การพยากรณ์การใช้ไฟฟ้าในประเทศไทยด้วยเทคนิคการเรียนรู้ของเครื่อง

Forecasting Electricity Consumption in Thailand Using Machine Learning Techniques

นงลักษณ์ เคลื่อนไธสงค์[1] อัชนัญญา ปิตะแสง[2] นันทิชา อุสาใจ[3] ศศิธร พินิรัมย์[4]

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบและทดลองประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ในการพยากรณ์ความต้องการใช้พลังงานไฟฟ้าในอนาคตของประเทศไทย โดยใช้ข้อมูลจำนวน 8,806 รายการ ซึ่งประกอบด้วยตัวแปรทั้งหมด 9 ตัว ได้แก่ ปริมาณการผลิตไฟฟ้าในแต่ละภูมิภาค การนำเข้าพลังงานไฟฟ้าจากประเทศเพื่อนบ้าน ฤดูกาล ปริมาณความต้องการใช้ไฟฟ้าทั้งหมด และช่วงเวลาของวัน เป็นต้น ข้อมูลดังกล่าวถูกนำมาวิเคราะห์และสร้างแบบจำลองด้วยเทคนิคการเรียนรู้ของเครื่อง 3 เทคนิค ได้แก่ ต้นไม้ป่าสุ่ม (Random Forest) โครงข่ายประสาทเทียม (Neural Network) และการถดถอยเชิงเส้น (Linear Regression) โดยใช้กระบวนการวิเคราะห์ข้อมูลเชิงพยากรณ์ (Regression Model) ผ่านโปรแกรม Altair AI Studio เพื่อเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลอง

ผลการวิจัยพบว่า แบบจำลองทั้งสามสามารถพยากรณ์ความต้องการใช้ไฟฟ้าได้อย่างมีประสิทธิภาพ โดยแบบจำลอง Random Forest ให้ผลการพยากรณ์ที่ดีที่สุด โดยมีค่า ค่าความคลาดเคลื่อนรากที่สองของค่าเฉลี่ย (Root Mean Square Error) เท่ากับ 514.03 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) เท่ากับ 264,225.81 ค่าความคลาดเคลื่อนเฉลี่ย (Absolute Error) เท่ากับ 399.71 ค่าผลรวมของความคลาดเคลื่อนกำลังสอง (Squared Error) เท่ากับ 264,356.34 และค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of Determination) เท่ากับ 96.70% ซึ่งแสดงให้เห็นว่าเทคนิค Random Forest มีศักยภาพสูงสุดในการคาดการณ์ความต้องการใช้พลังงานไฟฟ้าในอนาคตของประเทศไทยอย่างแม่นยำ อันจะเป็นประโยชน์ต่อการวางแผนและบริหารจัดการพลังงานไฟฟ้าในระดับประเทศได้อย่างมีประสิทธิภาพ

คำสำคัญ: พลังงานไฟฟ้า, การพยากรณ์, การเรียนรู้ของเครื่อง

ABSTRACT

The objective of this research is to develop and evaluate the performance of machine learning techniques for forecasting future electricity demand in Thailand. The study utilized a dataset consisting of 8,806 records with nine variables, including regional electricity generation, electricity imports from neighboring countries, seasonal factors, total electricity demand, and time of day, among others. These data were analyzed and modeled using three machine learning techniques: Random Forest, Neural Network, and Linear Regression, within the predictive analytics framework (Regression Model) implemented through the Altair AI Studio software, in order to compare the performance of each model.

The results revealed that all three models demonstrated effective forecasting capability, with the Random Forest model achieving the highest accuracy. Specifically, it yielded a Root Mean Square Error (RMSE) of 514.03, a Mean Square Error (MSE) of 264,225.81, an Absolute Error (AE) of 399.71, a Squared Error (SE) of 264,356.34, and a Coefficient of Determination (R^2) of 96.70%. These findings indicate that the Random Forest technique provides the most accurate prediction of future electricity demand in Thailand, which can serve as a valuable tool for energy planning and management at the national level.

Keywords: Electricity demand, Forecasting, Machine Learning

บทนำ

ในปัจจุบันการใช้พลังงานไฟฟ้าเป็นปัจจัยสำคัญในการพัฒนาประเทศ ทั้งในภาคอุตสาหกรรม พาณิชยกรรม และภาคครัวเรือน การคาดการณ์ความต้องการใช้พลังงานไฟฟ้าในอนาคตจึงมีความสำคัญอย่างยิ่งต่อการวางแผนการผลิต การจัดสรรทรัพยากร และการบริหารจัดการพลังงานอย่างมีประสิทธิภาพ [1] หากการพยากรณ์มีความคลาดเคลื่อน อาจส่งผลให้เกิดการผลิตไฟฟ้าเกินความต้องการ หรือเกิดภาวะขาดแคลนไฟฟ้า [2] ซึ่งล้วนส่งผลกระทบต่อเศรษฐกิจและคุณภาพชีวิตของประชาชนโดยรวม ในช่วงทศวรรษที่ผ่านมา ประเทศไทยมีแนวโน้มการใช้พลังงานไฟฟ้าเพิ่มขึ้นอย่างต่อเนื่อง [3] อันเป็นผลมาจากการขยายตัวทางเศรษฐกิจ การพัฒนาเมือง และความก้าวหน้าทางเทคโนโลยี ดังนั้น หน่วยงานที่เกี่ยวข้องกับพลังงาน เช่น การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.) จำเป็นต้องมีเครื่องมือที่สามารถพยากรณ์ความต้องการใช้ไฟฟ้าในอนาคตได้อย่างแม่นยำ [4] เพื่อรองรับการกำหนดนโยบายด้านพลังงาน การวางแผนก่อสร้างโรงไฟฟ้า และการจัดการระบบโครงข่ายไฟฟ้าอย่างยั่งยืน เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เป็นแนวทางที่ได้รับความสนใจอย่างมากในการพยากรณ์ เนื่องจากสามารถเรียนรู้รูปแบบและความสัมพันธ์ระหว่างข้อมูลจำนวนมากได้อย่างมีประสิทธิภาพ [5] โดยไม่จำเป็นต้องกำหนดสมการเชิงเส้นล่วงหน้า การประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่อง จึงมีศักยภาพสูงในการวิเคราะห์ข้อมูลที่มีความซับซ้อนและไม่เป็นเชิงเส้น เช่น ข้อมูลการใช้พลังงานไฟฟ้าที่ได้รับอิทธิพลจากหลายปัจจัยทั้งด้านภูมิภาค เวลา และฤดูกาล

งานวิจัยครั้งนี้จึงมุ่งศึกษาประสิทธิภาพของเทคนิคการถดถอย (Regression Techniques) 3 รูปแบบ ได้แก่ ต้นไม้ป่าสุ่ม (Random Forest) โครงข่ายประสาทเทียม (Neural Network) และการถดถอยเชิงเส้น (Linear Regression) [6] เพื่อใช้ในการพยากรณ์ความต้องการใช้พลังงานไฟฟ้าในอนาคตของประเทศไทย โดยใช้ข้อมูลจริงจำนวน 8,806 รายการ ซึ่งประกอบด้วยตัวแปรอิสระที่หลากหลาย เช่น ปริมาณการผลิตไฟฟ้า การนำเข้าพลังงานไฟฟ้า ฤดูกาล และช่วงเวลาของวัน จากนั้นทำการวิเคราะห์และเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองด้วยตัวชี้วัด เช่น ค่าความคลาดเคลื่อนรากที่สองของค่าเฉลี่ย ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ค่าความคลาดเคลื่อนเฉลี่ย ค่าผลรวมของความคลาดเคลื่อนกำลังสอง และค่าสัมประสิทธิ์การตัดสินใจ เพื่อค้นหาเทคนิคที่เหมาะสมที่สุดสำหรับการคาดการณ์พลังงานไฟฟ้าในบริบทของประเทศไทย ผลการศึกษาพบว่า เทคนิค Random Forest ให้ค่าความแม่นยำสูงสุด โดยมีค่า R^2 เท่ากับ 96.70% ซึ่งแสดงให้เห็นถึงศักยภาพของวิธีนี้ในการนำไปประยุกต์ใช้เป็นแบบจำลองต้นแบบสำหรับการพยากรณ์ความต้องการใช้ไฟฟ้าในอนาคต ทั้งนี้ผลลัพธ์ที่ได้สามารถช่วยสนับสนุนการวางแผนด้านพลังงานของประเทศให้มีประสิทธิภาพ และส่งเสริมการใช้เทคโนโลยีการวิเคราะห์ข้อมูลเชิงลึก (Data Analytics) เพื่อพัฒนาระบบพลังงานของประเทศให้ยั่งยืนในอนาคต

วัตถุประสงค์การวิจัย

1.เพื่อศึกษาการประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่อง สำหรับการพยากรณ์ความต้องการใช้พลังงานไฟฟ้าในอนาคตของประเทศไทย

2.เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง

วิธีดำเนินการวิจัย

งานวิจัยนี้ดำเนินการวิเคราะห์ข้อมูลตามกรอบแนวคิดของกระบวนการทำเหมืองข้อมูลแบบมาตรฐานสากล หรือ Cross-Industry Standard Process for Data Mining (CRISP–DM) ซึ่งเป็นขั้นตอนการวิจัยที่ได้รับการยอมรับอย่างแพร่หลายในการประยุกต์ใช้เทคโนโลยีการเรียนรู้ของเครื่อง เพื่อวิเคราะห์ข้อมูลและพยากรณ์แนวโน้มเชิงธุรกิจและพลังงาน ประกอบด้วย 6 ขั้นตอนหลัก

1.การทำความเข้าใจปัญหา (Process Understanding)

ปัจจุบันประเทศไทยมีความต้องการใช้พลังงานไฟฟ้าเพิ่มขึ้นอย่างต่อเนื่อง อันเป็นผลมาจากการเติบโตทางเศรษฐกิจ การขยายตัวของภาคอุตสาหกรรม และการเพิ่มขึ้นของจำนวนประชากรในเขตเมือง [7] การบริหารจัดการระบบผลิตและจ่ายไฟฟ้าจึงต้องอาศัยการพยากรณ์ความต้องการใช้ไฟฟ้าที่มีความถูกต้องและแม่นยำ เพื่อช่วยให้หน่วยงานที่เกี่ยวข้อง เช่น การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.) สามารถวางแผนการผลิต กำหนดนโยบาย และจัดสรรทรัพยากรด้านพลังงานได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม การคาดการณ์ด้วยวิธีทางสถิติดั้งเดิม

เช่น การถดถอยเชิงเส้นหรืออนุกรมเวลา มักไม่สามารถสะท้อนความซับซ้อนของข้อมูลได้อย่างครบถ้วน โดยเฉพาะในกรณีที่ข้อมูลมีความสัมพันธ์เชิงไม่เส้นตรง (Nonlinear Relationship) หรือได้รับอิทธิพลจากปัจจัยหลายด้านพร้อมกัน [8] งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาและเปรียบเทียบ

2. การทำความเข้าใจข้อมูล (Data Understanding)

ผู้วิจัยได้ดำเนินการรวบรวมข้อมูลที่ได้จากระบบไฟฟ้าไทย การผลิตไฟฟ้ารายชั่วโมง ความต้องการใช้ไฟฟ้า และกระแสไฟฟ้าข้ามพรมแดน Zenodo (2025) [9] โดยไฟล์ที่นำเข้าอยู่ในรูปแบบไฟล์ CSV เพื่อนำข้อมูลเข้าไปวิเคราะห์ถูกจัดเก็บตั้งแต่วันที่ 1 เดือนมกราคม พ.ศ.2567 ถึงวันที่ 1 เดือนมกราคม พ.ศ.2568 จำนวนข้อมูลทั้งสิ้น 8,806 แถว สำหรับข้อมูลที่นำมาใช้ในการวิเคราะห์นี้ คือ ปัจจัยที่ทำให้เกิดการใช้ไฟฟ้ารวม จำนวน 9 ตัวแปร ประกอบไปด้วย

ตารางที่ 1 ตารางข้อมูลที่ใช้ในการสร้างตัวแบบการพยากรณ์

ลำดับ	ตัวแปร	คำอภิบาย	ประเภท
1	date	วัน/เดือน/ปี (ID)	Date
2	north_generation	การผลิตไฟฟ้าภาคเหนือ (MW)	Real
3	metropolitan_generation	การผลิตไฟฟ้าในพื้นที่มหานคร (MW)	Real
4	central_generation	การผลิตไฟฟ้าภาคกลาง (MW)	Real
5	malaysia_IMP	การนำเข้าไฟฟ้าจากมาเลเซีย (MW)	Real
6	laos_IMP	การนำเข้าไฟฟ้าจากลาว (MW)	Real
7	hour_number	ชั่วโมงของวัน (0-23)	Integer
8	season	ฤดูกาล	Integer
9	total_demand	ความต้องการไฟฟ้ารวม (Label)	Real

3. การเตรียมข้อมูล (Data Preparation)

กระบวนการเตรียมข้อมูลเป็นขั้นตอนสำคัญก่อนการสร้างแบบจำลองการพยากรณ์ เพื่อให้ข้อมูลอยู่ในรูปแบบที่เหมาะสมต่อการนำไปวิเคราะห์ด้วยเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) โดยดำเนินการตามลำดับดังนี้

3.1 การนำเข้าข้อมูล (Read CSV)

ข้อมูลที่ใช้ในการวิจัยถูกรวบรวมและจัดเก็บในรูปแบบไฟล์ CSV (Comma-Separated Values) จากนั้นจึงนำเข้าข้อมูลเข้าสู่สภาพแวดล้อมการวิเคราะห์ในโปรแกรม RapidMiner Studio จุดประสงค์ของขั้นตอนนี้คือเพื่อให้ข้อมูลพร้อมสำหรับการตรวจสอบคุณภาพและการแปลงค่าขั้นต้น ในการนำเข้าข้อมูล มีการตรวจสอบ encoding ของไฟล์ เพื่อหลีกเลี่ยงปัญหาการอ่านตัวอักษรผิดรูป ตรวจสอบ หัวคอลัมน์ (header) ให้ถูกต้องตรงกับชื่อของตัวแปร และตรวจสอบ ชนิดข้อมูลเบื้องต้น (data types) เพื่อให้แน่ใจว่าข้อมูลแต่ละคอลัมน์ถูกตีความตามประเภทที่เหมาะสม [11]

3.2 การกำหนดบทบาทของตัวแปร (Set Role)

หลังจากนำเข้าข้อมูลแล้ว มีการกำหนดบทบาท ให้กับแต่ละคอลัมน์ของข้อมูล เพื่อให้กระบวนการวิเคราะห์และการสร้างแบบจำลองเป็นไปอย่างถูกต้อง การกำหนดบทบาทที่ชัดเจนช่วยให้ขั้นตอนการเลือกตัวแปร และการประเมินประสิทธิภาพของโมเดล Model Evaluation ทำงานได้อย่างถูกต้องและสอดคล้องกับวัตถุประสงค์การวิจัย [11]

3.3 การแปลงชนิดข้อมูลเป็นตัวเลข (Attributer/Features)

ในบางกรณี ข้อมูลที่ควรอยู่ในรูปแบบเชิงตัวเลขอาจถูกอ่านเป็นข้อความ (String) เนื่องจากการมีเครื่องหมายจุลภาค (Comma), สัญลักษณ์หน่วย (%), หรือรูปแบบวันที่ ที่ไม่เป็นมาตรฐาน ขั้นตอนนี้จึงดำเนินการแปลงค่าดังกล่าวให้เป็นชนิดข้อมูลเชิงตัวเลข (Numeric) เพื่อให้สามารถนำไปคำนวณและวิเคราะห์ได้อย่างถูกต้อง ทั้งนี้ต้องระมัดระวังเรื่อง สัญลักษณ์หน่วย, รูปแบบการแยกทศนิยม, และค่าที่อยู่นอกช่วง (Out-of-range) ซึ่งอาจส่งผลต่อความถูกต้องของการวิเคราะห์ [13]

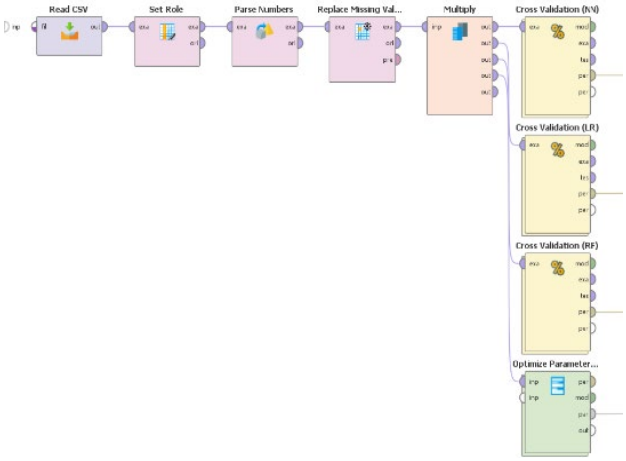
3.4 การจัดการค่าที่หายไป (Replace missing values)

ข้อมูลจริงมักประสบปัญหาค่าขาดหาย (Missing values) ซึ่งอาจเกิดจากการบันทึกข้อมูลไม่ครบถ้วน ความผิดพลาดระหว่างการเก็บข้อมูล หรือปัญหาทางเทคนิคอื่น ๆ การละเลยค่าที่หายไปอาจส่งผลให้แบบจำลองเกิดความคลาดเคลื่อนและความ

แม่นยำลดลง ดังนั้นจึงต้องมีขั้นตอนในการตรวจสอบและจัดการอย่างเหมาะสม [14]

4. การสร้างแบบจำลอง (Modeling)

ขั้นตอนนี้เป็นกระบวนการสร้างแบบจำลองเพื่อใช้ในการพยากรณ์ความต้องการใช้ไฟฟ้าในอนาคต โดยผู้วิจัยได้ใช้โปรแกรม RapidMiner Studio โดยมีการใช้เทคนิคการทำเหมืองข้อมูล 3 เทคนิค ได้แก่ เทคนิคต้นไม้ป่าสุ่ม (Random Forest) เทคนิคโครงข่ายประสาทเทียม (Neural Network) และเทคนิคการถดถอยเชิงเส้น (Linear Regression)

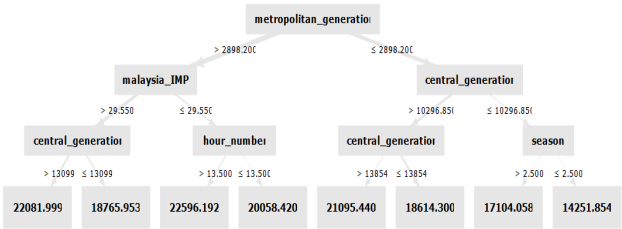


ภาพที่ 1 ขั้นตอนการสร้างตัวแบบและการวัดประสิทธิภาพของตัวแบบด้วยโปรแกรม Altair Ai Studio

4.1 เทคนิคต้นไม้ป่าสุ่ม (Random Forest)

เทคนิคต้นไม้ป่าสุ่ม (Random Forest) เป็นวิธีที่รวมผลของหลายต้นไม้ตัดสินใจ (Decision Trees) เข้าด้วยกัน โดยให้แต่ละต้นเรียนรู้จากชุดข้อมูลย่อยแบบสุ่ม (Bootstrap Aggregating) แล้วนำผลการทำนายทั้งหมดมาหาค่าเฉลี่ยเพื่อลดความคลาดเคลื่อนของแบบจำลอง การใช้วิธีนี้ช่วยลดปัญหา Overfitting และเพิ่มความเสถียรของการพยากรณ์ เหมาะกับข้อมูลที่มีความซับซ้อนและไม่เป็นเชิงเส้น [10]

จากภาพที่2 แสดงโครงสร้างของแบบจำลองที่ใช้สำหรับการพยากรณ์ความต้องการใช้ไฟฟ้า โดยแต่ละโหนดในแผนภูมิต้นไม้แสดงถึง เงื่อนไขการแบ่งข้อมูล ซึ่งเป็นเกณฑ์ที่โมเดลใช้ในการจำแนกข้อมูลออกเป็นกลุ่มย่อยตามค่าของตัวแปรอิสระ โหนดราก ของแบบจำลองคือ ปริมาณการผลิตไฟฟ้าในเขตมหานคร (metropolitan_generation) ซึ่งเป็นตัวแปรหลักที่ใช้แบ่งข้อมูลในระดับแรก โดยใช้เกณฑ์การตัดสินที่ค่า 2898.20 หากค่าของการผลิตไฟฟ้าในเขตมหานคร มากกว่า 2898.20 โมเดลจะพิจารณาตัวแปรถัดไป เช่น ปริมาณการนำเข้าไฟฟ้าจากประเทศมาเลเซีย(malaysia_IMP) และกำลังผลิตจากภาคกลาง (central_generation) ในทางกลับกัน หากค่าน้อยกว่าหรือเท่ากับ 2898.20 โมเดลจะพิจารณาตัวแปร กำลังผลิตจากภาคกลาง และ ฤดูกาล(season) ค่าตัวเลขที่ปรากฏในแต่ละกล่องปลายทาง (Leaf Node) เช่น 22081.999, 22596.192 และ 18614.300 แสดงถึง ค่าการพยากรณ์เฉลี่ยของปริมาณความต้องการใช้ไฟฟ้า มีหน่วยเป็น เมกะวัตต์ ที่โมเดลคำนวณได้เมื่อข้อมูลเข้าสู่เส้นทางการตัดสินใจนั้น ๆ



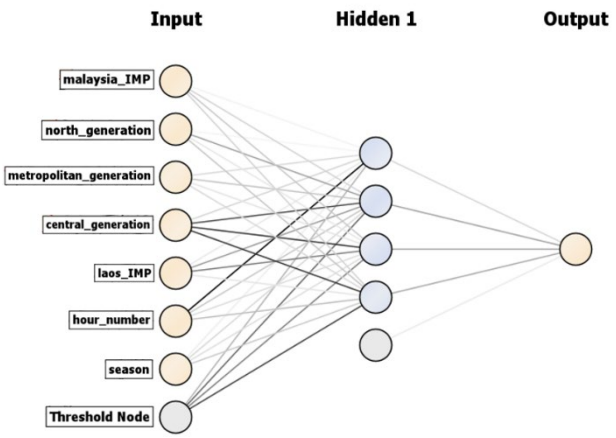
ภาพที่ 2 ตัวอย่างการทำงานของ Random Forest

4.2 เทคนิคโครงข่ายประสาทเทียม (Neural Network)

เทคนิคโครงข่ายประสาทเทียม (Neural Network) เป็นเทคนิคที่จำลองการทำงานของเซลล์ประสาทในสมองมนุษย์ โดยมีโครงสร้างประกอบด้วยชั้นอินพุต (Input Layer), ชั้นซ่อน (Hidden Layer) และชั้นเอาต์พุต (Output Layer) [26]

จากภาพที่ 3 แสดงโครงสร้างของแบบจำลองโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) ที่ใช้ในการพยากรณ์ความต้องการใช้ไฟฟ้า โดยประกอบด้วยสามชั้นหลัก ได้แก่ ชั้นข้อมูลนำเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นผลลัพธ์ (Output Layer) ชั้นข้อมูลนำเข้าประกอบด้วยตัวแปรอิสระที่มีผลต่อความต้องการใช้ไฟฟ้า เช่น ปริมาณการผลิตไฟฟ้าในแต่ละภูมิภาค การนำเข้าพลังงานจากประเทศเพื่อนบ้าน ชั่วโมงของวัน และฤดูกาล รวมถึง Threshold Node ซึ่งทำหน้าที่ปรับสมดุลของแบบจำลอง ข้อมูลจากชั้นนำเข้าจะถูกประมวลผลผ่านชั้นซ่อน โดยใช้ค่าน้ำหนัก (Weights) และฟังก์ชันกระตุ้น (Activation Function) เพื่อสร้างความสัมพันธ์ไม่เชิงเส้นระหว่างตัวแปร ก่อนส่งต่อไปยังชั้นผลลัพธ์ที่ให้ค่าพยากรณ์ความต้องการใช้ไฟฟ้าในอนาคต แบบจำลองโครงข่าย

ประสาทเทียมสามารถเรียนรู้จากข้อมูลในอดีตและปรับค่าพารามิเตอร์ภายในอย่างต่อเนื่อง เพื่อเพิ่มประสิทธิภาพและความแม่นยำของการพยากรณ์ได้อย่างเหมาะสม



ภาพที่ 3 ตัวอย่างการทำงานของ Neural Network

4.3 เทคนิคการถดถอยเชิงเส้น (Linear Regression)

เทคนิคนี้เป็นการสร้างสมการเชิงเส้นเพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรตาม (ปริมาณการใช้ไฟฟ้า) กับตัวแปรอิสระ สมการพื้นฐานของแบบจำลองสามารถเขียนได้ดังนี้ [12]

$$Y = \beta_0+\beta_1X_1+\beta_2X_2+...+\beta_nX_n+\varepsilon \qquad (1)$$

โดยที่ Y คือ ค่าความต้องการใช้ไฟฟ้าที่คาดการณ์ได้

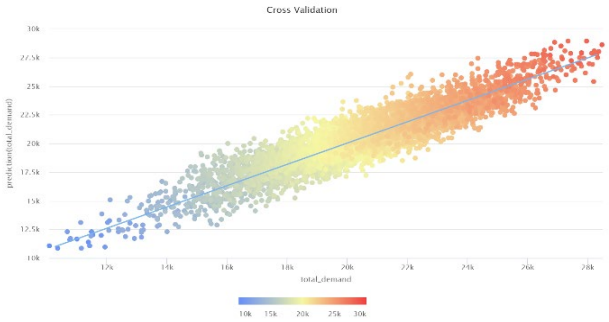
β_0 คือ ค่าคงที่ หรือ จุดตัดแกน (Intercept)

$\beta_1, \beta_2, ..., \beta_n$ คือ ค่าสัมประสิทธิ์ของตัวแปรต้น

$X_1, X_2, ..., X_n$ คือ ตัวแปรอิสระ

ε (epsilon) คือ ค่าความคลาดเคลื่อน หรือ ค่าความผิดพลาด (Error Term / Residual)

จากภาพที่ 4 แสดงความสัมพันธ์ระหว่างค่าความต้องการใช้ไฟฟ้าจริง (total_demand) บนแกน X และค่าที่โมเดลพยากรณ์ได้ (prediction(total_demand)) บนแกน Y โดยใช้เทคนิค Linear Regression ซึ่งเป็นแบบจำลองเชิงเส้นที่นิยมใช้ในการพยากรณ์ข้อมูลเชิงปริมาณ จากภาพจะเห็นได้ว่าจุดข้อมูล (scatter points) มีการกระจายตัวอยู่ใกล้กับเส้นสมการเชิงเส้น (regression line) ซึ่งแสดงถึงแนวโน้มเชิงบวกระหว่างค่าจริงและค่าที่พยากรณ์ได้ เมื่อความต้องการใช้ไฟฟ้าจริงเพิ่มขึ้น ค่าที่โมเดลพยากรณ์ก็เพิ่มขึ้นตามไปด้วยในทิศทางเดียวกัน สะท้อนว่าโมเดลสามารถอธิบายความสัมพันธ์ของข้อมูลได้อย่างเหมาะสม จุดแต่ละจุดในกราฟถูกระบายสีตามค่าความต้องการไฟฟ้า โดยใช้ไล่ระดับสีจากสีน้ำเงิน เหลือง ไปจนถึงแดง ซึ่งช่วยให้เห็นการกระจายของข้อมูลในแต่ละช่วงได้ชัดเจนมากขึ้น โดยสีน้ำเงินแทนค่าความต้องการไฟฟ้าระดับต่ำประมาณ 10,000–15,000 หน่วย สีเหลืองแทนค่าความต้องการระดับกลาง ประมาณ 18,000–22,000 หน่วย และสีแดงแทนค่าความต้องการสูงมากกว่า 25,000 หน่วยขึ้นไป ลักษณะของกราฟที่ข้อมูลเรียงตัวใกล้เส้นพยากรณ์นี้บ่งบอกว่าโมเดลมีความแม่นยำในระดับที่ดีและสามารถอธิบายความสัมพันธ์ระหว่างตัวแปรได้อย่างเหมาะสม โดยเฉพาะเมื่อผ่านการตรวจสอบความถูกต้องแบบ (Cross Validation) ซึ่งช่วยลดความเอนเอียงของแบบจำลอง เพิ่มความน่าเชื่อถือของผลการพยากรณ์ และยืนยันว่าโมเดลสามารถนำไปใช้พยากรณ์ความต้องการใช้ไฟฟ้าในอนาคตได้อย่างมีประสิทธิภาพ



ภาพที่ 4 ตัวอย่างการทำงานของ Linear Regression

5. การประเมินประสิทธิภาพของการพยากรณ์ค่า

การประเมินประสิทธิภาพของแบบจำลองพยากรณ์มีความสำคัญอย่างยิ่ง เนื่องจากช่วยตรวจสอบความถูกต้องและความแม่นยำของผลการคาดการณ์ โดยเปรียบเทียบค่าที่แบบจำลองพยากรณ์ได้กับค่าจริง เพื่อดูระดับความคลาดเคลื่อนของผลลัพธ์

ในการศึกษานี้ได้ใช้เทคนิค Cross Validation เพื่อประเมินประสิทธิภาพของแบบจำลองการพยากรณ์ โดยแบ่งข้อมูลออกเป็นส่วน ๆ สำหรับการฝึกสอนและทดสอบโมเดลอย่างเป็นระบบภาพแสดงกระบวนการของ Cross Validation Operators จำนวน 3 ชุด ซึ่งใช้กับแบบจำลอง Neural Network, Linear Regression และRandom Forest โดยแต่ละชุดมีโครงสร้าง

เหมือนกัน ประกอบด้วยสองขั้นตอนหลัก คือ การฝึกสอนโมเดล (Training) และ การทดสอบโมเดล (Testing) ในขั้นตอนการฝึกสอน ใช้ข้อมูลชุดฝึก (Train set) เพื่อสร้างแบบจำลอง เช่น การใช้ Neural Net Operator โดยตั้งค่าชั้นซ่อน (Hidden Layers) จำนวน 4 ชั้น ส่วนขั้นตอนการทดสอบ ใช้ Apply Model Operator เพื่อทำนายข้อมูลชุดทดสอบ และประเมินผลด้วย Performance (Regression) เพื่อวัดค่าประสิทธิภาพของแบบจำลองในด้านต่าง ๆ การตั้งค่า Cross Validation กำหนดจำนวนรอบตรวจสอบ (Number of folds) เท่ากับ 10 หมายถึงการแบ่งข้อมูลเป็น 10 ส่วน พร้อมกำหนดค่า Random Seed เพื่อให้ผลลัพธ์สามารถทำซ้ำได้ หลังจากดำเนินการ Cross Validation กับทั้งสามโมเดลแล้ว ได้นำค่าประสิทธิภาพที่ได้มาเปรียบเทียบ และปรับค่าพารามิเตอร์ด้วย Optimize Parameter เพื่อค้นหาค่าที่เหมาะสมที่สุด จากนั้นส่งผลลัพธ์ผ่านพอร์ตแสดงผลเพื่อสรุปค่าประสิทธิภาพรวมของแต่ละแบบจำลอง สรุปได้ว่า เทคนิค Cross Validation (10-fold) มีบทบาทสำคัญในการประเมินและเปรียบเทียบความแม่นยำและความเสถียรของแบบจำลองทั้งสาม เพื่อคัดเลือกแบบจำลองที่มีประสิทธิภาพสูงสุดในการพยากรณ์ความต้องการใช้ไฟฟ้าในประเทศไทย

5.1 ค่าความคลาดเคลื่อนรากที่สองของความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error: RMSE)

ค่าความคลาดเคลื่อนรากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย ใช้บ่งชี้ระดับความแตกต่างระหว่างค่าที่แบบจำลองพยากรณ์ได้กับค่าจริง ยิ่งค่ามีค่าน้อย แสดงว่าแบบจำลองสามารถพยากรณ์ได้แม่นยำมากขึ้น [17]

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

โดยที่ n= จำนวนข้อมูลทั้งหมด

y_i= ค่าจริง (Actual value) ของตัวอย่างที่ i

$\hat{y}_i$= ค่าที่แบบจำลองพยากรณ์ได้ (Predicted value) ของตัวอย่างที่ i

5.2 ค่าความคลาดเคลื่อนเฉลี่ยยกกำลังสอง (Mean Squared Error)

ค่าความคลาดเคลื่อนเฉลี่ยในรูปยกกำลังสอง เป็นตัวชี้วัดความถูกต้องของแบบจำลองในภาพรวม โดยค่าที่ต่ำกว่าหมายถึงแบบจำลองมีความผิดพลาดในการพยากรณ์น้อย [24]

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (3)$$

โดยที่ n= จำนวนข้อมูลทั้งหมด

y_i= ค่าจริง (Actual value) ของตัวอย่างที่ i

$\hat{y}_i$= ค่าที่แบบจำลองพยากรณ์ได้ (Predicted value) ของตัวอย่างที่ i

5.3 ค่าความคลาดเคลื่อนสัมบูรณ์ (Absolute Error: AE)

ค่าความคลาดเคลื่อนสัมบูรณ์ระหว่างค่าจริง (Actual value) และค่าที่แบบจำลองพยากรณ์ได้ (Predicted value) ซึ่งใช้วัดว่าผลการพยากรณ์ของแบบจำลองมีความแตกต่างจากค่าจริงมากน้อยเพียงใดโดยไม่สนทิศทางของความคลาดเคลื่อน [15]

$$\mathrm{AE}_i = | y_i - \hat{y}_i | \quad (4)$$

โดยที่ y_i= ค่าจริง (Actual value) ของตัวอย่างที่ i

$\hat{y}_i$= ค่าที่แบบจำลองพยากรณ์ได้ (Predicted value)

$|\cdot|$= ค่าสัมบูรณ์ (Absolute value)

5.4 ความคลาดเคลื่อนยกกำลังสอง (Squared Error: SE)

เป็นตัวชี้วัดที่ใช้ประเมินความแตกต่างระหว่างค่าจริงและค่าที่พยากรณ์จากโมเดลในงานวิเคราะห์ข้อมูลและการเรียนรู้ของเครื่อง (machine learning) โดยเฉพาะในงานที่เกี่ยวข้องกับการพยากรณ์ค่าต่อเนื่อง เช่น การถดถอยเชิงเส้น หรือโมเดลเชิงเส้นอื่นๆ [16]

$$\mathrm{SE}_i = (y_i - \hat{y}_i)^2 \quad (5)$$

โดยที่ y_i= ค่าจริง (Actual value) ของตัวอย่างที่ i

$\hat{y}_i$= ค่าที่แบบจำลองพยากรณ์ได้ (Predicted value)

5.5 ค่าสัมประสิทธิ์การตัดสินใจ (R-Squared (R^2))

เป็นค่าที่บ่งชี้ว่าแบบจำลองสามารถอธิบายความแปรปรวนของข้อมูลได้มากเพียงใน โดยค่าของ R^2 ยิ่งเข้าใกล้ 1 แสดงถึงความแม่นยำของแบบจำลองที่สูง ซึ่งคำนวณจากสมการ[18]

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (6)$$

โดยที่ y_i = ค่าจริง (Actual value) ของตัวอย่างที่ i

$\hat{y}_i$ = ค่าที่แบบจำลองพยากรณ์ได้ (Predicted value)

$\bar{y}$ = ค่าเฉลี่ยของค่าจริงทั้งหมด

6. การนำไปใช้ (Deployment)

จากผลการวิจัยพบว่าแบบจำลองเทคนิคต้นไม้ป่าสุ่ม เป็นแบบจำลองที่มีประสิทธิภาพสูงที่สุดในการพยากรณ์ความต้องการใช้ไฟฟ้าในประเทศไทย โดยให้ค่าความแม่นยำสูงและมีความคลาดเคลื่อนต่ำที่สุดเมื่อเทียบกับแบบจำลองอื่น[21][22] ซึ่งสามารถนำผลลัพธ์จากแบบจำลองนี้ไปประยุกต์ใช้ได้จริงในหลายด้าน โดยเฉพาะในงานวางแผนและบริหารจัดการพลังงานของประเทศ ในทางปฏิบัติ หน่วยงานที่เกี่ยวข้องกับพลังงาน เช่น การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.) การไฟฟ้านครหลวง (กฟน.) หรือการไฟฟ้าส่วนภูมิภาค (กฟภ.) สามารถนำแบบจำลองนี้ไปใช้ช่วยพยากรณ์แนวโน้มการใช้ไฟฟ้าในแต่ละช่วงเวลาได้อย่างแม่นยำมากขึ้น[19][20] ข้อมูลที่ได้จะช่วยให้การจัดสรรกำลังการผลิตเป็นไปอย่างมีประสิทธิภาพ ป้องกันปัญหาการผลิตไฟฟ้าเกินความต้องการหรือขาดแคลนไฟฟ้าในบางช่วงเวลา และสามารถแสดงผลการพยากรณ์ผ่านแดชบอร์ดหรือเว็บไซต์ [23] เพื่อให้ผู้บริหารหรือผู้มีส่วนเกี่ยวข้องสามารถเข้าถึงข้อมูลและตัดสินใจได้อย่างรวดเร็วและแม่นยำ นอกจากนี้ผลการพยากรณ์จากแบบจำลองยังสามารถนำไปใช้ในระดับนโยบาย เพื่อช่วยกำหนดแนวทางในการวางแผนพลังงานของประเทศ [20] เช่น การประเมินปริมาณการใช้ไฟฟ้าในอนาคตเพื่อเตรียมความพร้อมในการก่อสร้างโรงไฟฟ้าใหม่ การวางแผนพลังงานหมุนเวียนให้สอดคล้องกับความต้องการในแต่ละช่วงเวลา และการบริหารจัดการทรัพยากรพลังงานให้เกิดประโยชน์สูงสุด การนำแบบจำลองไปใช้งานจริงจะช่วยให้หน่วยงานด้านพลังงานสามารถบริหารจัดการระบบไฟฟ้าได้อย่างมีประสิทธิภาพมากขึ้น

ผลการวิจัย

ผู้วิจัยได้ทำการเปรียบเทียบค่าที่ใช้ทดสอบประสิทธิภาพของตัวแบบการพยากรณ์การใช้ไฟฟ้าในประเทศไทยด้วยเทคนิคการเรียนรู้ของเครื่องมือทั้ง 3 เทคนิคได้แก่ เทคนิคต้นไม้ป่าสุ่ม (Random Forest) เทคนิคโครงข่ายประสาทเทียม (Neural Network) และเทคนิคการถดถอยเชิงเส้น (Linear Regression) และได้มีการทำการทดสอบประสิทธิภาพของการพยากรณ์ค่าด้วยค่าความคลาดเคลื่อนรากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย(RMSE), ค่าความคลาดเคลื่อนเฉลี่ยในรูปกำลังสอง (MSE), ค่าความคลาดเคลื่อนสัมบูรณ์ (AE), ค่าความคลาดเคลื่อนกำลังสอง(SE), ค่าสัมประสิ(ทธิ์การตัดสินใจพหุคูณ (R^2) ซึ่งผลการวิเคราะห์ข้อมูลได้ผลแสดงดังตารางที่1

ตารางที่ 2 การเปรียบเทียบค่าทดสอบประสิทธิภาพตัวแบบ

Regression Technique	Regression Performance				
	RMSE	MSE	AE	SE	R^2
Random Forest**	514.03	264225.81	399.71	2643 36.33	96.7%**
Neural Network	716.29	513065.63	565.74	5133 87.09	93.6%
Linear Regression	787.13	619567.34	621.36	6199 70.17	92.1%

หมายเหตุ ** คือ ค่าเทคนิคที่มีความเหมาะสมสำหรับนำมาสร้างตัวแบบการพยาการณ์ความต้องการใช้ไฟฟ้าในประเทศไทย

จากตารางที่ 2 แสดงผลการประเมินประสิทธิภาพของแบบจำลองการพยากรณ์การใช้ไฟฟ้าในประเทศไทยด้วยเทคนิคการเรียนรู้ของเครื่องทั้งสามแบบ ได้แก่ ต้นไม้ป่าสุ่ม (Random Forest), เทคนิคโครงข่ายประสาทเทียม (Neural Network) และเทคนิคการถดถอยเชิงเส้น (Linear Regression) โดยใช้ตัวชี้วัดทางสถิติ 5 ค่า ได้แก่ RMSE, MSE, AE, SE และ R^2 เพื่อเปรียบเทียบความแม่นยำและความเสถียรของแบบจำลอง ผลการวิเคราะห์พบว่า แบบจำลอง Random Forest ให้ผลลัพธ์ที่ดีที่สุด โดยมีค่า RMSE ต่ำที่สุดเท่ากับ 514.03 แสดงถึงความคลาดเคลื่อนระหว่างค่าที่พยากรณ์ได้กับค่าจริงในระดับต่ำ และมีค่า R^2 สูงสุดที่ 96.70% ซึ่งสะท้อนว่าแบบจำลองสามารถอธิบายความแปรปรวนของข้อมูลได้ถึงเกือบทั้งหมด ทำให้ Random Forest มีความแม่นยำและความเสถียรสูงที่สุดในบรรดาแบบจำลองทั้งหมด ในขณะที่ แบบจำลองNeural Network

(NN) ให้ค่า R^2 เท่ากับ 93.60% และ RMSE เท่ากับ 716.29 แม้ค่าความคลาดเคลื่อนจะสูงกว่า Random Forest เล็กน้อย แต่ยังถือว่ามีประสิทธิภาพที่ดี สามารถเรียนรู้รูปแบบข้อมูลที่ไม่เป็นเชิงเส้นได้ในระดับหนึ่ง ส่วน แบบจำลอง Linear Regression (LR) ซึ่งเป็นวิธีการเชิงเส้น ให้ค่า R^2 เท่ากับ 92.10% และ RMSE เท่ากับ 787.13 ถือว่ามีความแม่นยำน้อยที่สุดเมื่อเทียบกับอีกสองเทคนิค สาเหตุเนื่องจากสมมติฐานเชิงเส้นของแบบจำลองไม่สามารถจับความซับซ้อนของข้อมูลการใช้ไฟฟ้าที่ได้รับอิทธิพลจากหลายปัจจัยพร้อมกันได้ ผลการเปรียบเทียบชี้ให้เห็นว่า Random Forest เป็นเทคนิคที่มีประสิทธิภาพสูงสุดในการพยากรณ์ความต้องการใช้ไฟฟ้าในประเทศไทย เนื่องจากสามารถจัดการกับความซับซ้อนของข้อมูลขนาดใหญ่ได้ดี มีความสามารถในการลดการเกิด Overfitting และให้ค่าความคลาดเคลื่อนต่ำสุด ซึ่งทำให้เหมาะสมต่อการนำไปใช้ในงานวางแผนและบริหารการจัดการพลังงานในระดับประเทศอย่างมีประสิทธิภาพ

ตารางที่ 3 ผลการวิเคราะห์สมการถดถอยพหุคูณ (Multiple Linear Regression Analysis Table)

Attribute	COEF	Std. Error	t-Stat	p-Value
Malaysia_IMP	-8.458	0.651	-12.999	0.000**
North_Generation	1.252	0.022	57.189	0.000**
Metropolitan Generation	1.244	0.025	49.771	0.000**
Central_ Generation	0.901	0.005	193.231	0.000**
Laos_IMP	3.305	0.060	55.520	0.000**
Hour_Number	18.439	1.489	12.382	0.000**
Season	297.377	14.803	20.090	0.000**
Intercept	499.614	87.315	5.722	0.000

หมายเหตุ: ** หมายถึง ผลการวิเคราะห์ที่มีนัยสำคัญทางสถิติที่ระดับ 0.005 แสดงว่าตัวแปรมีอิทธิพลต่อแบบจำลองอย่างมีนัยสำคัญ

จากตารางที่ 3 ผลการวิเคราะห์สมการถดถอยพหุคูณ (Multiple linear regression) แสดงถึงความสัมพันธ์ระหว่างตัวแปรอิสระได้แก่ การนำเข้าไฟฟ้าจากประเทศมาเลเซีย (Malaysia_IMP) การผลิตไฟฟ้าภาคเหนือ (North_generation) การผลิตไฟฟ้าในเขตมหานคร (Metropoltan_generation) การผลิตไฟฟ้าภาคกลาง (Central_generation) การนำเข้าไฟฟ้าจากประเทศลาว (Laos_IMP) ช่วงเวลาของวัน (Hour_number) และฤดูกาล (Season) ต่อค่าความต้องการใช้ไฟฟ้ารวมของประเทศไทย โดยผลการวิเคราะห์พบว่าตัวแปรอิสระทั้งหมดมีค่า P-Value เท่ากับ 0.000 ซึ่งน้อยกว่า 0.05 แสดงว่าทุกตัวแปรมีนัยสำคัญทางสถิติในการอธิบายการเปลี่ยนแปลงของค่าความต้องการใช้ไฟฟ้าในระดับความเชื่อมั่น 95 เปอร์เซ็นต์ ผลลัพธ์จากค่าความสัมพันธ์เชิงสถิติ (Coefficlent) พบว่าตัวแปรส่วนใหญ่มีค่าสัมประสิทธิ์เป็นบวก แสดงถึงความสัมพันธ์เชิงบวกระหว่างตัวแปรกับค่าความต้องการใช้ไฟฟ้า โดยเฉพาะการผลิตไฟฟ้าในภาคเหนือ ภาคกลาง และพื้นที่มหานคร ซึ่งมีค่าสัมประสิทธิ์เท่ากับ 1.252, 0.901 และ 1.244 ตามลำดับ หมายความว่าการผลิตไฟฟ้าในภูมิภาคเหล่านี้เพิ่มขึ้น 1 หน่วย จะส่งผลให้ความต้องการใช้ไฟฟ้าเพิ่มขึ้นประมาณ 1 หน่วยเช่นกัน ทั้งนี้ภาคกลางมีค่าความสัมพันธ์สูงที่สุด สะท้อนให้เห็นว่าพื้นที่ภาคกลางเป็นศูนย์กลางของความต้องการใช้ไฟฟ้าของประเทศ นอกจากนี้ การนำเข้าไฟฟ้าจากประเทศลาวมีค่าสัมประสิทธิ์เป็นบวกเท่ากับ 3.305 แสดงถึงการเพิ่มขึ้นของปริมาณไฟฟ้าที่นำเข้าซึ่งสัมพันธ์กับความต้องการใช้ไฟฟ้าที่เพิ่มสูงขึ้น ในทางตรงกันข้าม การนำเข้าพลังงานไฟฟ้าจากประเทศมาเลเซียมีค่าสัมประสิทธิ์เป็นลบเท่ากับ -8.458 บ่งชี้ถึงความสัมพันธ์เชิงลบ ซึ่งอาจเกิดจากรูปแบบการส่งออกหรือการชดเชยพลังงานระหว่างระบบไฟฟ้าในบางช่วงเวลา ส่วนตัวแปรช่วงเวลาของวัน (Hour_number) มีค่าสัมประสิทธิ์ 18.439 แสดงให้เห็นว่าช่วงเวลามีผลต่อการใช้ไฟฟ้า โดยเฉพาะในช่วงเวลาที่มีกิจกรรมของภาคครัวเรือนและภาคธุรกิจสูง ขณะที่ตัวแปรฤดูกาล (Season) มีค่าสัมประสิทธิ์สูงถึง 297.377 ซึ่งสะท้อนว่า ฤดูกาลมีผลอย่างมากต่อความต้องการใช้ไฟฟ้า โดยเฉพาะในช่วงฤดูร้อนที่มีการใช้เครื่องปรับอากาศเพิ่มขึ้นอย่างมากทั่วประเทศ สำหรับค่า Intercept T เท่ากับ 499.614 หมายถึงค่าความต้องการใช้ไฟฟ้าพื้นฐานเมื่อค่าของตัวแปรอิสระทั้งหมดเป็นศูนย์ แสดงให้เห็นว่ามีระดับการใช้ไฟฟ้าขั้นต่ำที่เกิดขึ้นโดยไม่ขึ้นอยู่กับตัวแปรใด ๆ ผลลัพธ์จากการวิเคราะห์นี้แสดงให้เห็นว่าแบบจำลองถดถอยพหุคูณสามารถอธิบายความสัมพันธ์ระหว่างตัวแปรด้านภูมิภาค การนำเข้าไฟฟ้า ช่วงเวลา และฤดูกาลต่อความต้องการใช้ไฟฟ้าในประเทศไทยได้อย่างมีนัยสำคัญทางสถิติ ซึ่งสามารถนำผลลัพธ์ไปใช้เป็นพื้นฐานในการพัฒนาแบบจำลองการพยากรณ์ขั้นสูง เช่น Random forest หรือ Neural network

เพื่อเพิ่มประสิทธิภาพและความแม่นยำของการคาดการณ์การใช้พลังงานไฟฟ้าในอนาคตของประเทศได้ต่อไป

สรุปผลการวิจัย

งานวิจัยเรื่อง การวิเคราะห์และพยากรณ์การใช้ไฟฟ้าในประเทศไทยด้วยแบบจำลองเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) นี้มีวัตถุประสงค์หลักเพื่อ ศึกษาและเปรียบเทียบประสิทธิภาพ ของเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ที่ได้รับความนิยม 3 รูปแบบ ได้แก่ ต้นไม้ป่าสุ่ม (Random forest), โครงข่ายประสาทเทียม (Neural network: NN), และ การถดถอยเชิงเส้น (Linear Regression: LR) เพื่อประยุกต์ใช้ในการพยากรณ์ความต้องการใช้พลังงานไฟฟ้าในประเทศไทยอย่างแม่นยำ โดยใช้ชุดข้อมูลจริงที่เก็บรวบรวมได้จำนวนทั้งสิ้น 8,806 รายการ ซึ่งครอบคลุมตัวแปรสำคัญที่ส่งผลต่อการใช้ไฟฟ้า เช่น ปริมาณการผลิตไฟฟ้าในภูมิภาค การนำเข้าพลังงานไฟฟ้า ฤดูกาล และช่วงเวลาของวัน เมื่อพิจารณาจากตัวชี้วัดความแม่นยำทางสถิติที่หลากหลาย แบบจำลอง Random forest ได้พิสูจน์แล้วว่าให้ผลการพยากรณ์ที่โดดเด่นและมีความแม่นยำสูงที่สุดเหนือกว่าเทคนิคอื่น ๆ อย่างชัดเจนความแม่นยำของแบบจำลอง Random forest ถูกยืนยันด้วย ค่าสัมประสิทธิ์การตัดสินใจ (R^2) ที่สูงถึง 96.70% ซึ่งเป็นตัวเลขที่มีนัยสำคัญอย่างยิ่ง เนื่องจากมันแสดงให้เห็นว่าแบบจำลองนี้สามารถอธิบายและจับความแปรปรวน (Variation) ของความต้องการใช้ไฟฟ้าในชุดข้อมูลได้เกือบทั้งหมด และมีความสามารถในการทำนายที่เชื่อถือได้สูงสุด นอกจากนี้ แบบจำลอง Random forest ยังมี ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยราก (RMSE) ต่ำที่สุดที่ 514.03 ซึ่งเป็นดัชนีสำคัญที่บ่งชี้ว่าค่าที่แบบจำลองพยากรณ์ได้มีความเบี่ยงเบนจากค่าความต้องการใช้ไฟฟ้าจริงน้อยที่สุด ในขณะเดียวกัน ประสิทธิภาพของแบบจำลองอื่น ๆ ก็ยังคงอยู่ในระดับดีเยี่ยม โดย แบบจำลอง Neural network ตามมาเป็นลำดับที่สอง ด้วย ค่า R^2 เท่ากับ 93.60% ซึ่งแสดงให้เห็นถึงความสามารถที่ดีเยี่ยมในการเรียนรู้ความสัมพันธ์แบบไม่เป็นเชิงเส้นที่ซับซ้อนของข้อมูล และสุดท้าย แบบจำลอง Linear Regression ซึ่งเป็นเทคนิคที่เรียบง่ายที่สุด ก็ยังคงให้ผลลัพธ์ที่น่าพอใจ โดยมี ค่า R^2 เท่ากับ 92.10% ซึ่งเป็นการยืนยันว่าปัจจัยที่ใช้ในการวิเคราะห์มีความสัมพันธ์เชิงเส้นที่ชัดเจนและมีนัยสำคัญต่อความต้องการใช้ไฟฟ้าโดยรวม

อภิปรายการวิจัย

ผลการศึกษาครั้งนี้สอดคล้องกับงานวิจัยก่อนหน้า โดยงานวิจัยของ ภูมิพัฒน์ หงส์ธนากรหิรัญ เรื่อง การทำนายการผลิตพลังงานไฟฟ้าจากเซลล์แสงอาทิตย์ด้วยการเรียนรู้ของเครื่อง พบว่าแบบจำลอง Random Forest (RF) สามารถให้ค่าความแม่นยำของการพยากรณ์อยู่ที่ 94% แสดงให้เห็นถึงศักยภาพของโมเดลในการวิเคราะห์ข้อมูลเชิงซับซ้อน นอกจากนี้งานวิจัยของ อังคณา ชมภูศรี เรื่อง การประเมินศักยภาพพลังงานไฟฟ้าโซลาร์เซลล์บนหลังคาในเขตพื้นที่เมืองด้วยแบบจำลองความสูงของชั้นเรือนยอด พบว่าแบบจำลอง Random Forest ให้ค่าความแม่นยำเท่ากับ 77.581% (ประมาณ 77%) ซึ่งแม้จะต่ำกว่าผลลัพธ์ในงานวิจัยนี้ที่ได้ค่าความแม่นยำ 96.7% แต่ก็ยังแสดงให้เห็นว่า Random Forest เป็นแบบจำลองที่มีประสิทธิภาพสูงในการพยากรณ์ข้อมูลด้านพลังงาน เมื่อเปรียบเทียบผลการศึกษาทั้งหมดสามารถสรุปได้ว่า แบบจำลอง Random Forest (RF) เป็นโมเดลที่เหมาะสมที่สุดในการวิเคราะห์และพยากรณ์ความต้องการใช้ไฟฟ้าในประเทศไทย เนื่องจากให้ค่าความแม่นยำสูงสุดและสามารถประมวลผลความสัมพันธ์ของข้อมูลได้อย่างมีประสิทธิภาพ

ข้อเสนอแนะ

1.หากมีผู้ที่สนใจจะนำแนวคิดการสร้างตัวแบบ Machine Learning ไปประยุกต์ใช้ในการพยากรณ์ปัญหาอื่น ๆ ในอนาคต จำเป็นต้องตระหนักว่าชุดข้อมูลที่ใช้ และตัวแปรที่เกี่ยวข้องทั้งหมดจะมีความแตกต่างอย่างสิ้นเชิง

2.ผลการวิเคราะห์และแบบจำลองการพยากรณ์ที่ได้จากงานวิจัยนี้ Random forest, Neural network, Linear Regression ถูกพัฒนาและประเมินประสิทธิภาพโดยอ้างอิงจาก ชุดข้อมูลความต้องการใช้ไฟฟ้าในประเทศไทยเท่านั้น ดังนั้น ผลลัพธ์ที่ได้จึงมีความจำเพาะเจาะจงกับบริบทของปัญหาด้านพลังงาน

คำชี้แจงการใช้ AI

ผู้เขียนขอยืนยันว่า การใช้เทคโนโลยีปัญญาประดิษฐ์ (AI) ในงานวิจัยฉบับนี้ไม่ได้ส่งผลกระทบต่อความถูกต้องของกระบวนการวิจัย และองค์ความรู้ทั้งหมดในบทความยังคงเป็นความรับผิดชอบของผู้เขียนแต่เพียงผู้เดียว ไม่มีการใช้เครื่องมือ AI เพื่อสร้างสมมติฐานทางวิทยาศาสตร์หรือแปลผลการทดลอง โดยปราศจากการกำกับดูแลจากมนุษย์ ต้นฉบับนี้มีเนื้อหาที่สร้าง

ขึ้นหรือได้รับการสนับสนุนจากเครื่องมือปัญญาประดิษฐ์ ได้แก่ ChatGPT (OpenAI, 2025) และGemini ซึ่งถูกนำมาใช้เพื่อช่วยในการดำเนินงาน เช่น การปรับแก้ภาษา การเรียบเรียงข้อความให้อยู่ในรูปแบบทางวิชาการ การช่วยอธิบายผลลัพธ์เชิงเทคนิค และการจัดรูปแบบเอกสาร ทั้งนี้ผู้เขียนได้ตรวจสอบเนื้อหาที่สร้างโดย AI อย่างรอบคอบ เพื่อให้มั่นใจในความถูกต้อง ความเป็นต้นฉบับ และการปฏิบัติตามหลักจริยธรรมทางวิชาการทุกประการ

เอกสารอ้างอิง

[1] สำนักงานนโยบายและแผนพลังงาน. (2566). รายงานแนวโน้มการใช้พลังงานของประเทศไทย ปี 2566–2570. กระทรวงพลังงาน. https://www.eppo.go.th/index.php/th/

[2] การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย. (2565). แผนพัฒนากำลังผลิตไฟฟ้าของประเทศไทย พ.ศ. 2561–2580 (PDP2018 Revision 1). https://www.egat.co.th/home/

[3] สำนักงานนโยบายและแผนพลังงาน. (2567). รายงานสถิติพลังงานประเทศไทย ปี 2567. กระทรวงพลังงาน https://www.eppo.go.th/index.php/th/

[4] Ryu, S., Noh, J., & Kim, H. (2017). Deep Neural Network-Based Demand Side Short Term Load Forecasting. Energies, 10(1), https://www.mdpi.com/1996-1073/10/1/3

[5] Kermanshahi, B., & Iwamiya, H. (2011). Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach. Energies, 4(8), 1246–1257. https://www.mdpi.com/1996-1073/4/8/1246

[6] Chujai, P., & Kerdprasop, K. (2022). Forecasting Daily Electricity Consumption in Thailand Using Regression, Artificial Neural Network, Support Vector Machine, and Hybrid Models. Energies, 15(9), 3105. https://www.mdpi.com/1996-1073/15/9/3105

[7] Ahmed, I., & Khan, S. (2023). Hybrid Machine Learning Model for Electricity Consumption Prediction Using Random Forest and Artificial Neural Networks. Journal of Electrical and Computer Engineering, 2023, Article ID 1562942. https://onlinelibrary.wiley.com/doi/10.1155/2022/1562942

[8] การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย. (2023). รายงานประจำปี 2566. สืบค้นจาก https://www.egat.co.th/home/annual-report

[9] ระบบไฟฟ้าไทย: การผลิตไฟฟ้ารายชั่วโมง ความต้องการใช้ไฟฟ้า และกระแสไฟฟ้าข้ามพรมแดน สืบค้น จาก:https://zenodo.org/records/17109911

[10] สืบค้นจาก https://mljar.com/glossary/random-forest/

[11] Big Data Institute (ประเทศไทย). (2024). *เตรียมข้อมูลอย่างไรให้โมเดล Linear Regression ทำงานได้อย่างถูกต้อง.* สืบค้นจาก: https://bdi.or.th/big-data-101/data-preparation-for-linear-regression/

[12] เตรียมข้อมูลอย่างไรให้โมเดล Linear Regression สืบค้นจาก https://bdi.or.th/big-data-101/data-preparation-for-linear-regression/

[13] เอกสารคู่มือการใช้งาน RapidMiner. (2024). RapidMiner Studio Tutorial. สืบค้นจาก: https://docs.rapidminer.com/latest/studio

[14] การวิเคราะห์ทางสถิติเมื่อข้อมูลมีค่าขาดหาย Little, R. J. A., & Rubin, D. B. (2019). Statistical Analysis with Missing Data. 3rd ed., Wiley. สืบค้นจาก: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119482260

[15] พรทิพย์ แซ่ตั้ง. (2564). การเปรียบเทียบแบบจำลองการพยากรณ์ความต้องการพลังงานไฟฟ้าในประเทศ

ไทยด้วยวิธีการทางสถิติและการเรียนรู้ของเครื่อง. วารสารวิชาการพระจอมเกล้าพระนครเหนือ, 31(3), 412–424.

[16] Calculator. (2025, March 27). Squared Error Calculator. สืบค้นจาก: https://shorturl.asia/lbMnc

[17] ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE)สืบค้าจาก https://statisticsbyjim.com/regression/root-mean-square-error-rmse/?utm_source=chatgpt.com

[18] GeeksforGeeks. (n.d.). Python Tutorial – Learn Python Programming Language. Retrieved October 13, 2025,สืนค้าจาก: https://www.geeksforgeeks.org/python/python-programming-language-tutorial/

[19] การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย. (2566). รายงานประจำปี 2566.สืบค้นจาก: https://www.egat.co.th/home/annual-report

[20] สำนักงานนโยบายและแผนพลังงาน. (2567). รายงานแนวโน้มการใช้พลังงานของประเทศไทย ปี 2566–2570.กระทรวงพลังงาน. สืบค้นจาก: https://www.eppo.go.th/index.php/th/]

[21] Ahmed, I., & Khan, S. (2023). Hybrid Machine Learning Model for Electricity Consumption Prediction Using Random Forest and Artificial Neural Networks. Journal of Electrical and Computer Engineering, 2023, Article ID 1562942. สืบค้นจาก: https://onlinelibrary.wiley.com/doi/10.1155/2022/1562942

[22] Chujai, P., & Kerdprasop, K. (2022). Forecasting Daily Electricity Consumption in Thailand Using Regression, Artificial Neural Network, Support Vector Machine, and Hybrid Models. Energies, 15(9), 3105. สืบค้นจาก: https://www.mdpi.com/1996-1073/15/9/3105

[23] RapidMiner Documentation. (2024). RapidMiner Studio Tutorial & Deployment. สืบค้นจาก: https://docs.rapidminer.com/latest/studio

[24] สืบค้นจาก : https://shorturl.asia/rAL5K

[25] Ryu, S., Noh, J., & Kim, H. (2017). *Deep Neural Network-Based Demand Side Short Term Load Forecasting.* Energies, 10(1). สืบค้นจาก: https://www.mdpi.com/1996-1073/10/1/3