและ 10 โมง หลังจากนั้นจะมีค่าความผิดพลาดที่สูงกว่าความเป็นจริง

4. เน็ตเวิร์ก GA-CuDNNGRU

เน็ตเวิร์ก GA-CuDNNGRU ให้ค่าความผิดพลาดอยู่ที่ 7.87% และฝึกสอนได้อย่างรวดเร็ว

ตารางที่ 17 ความผิดพลาดในแต่ละ time-steps ในแต่ละเน็ตเวิร์ก

Model	RMSE (%) of Time Steps							
	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8
DNN	4.60	5.11	5.28	6.13	10.05	11.41	11.75	12.2
GRU	1.70	1.53	1.53	2.56	8.69	10.73	11.58	12.10
LSTM	2.56	2.39	4.09	3.92	8.86	10.56	11.41	11.75
CuDNNGRU	1.70	2.04	1.53	1.87	8.69	10.90	11.75	12.27

ตารางที่ 18 ค่าความผิดพลาดในแต่ละช่วงเวลาของวันของแต่ละเน็ตเวิร์ก

Model	RMSE (%) of Hours												
	6	7	8	9	10	11	12	13	14	15	16	17	18
DNN	4.60	5.11	6.47	8.01	9.37	8.18	7.50	6.64	5.45	1.87	2.21	2.21	2.56
GRU	2.56	2.21	1.70	1.36	1.36	1.53	1.87	2.21	1.70	1.53	1.19	0.34	0.34
LSTM	3.24	3.24	3.41	3.41	3.58	3.58	3.92	3.41	2.90	2.21	1.02	1.02	1.19
CuDNNGRU	2.73	2.90	2.56	2.90	2.21	1.87	1.87	2.21	1.70	1.87	1.19	1.70	2.04

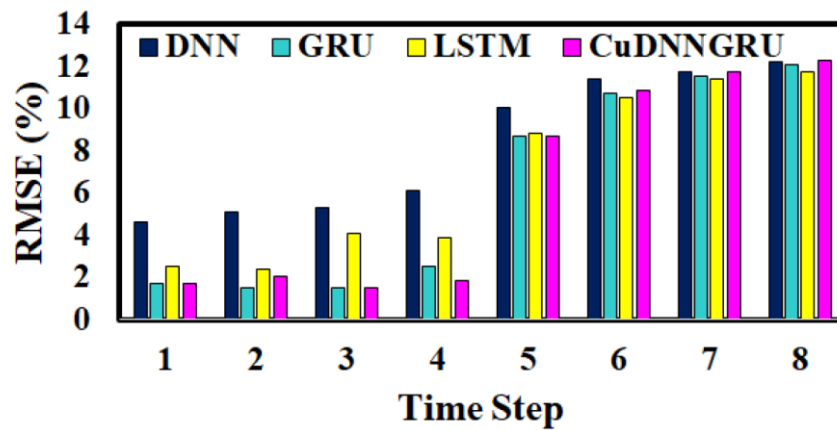
ตารางที่ 19 ค่า MBE ในแต่ละช่วงเวลาของวันของแต่ละเน็ตเวิร์ก

Model	RMSE (%) of Hours												
	6	7	8	9	10	11	12	13	14	15	16	17	18
DNN	-0.34	1.70	2.04	5.11	7.16	5.45	4.09	3.41	3.58	0.34	0.17	0.00	-0.00
GRU	-1.53	-0.85	-0.51	-0.51	0.51	1.19	1.36	1.53	1.36	0.85	0.68	0.00	-0.17
LSTM	-0.85	0.51	1.02	1.53	2.21	2.04	2.56	1.19	0.17	-0.51	-0.17	-0.51	-0.34
CuDNN GRU	-2.04	-2.39	-2.21	-2.21	-1.36	-0.85	-0.68	-1.02	-0.85	-1.36	-0.68	-1.36	-1.70

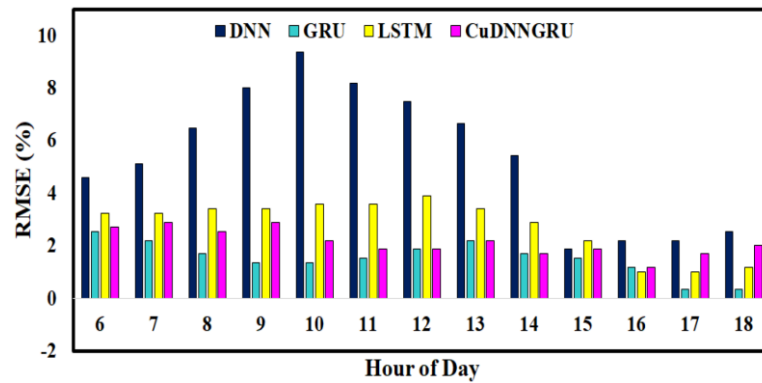
CHULALONGKORN UNIVERSITY

ตารางที่ 20 เวลาที่ใช้สำหรับการฝึกสอนของเน็ตเวิร์กแต่ละประเภท

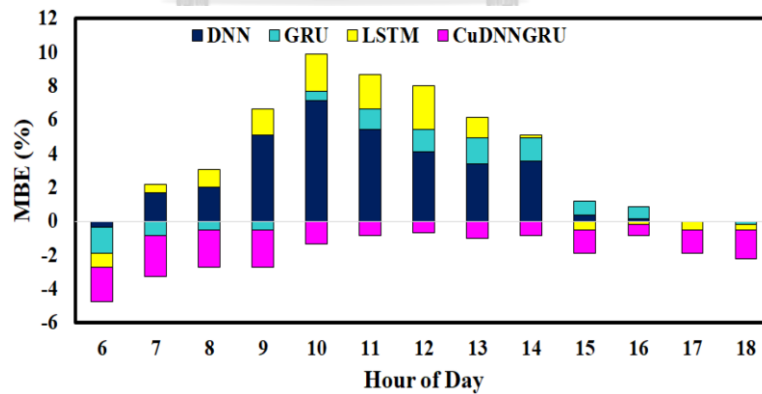
Model	DNN	GRU	LSTM	CuDNNGRU
Time (Sec)	42	126	603	51



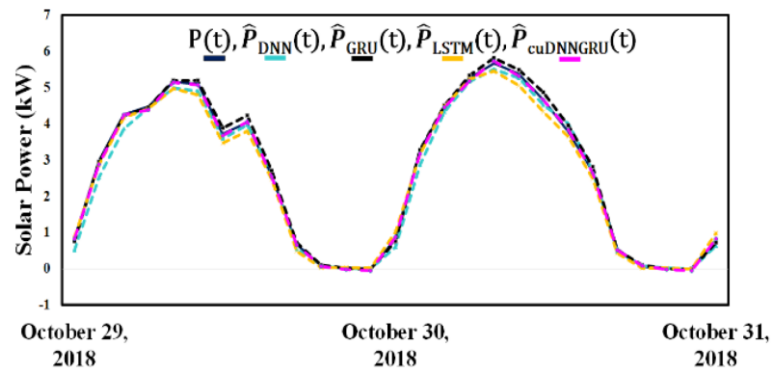
รูปที่ 72 ค่า RMSE ในแต่ละ time-steps ของเน็ตเวิร์กแต่ละประเภท



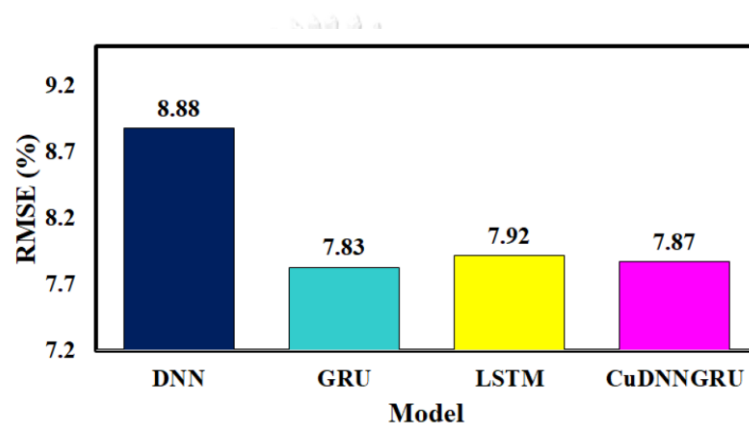
รูปที่ 73 ค่า RMSE ในแต่ละช่วงเวลาของวันของเน็ตเวิร์กแต่ละประเภท



รูปที่ 74 ค่า MBE ในแต่ละช่วงเวลาของวันของเน็ตเวิร์กแต่ละประเภท



รูปที่ 75 การเปรียบเทียบค่าจริงและค่าพยากรณ์กำลังผลิตไฟฟ้าจากเน็ตเวิร์กแต่ละประเภท



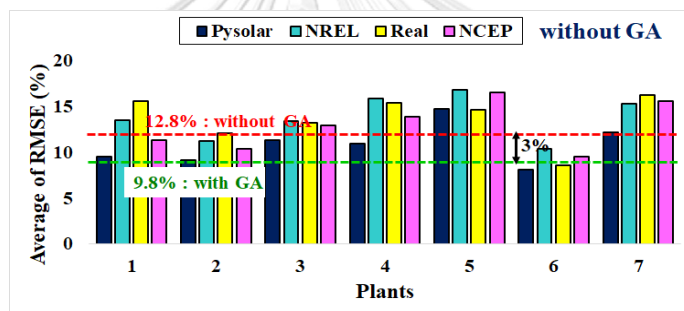
รูปที่ 76 การเปรียบเทียบค่า RMSE ของเน็ตเวิร์กแต่ละประเภท

4.2 แบบจำลอง Regional Solar Power Forecast (RSPF)

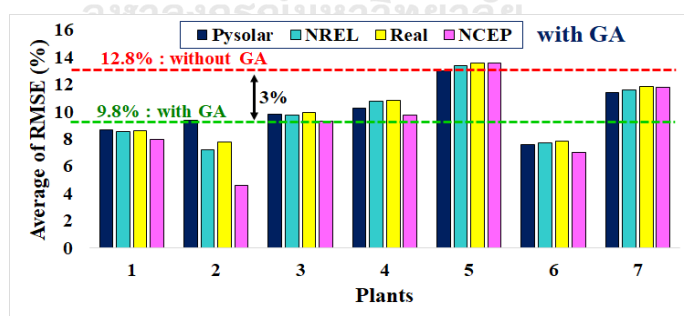
แบบจำลอง Regional Solar Power Forecast (SSPF) เป็นแบบจำลองที่สร้างขึ้นเพื่อทำการพยากรณ์กำลังผลิตไฟฟ้ารวมของโรงไฟฟ้า 7 แห่ง

4.2.1 ผลลัพธ์การค้นหาไฮเปอร์พารามิเตอร์ที่เหมาะสม

กรณีศึกษานี้ทำการศึกษาเพื่อเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองที่ไม่มีการค้นหาไฮเปอร์พารามิเตอร์กับแบบจำลองที่มีการค้นหาไฮเปอร์พารามิเตอร์ด้วยขั้นตอนวิธีเชิงพันธุกรรม ประสิทธิภาพของแบบจำลองที่ไม่มีการเลือกไฮเปอร์พารามิเตอร์ที่เหมาะสมแสดงดังรูปที่ 77 และประสิทธิภาพของแบบจำลองที่มีการคัดเลือกไฮเปอร์พารามิเตอร์ด้วย GA แสดงดังรูปที่ 78 ซึ่งจะเห็นได้ว่าประสิทธิภาพของแบบจำลองที่มีการค้นหาไฮเปอร์พารามิเตอร์จะมีประสิทธิภาพดีกว่าประมาณ 3 เปอร์เซ็นต์



รูปที่ 77 RMSE ของแบบจำลองที่ไม่มีการคัดเลือกไฮเปอร์พารามิเตอร์



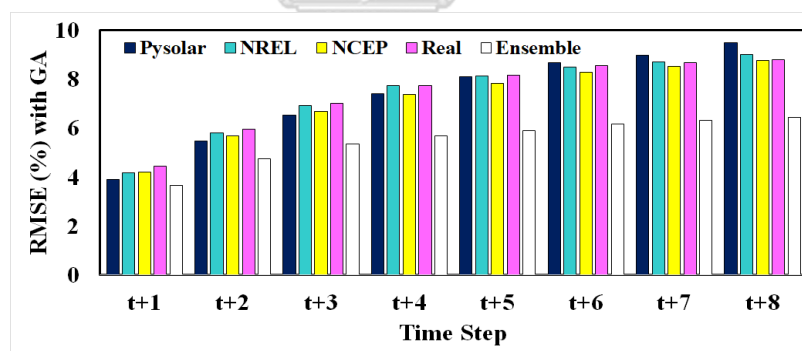
รูปที่ 78 RMSE ของแบบจำลองที่มีการคัดเลือกไฮเปอร์พารามิเตอร์ด้วย GA

4.2.2 การเรียนรู้แบบรวมกลุ่ม

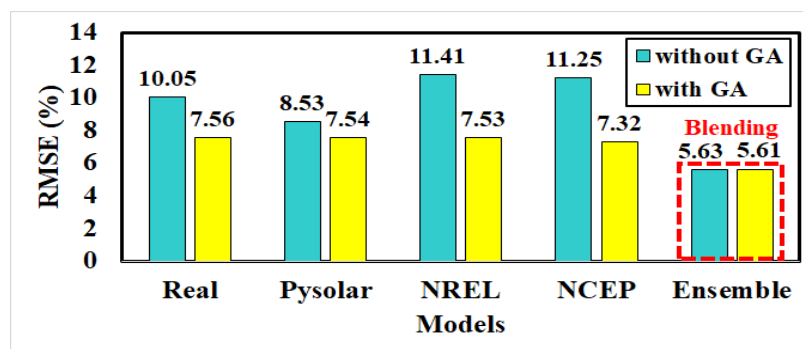
วิธีการนี้เป็นเทคนิคที่ช่วยหาความสมดุลระหว่างค่าความเอนเอียง (bias) และ ค่าความแปรปรวน (variation) ซึ่งเปอร์เซ็นต์ความผิดพลาด RMSE ของแบบจำลองแต่ละประเภทในแต่ละโรงไฟฟ้าแสดงดังรูปที่ 79 จากการผสมผสานผลลัพธ์การพยากรณ์จากแบบจำลองแต่ละประเภทเข้าด้วยกันพบว่าผลลัพธ์จากการพยากรณ์ด้วยวิธีการเรียนรู้แบบรวมกลุ่มช่วยปรับปรุงความแม่นยำในการพยากรณ์ดังแสดงในรูปที่ 80 และจากตารางที่ 21 แสดงค่าความผิดพลาดในแต่ละ time-Step โดยดัชนีที่ใช้ในการประเมินผลของการศึกษานี้คือค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error; RMSE) ตาม (8) และการเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์ของแบบจำลองแต่ละประเภทแสดงดังรูปที่ 81

ตารางที่ 21 เปอร์เซนต์ RMSE ในแต่ละ Time-step ของแบบจำลองแต่ละประเภท

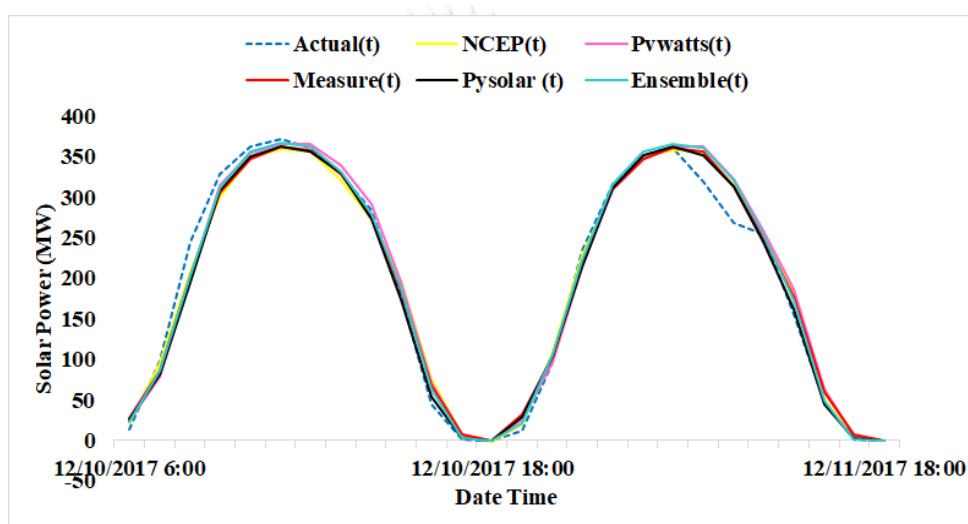
Models	Percentage of RMSE in each time-step							
	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8
	0.37	0.36	0.28	0.32	0.41	0.27	0.33	0.34
PVWatts	0.38	0.31	0.27	0.29	0.23	0.34	0.24	0.37
Measure	0.21	0.20	0.26	0.30	0.33	0.24	0.34	0.25
Pysolar	0.29	0.21	0.20	0.23	0.19	0.28	0.24	0.22
Blending	0.15	0.20	0.37	0.13	0.19	0.16	0.14	0.27



รูปที่ 79 RMSE ของแบบจำลองแต่ละประเภทในแต่ละ time-Step



รูปที่ 80 RMSE ของแบบจำลองแต่ละประเภทเมื่อเปรียบเทียบระหว่างการใช้และไม่ใช้ GA



รูปที่ 81 การเปรียบเทียบระหว่างค่าจริงและค่าพยากรณ์จากแต่ละแบบจำลอง

จุฬาลงกรณ์มหาวิทยาลัย
บทที่ 5
สรุปผลการวิจัย

วิทยานิพนธ์ฉบับนี้นำเสนอวิธีการประยุกต์ใช้นิวรอนเน็ตเวิร์กสำหรับการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ 3 วิธีได้แก่

1. การพยากรณ์กำลังผลิตไฟฟ้าระยะสั้นมากด้วยการพยากรณ์อากาศเชิงตัวเลข (Very Short-Term Solar Power Forecast using Data from NWP Model)
2. การพยากรณ์กำลังผลิตไฟฟ้าระยะสั้นมากด้วยนิวรอนเน็ตเวิร์กและขั้นตอนวิธีเชิงพันธุกรรม (Very Short-Term Solar Power Forecasting Using Genetic Algorithm Based Deep Neural Network)
3. การพยากรณ์กำลังผลิตไฟฟ้ารายภาคจากพลังงานแสงอาทิตย์โดยใช้โครงข่ายประสาทเทียมร่วมกับการปรับจูนไฮเปอร์พารามิเตอร์ (Deep Neural Network based on

Genetic Algorithm and Ensemble Methods for Regional Solar Power Forecasting in Thailand)

สำหรับการศึกษาวิจัยในวิธีการแรกพบว่าเน็ตเวิร์ก FCL มีประสิทธิภาพดีกว่าเน็ตเวิร์ก RNN จากการเพิ่มมิติของเวลาให้กับเน็ตเวิร์ก FCL ทำให้เน็ตเวิร์ก FCL มีค่าเปอร์เซ็นต์ RMSE ต่ำกว่า RNN อยู่ 0.35% และเมื่อเปรียบเทียบประสิทธิภาพระหว่าง FCL ที่มีและไม่มีมิติของเวลาพบว่า FCL ที่มีมิติของเวลามีประสิทธิภาพที่ดีกว่า FCL ที่ไม่มีมิติของเวลาอยู่ 0.86% สำหรับการปรับปรุงประสิทธิภาพด้วยการใช้ค่าพยากรณ์อากาศและค่าวัดจริงช่วยปรับปรุงประสิทธิภาพให้ดีขึ้น 0.12% และการปรับปรุงประสิทธิภาพด้วยการใส่ค่าเฉลี่ยเคลื่อนที่แบบฟังก์ชันเอกซ์โพเนนเชียลของความเข้มแสงอาทิตย์ที่วัดได้ช่วยปรับปรุงประสิทธิภาพให้ดีขึ้นอีก 0.13% เมื่อใช้เทคนิคการปรับปรุงประสิทธิภาพทั้งหมดที่นำเสนอจะช่วยปรับปรุงประสิทธิภาพให้ดีขึ้น 0.25% หรือมีความผิดพลาดเพียง 0.65 kW ซึ่งจะทำให้แบบจำลองพยากรณ์อากาศสามารถพยากรณ์ได้แม่นยำขึ้นและเมื่อพิจารณาค่าความคลาดเคลื่อนทางสถิติ MBE ในแต่ละช่วงเวลาของวันพบว่าแบบจำลอง FCL จะให้ค่าพยากรณ์ที่สูงกว่าความเป็นจริงในขณะที่แบบจำลอง RNN ให้ค่าพยากรณ์ที่ต่ำกว่าความเป็นจริง แต่อย่างไรก็ตามค่าพยากรณ์เหล่านี้ก็สูงหรือต่ำกว่าค่าจริงเพียงเล็กน้อยซึ่งจะเห็นได้จากค่า RMSE

สำหรับการศึกษาในวิธีการที่สองแสดงให้เห็นว่าเน็ตเวิร์ก GRU ให้การพยากรณ์ที่ดีในแง่ของความแม่นยำด้วยค่า RMSE 7.83% หรือมีความผิดพลาด 0.626 kW แต่หากพิจารณาในแง่ของความเร็วในการฝึกสอน เน็ตเวิร์ก FCL ใช้เวลาในการฝึกสอนน้อยที่สุดเมื่อเทียบกับแบบจำลองอื่น โดยมีค่าอยู่ที่ 42 วินาที ถ้าพิจารณาทั้งในแง่ของเวลาในการฝึกสอนและความแม่นยำในการพยากรณ์แบบจำลอง CuDNNGRU จะให้ผลลัพธ์ที่ดีที่สุดด้วย RMSE 7.87% หรือมีความผิดพลาดไม่เกิน 0.63 kW และเวลาในการฝึกสอนอยู่ที่ 51 วินาที อย่างไรก็ตามที่รู้กันว่าความแม่นยำของเน็ตเวิร์กขึ้นอยู่กับปริมาณและคุณภาพของข้อมูลที่ใช้ในการฝึกสอน ดังนั้นในกรณีนี้การเลือกพารามิเตอร์ของ GA ที่เหมาะสมก็เป็นสิ่งสำคัญ ยกตัวอย่างเช่น ค่าความน่าจะเป็นในการสลับสายพันธุ (crossover probability) ควรจะมีค่าอยู่ระหว่าง 0.6 และ 0.95 (บางการศึกษาแนะนำ 0.85-0.95 [34]) ค่าความน่าจะเป็นในการกลายพันธุ (mutation probability) ควรมีค่าอยู่ระหว่าง 0 และ 0.01 ต่อตำแหน่งโครโมโซม เนื่องจากถ้าเรากำหนดค่าสูงเกินไป GA จะมีลักษณะคล้ายกับการค้นหาด้วยการสุ่ม (random search) ถ้าขนาดของประชากรใหญ่เกินไปเวลาในการฝึกสอนก็ยากที่จะยอมรับได้ (ขนาดของประชากรควรเป็น 10 เท่าของจำนวนตัวแปร [35]) ข้อเท็จจริงเหล่านี้แสดงให้เห็นว่า GA ยังคงมีข้อจำกัดอยู่ ดังนั้นอัลกอริทึมที่ใช้ในการค้นหาไฮเปอร์พารามิเตอร์ยังคงมีการพัฒนาต่อไป

สำหรับการศึกษาวิจัยในวิธีการที่สามพบว่าโครงข่ายประสาทเทียมแบบลึกด้วยขั้นตอนวิธีเชิงพันธุกรรมและวิธีการเรียนรู้แบบรวมกลุ่มสำหรับการพยากรณ์กำลังผลิตไฟฟ้าของโรงไฟฟ้าพลังงานแสงอาทิตย์ในประเทศไทยล่วงหน้าหนึ่งวันที่ประกอบไปด้วย 1) การประยุกต์ใช้อินพุตจากหลายโมเดลพยากรณ์อากาศเพื่อผสมผสานข้อดีของแต่ละโมเดลด้วยวิธีการเรียนรู้แบบรวมกลุ่ม (Ensemble Learning Method) เพื่อปรับปรุงค่า Bias จากแต่ละโมเดล 2) การประยุกต์ใช้ GA ในการคัดเลือกตัวแปรเพื่อคัดเลือกตัวแปรที่เกี่ยวข้องและสำคัญต่อการพยากรณ์เนื่องจากตัวแปรที่ไม่เกี่ยวข้องจะทำให้ประสิทธิภาพของโมเดลแย่ลง 3) การประยุกต์ใช้เคมีนคลัสเตอร์ริง (K-means clustering) เพื่อคัดเลือกค่าพยากรณ์อากาศรอบโรงไฟฟ้าเพื่อลดค่าความผันผวน 4) การประยุกต์ใช้ GA ปรับจูนไฮเปอร์พารามิเตอร์ของโมเดลเพื่อปรับปรุงประสิทธิภาพของโมเดลและ 5) การผสมผสานค่าพยากรณ์จากแต่ละโมเดลเข้าด้วยกันด้วยนิเวศน์วิวิธวิธีเพื่อลดความผันผวนของค่าพยากรณ์ สำหรับกรณีศึกษาการพยากรณ์กำลังผลิตไฟฟ้าของโรงไฟฟ้าพลังงานแสงอาทิตย์ในประเทศไทยทั้งหมด 7 โรงไฟฟ้า พบว่าการศึกษานี้ให้ค่า RMSE อยู่ที่ 6.94% หรือมีความผิดพลาด 30.26 MW ซึ่งลดลง 0.62% เมื่อเทียบกับโมเดลเดี่ยวที่ใช้ค่าวัดจริงที่ดีที่สุดซึ่งการศึกษานี้สามารถนำไปประยุกต์ใช้งานจริงได้เพราะสามารถใช้ได้กับทั้งโรงไฟฟ้าที่มีและไม่มีอุปกรณ์ตรวจวัดสภาพอากาศเนื่องจากสามารถใช้ค่าพยากรณ์อากาศแทนได้ซึ่งให้ประสิทธิภาพใกล้เคียงกัน

แนวทางการศึกษาในอนาคตคือการพยากรณ์กำลังผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ทั้งแบบโซลาร์ฟาร์มและโซลาร์รูฟท็อปที่อยู่นอกระบบของการไฟฟ้าฝ่ายผลิตแห่งประเทศไทยหรือการไฟฟ้าส่วนภูมิภาคเนื่องจากกำลังผลิตไฟฟ้าส่วนนี้เป็นส่วนที่อยู่นอกระบบทำให้ผู้ผลิตไฟฟ้าบริหารจัดการพลังงานส่วนนี้ได้ยาก

บรรณานุกรม

1. Suksamosorn, S., N. Hoonchareon, and J. Songsiri, *Influential Variable Selection for Improving Solar Forecasts from Numerical Weather Prediction*, in *2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. 2018: Chiang Rai, Thailand. p. 333-336.
2. Fonseca Junior, J.G.d.S., et al., *Regional forecasts of photovoltaic power generation according to different data availability scenarios: a study of four methods*. *Progress in Photovoltaics: Research and Applications*, 2015. **23**(10): p. 1203-1218.
3. Zhang, Y., et al., *Day-Ahead Power Output Forecasting for Small-Scale Solar Photovoltaic Electricity Generators*. *IEEE Transactions on Smart Grid*, 2015. **vol. 6, no. 5**: p. pp. 2253-2262.
4. Haupt, S.E. and B. KosoviĆ, *Variable Generation Power Forecasting as a Big Data Problem*. *IEEE Transactions on Sustainable Energy*, 2017. **vol. 8, no. 2**: p. pp. 725-732.
5. Nitisanon, S. and N. Hoonchareon, *Solar power forecast with weather classification using self-organized map*, in *2017 IEEE Power & Energy Society General Meeting*. 2017: Chicago, IL. p. pp. 1-5.
6. Zhang, X., et al., *A Solar Time Based Analog Ensemble Method for Regional Solar Power Forecasting*. *IEEE Transactions on Sustainable Energy*, 2019. **vol. 10, no. 1**: p. pp. 268-279.
7. Li, Y. and e. al., *A machine-learning approach for regional photovoltaic power forecasting*, in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. 2016: Boston, MA. p. 1-5.
8. Omain, et al. *Predicting daily incoming solar energy from weather data*. 2013.
9. Chu, Y., et al., *Short-term reforecasting of power output from a 48 MWe solar PV plant*. *Solar Energy*, 2015. **112**: p. 68-77.
10. Bae, K.Y., H.S. Jang, and D.K. Sung, *Hourly Solar Irradiance Prediction Based on Support Vector Machine and Its Error Analysis*. *IEEE Transactions on Power Systems*, 2017. **32**: p. 935-945.
11. Zhang, P., B. Verma, and K. Kumar, *A neural-genetic algorithm for feature selection and*

- breast abnormality classification in digital mammography*. 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), 2004. **3**: p. 2303-2308 vol.3.
12. Jiang, S., et al., *Modified genetic algorithm-based feature selection combined with pre-trained deep neural network for demand forecasting in outpatient department*. Expert Syst. Appl., 2017. **82**: p. 216-230.
 13. Kazemi, S.M.R., et al., *An evolutionary-based adaptive neuro-fuzzy inference system for intelligent short-term load forecasting*. International Transactions in Operational Research, 2014. **21**(2): p. 311-326.
 14. Sheikhan, M. and N. Mohammadi, *Neural-based electricity load forecasting using hybrid of GA and ACO for feature selection*. Neural Computing and Applications, 2011. **21**: p. 1961-1970.
 15. Urraca-Valle, R., et al., *Improving hotel room demand forecasting with a hybrid GA-SVR methodology based on skewed data transformation, feature selection and parsimony tuning*. Logic Journal of the IGPL, 2017. **25**: p. 877-889.
 16. Liu, Y., et al., *Wrapper Feature Selection Optimized SVM Model for Demand Forecasting*, in *2008 The 9th International Conference for Young Computer Scientists, Human*. 2008. p. 953-958.
 17. Sexton, R.S., R.S. Sriram, and H. Etheridge, *Improving Decision Effectiveness of Artificial Neural Networks: A Modified Genetic Algorithm Approach*. Decision Sciences, 2003. **34**(3): p. 421-442.
 18. Tsai, C.-F., W. Eberle, and C.-Y. Chu, *Genetic algorithms in feature and instance selection*. Knowledge-Based Systems, 2013. **39**: p. 240-247.
 19. Tong, D.L. and R. Mintram, *Genetic Algorithm-Neural Network (GANN): a study of neural network activation functions and depth of genetic algorithm search applied to feature selection*. International Journal of Machine Learning and Cybernetics, 2010. **1**: p. 75-87.
 20. H, C., *Using artificial intelligence in chemistry and biology: A practical guide*. In: *Chapter Evolutionary Algorithms*, CRC Press, Taylor & Francis Group, Boca Raton, London pp 113–172. 2008: Boca Raton: CRC Press.
 21. De Jong, K.A. and W.M. Spears. *An analysis of the interacting roles of population size and crossover in genetic algorithms*. 1991. Berlin, Heidelberg: Springer Berlin Heidelberg.

22. N.Moon, S. and D.N. Bawane, *Optimal Feature Selection By Genetic Algorithm For Classification Using Neural Network*. International Research Journal of Engineering and Technology (IRJET), 2015. **Volume: 02 Issue: 05**.
23. Heinis, S., et al., *Of Genes and Machines: Application of a Combination of Machine Learning Tools to Astronomy Data Sets*. The Astrophysical Journal, 2016. **821**.
24. Bouktif, S., et al. *Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches*. 2018.
25. Erguzel, T.T., et al., *Feature Selection and Classification of Electroencephalographic Signals: An Artificial Neural Network and Genetic Algorithm Based Approach*. Clinical EEG and neuroscience, 2015. **46 4**: p. 321-6.
26. Eiben, A.E. and J.E. Smith, *Introduction to Evolutionary Computing*. 2003: SpringerVerlag.
27. Holland, J.H., *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. 1992: MIT Press. 228.
28. Gen, M. and R. Cheng, *Genetic Algorithm and Engineering Design*. 1997, John & Wiley Sons, New York.
29. Zhong, J., et al., *Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms*, in *Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce Vol-2 (CIMCA-IAWTIC'06) - Volume 02*. 2005, IEEE Computer Society. p. 1115-1121.
30. Goldberg, D.E. and J. Robert Lingle, *AllelesLoci and the Traveling Salesman Problem*, in *Proceedings of the 1st International Conference on Genetic Algorithms*. 1985, L. Erlbaum Associates Inc. p. 154-159.
31. Saha, S., et al., *The NCEP Climate Forecast System Version 2*. Journal of Climate, 2014. **27(6)**: p. 2185-2208.
32. Fran, et al., *DEAP: a python framework for evolutionary algorithms*, in *Proceedings of the 14th annual conference companion on Genetic and evolutionary computation*. 2012, ACM: Philadelphia, Pennsylvania, USA. p. 85-92.
33. Wilson, A.C., et al., *The Marginal Value of Adaptive Gradient Methods in Machine*

Learning. 2017: p. 4148-4158.

34. Zhong, J., et al., *Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms*. International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06), 2005. 2: p. 1115-1121.
35. Bergstra, J., et al., *Algorithms for hyper-parameter optimization*, in *Proceedings of the 24th International Conference on Neural Information Processing Systems*. 2011, Curran Associates Inc.: Granada, Spain. p. 2546-2554.





จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ประวัติผู้เขียน

ชื่อ-สกุล	นาย สุกฤษฎ์ ใจดี
วัน เดือน ปี เกิด	15 สิงหาคม 2534
สถานที่เกิด	โรงพยาบาลรามาริบัติ กรุงเทพมหานคร
วุฒิการศึกษา	ปริญญาตรีวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
ที่อยู่ปัจจุบัน	4/1 ม.2 ต.บางสะแก อ.บางคนที จ.สมุทรสงคราม 75120
ผลงานตีพิมพ์	<p>S. Jaidee and W. Pora, "Very Short-Term Solar Power Forecasting Using Genetic Algorithm Based Deep Neural Network," 2019 4th International Conference on Information Technology (InCIT), Bangkok, Thailand, 2019, pp. 184-189.</p> <p>S. Jaidee and W. Pora, "Very Short-Term Solar Power Forecast using Data from NWP Model," 2019 4th International Conference on Information Technology (InCIT), Bangkok, Thailand, 2019, pp. 44-49.</p> <p>S. Jaidee and W. Pora, "Deep Neural Network based on Genetic Algorithm and Ensemble Methods for Regional Solar Power Forecasting in Thailand," 2019 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), Bangkok, Thailand, 2019, pp. 377-380.</p>
รางวัลที่ได้รับ	Echonect Lite Hackathon 2019 by Chula, Second Runner-Up,