บทความวิจัย (Research Article)

การทำนายสภาวะที่เหมาะสมในการผลิตไบโอดีเซลด้วยเทคนิคโครงข่ายประสาทและการจำลองการอบเหนียว

Optimal Condition Prediction of Biodiesel Production using Neural Network and Simulated Annealing Techniques

ธิติมา ตาสาโรจน์[1], กนกชัย พลแสน[1] และ บัณฑิต บุญขาว[2]*

[1] สาขาวิชาวิทยาศาสตร์ คณะครุศาสตร์ มหาวิทยาลัยนครพนม
[2] สาขาวิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนครพนม
* ผู้รับผิดชอบบทความ
bundit@npu.ac.th

Received: 14 Jan 2021
Revised: 13 Apr 2021
Accepted: 29 Apr 2021

บทคัดย่อ

บทความนี้เป็นการนำเทคนิคการขับเคลื่อนด้วยข้อมูลมาประยุกต์ใช้ในการทำนายสภาวะที่เหมาะสมของการสังเคราะห์ไบโอดีเซล เป็นที่ทราบกันดีว่าราคาของไบโอดีเซลนั้นไม่สามารถแข่งขันกับราคาน้ำมันดีเซลได้ เนื่องจากข้อจำกัดด้านประสิทธิภาพการสังเคราะห์ซึ่งขึ้นอยู่กับสภาวะเงื่อนไขการสังเคราะห์ แม้ว่าการออกแบบการทดลองเชิงสถิติสามารถนำไปใช้เพื่อหาสภาวะการสังเคราะห์ที่เหมาะสมและลดต้นทุนในการดำเนินการทางห้องปฏิบัติการให้น้อยที่สุดได้ แต่ก็ยังใช้เวลาและค่าใช้จ่ายในการดำเนินการทางห้องปฏิบัติการ ดังนั้นบทความนี้จึงเสนอแนวทางในการค้นหาสภาวะการสังเคราะห์ที่เหมาะสมโดยใช้เทคนิคการขับเคลื่อนด้วยข้อมูล เนื่องจากมีการรายงานสภาวะการสังเคราะห์ไบโอดีเซลเป็นจำนวนมากในวารสารและหนังสือทางวิชาการ ข้อมูลเหล่านี้สามารถนำมาใช้เป็นข้อมูลดิบสำหรับการเรียนรู้และฝึกเครื่องจักรเพื่อสร้างแบบจำลองที่จะนำไปใช้ในการทำนายหรือค้นหาสภาวะที่เหมาะสม ในบทความนี้ได้ใช้สภาวะการสังเคราะห์จำนวน 80 รายการและร้อยละของผลผลิตของการสังเคราะห์ไบโอดีเซลจากฐานข้อมูลบทความทางวิชาการถูกกำหนดให้เป็นข้อมูลป้อนเข้าและข้อมูลผลลัพธ์ตามลำดับ ข้อมูลดังกล่าวถูกใช้ในการฝึกการเรียนรู้ของเครื่องจักรโดยเลือกใช้เทคนิคโครงข่ายประสาทเป็นเทคนิคการเรียนรู้ของเครื่องจักร แบบจำลองที่ได้จากการเรียนรู้มีความคลาดเคลื่อนสูงสุดที่ร้อยละ 2.44 เมื่อทำการตรวจสอบความถูกต้องด้วยการทำนายสภาวะการสังเคราะห์ที่เป็นข้อมูลชุดใหม่จำนวน 8 รายการ จากนั้นแบบจำลองการเรียนรู้จะถูกนำไปใช้ในการทำนายและค้นหาสภาวะการสังเคราะห์ที่คาดว่าจะให้ร้อยละผลผลิตสูงสุดด้วยเทคนิคการหาค่าที่เหมาะสม โดยพบว่าจากการค้นหาสภาวะการสังเคราะห์ที่คาดว่าจะให้ร้อยละของผลผลิตสูงสุดนั้นจะมีค่าไม่เกินค่าสูงสุดที่ใช้ในการฝึกการเรียนรู้ของเครื่องจักร ซึ่งการค้นหาสภาวะที่เหมาะสมจะค้นหาอยู่เพียงแค่ภายในขอบเขตของข้อมูลที่ใช้ฝึกการเรียนรู้เท่านั้น ดังนั้นหากต้องการให้การค้นหาสภาวะที่เหมาะสมอยู่นอกเหนือจากข้อมูลที่ใช้ฝึกการเรียนรู้จะต้องหาวิธีการที่จะทำให้การค้นหาอยู่นอกสภาวะที่ใช้ในการฝึกการเรียนรู้ ปัญหาดังกล่าวถือว่าเป็นโจทย์ที่ท้าทายสำหรับการดำเนินการวิจัยต่อไป

คำสำคัญ: ไบโอดีเซล เทคนิคการขับเคลื่อนด้วยข้อมูล โครงข่ายประสาท การจำลองการอบเหนียว การหาสภาวะที่เหมาะสม

Abstract

In this article, an implementation of the data driven technique has been applied for optimal condition prediction of biodiesel synthesis. It is known that price of biodiesel still has not been

competed with petroleum diesel due to the efficiency synthesis limitation. Although, statistical design of experiment can be applied for optimising synthesis condition and minimising cost of investigation, it still consumes time and cost of investigation. Therefore, this article proposes a guideline for determining optimal synthesis condition using data driven technique. As there are many synthesis conditions reported in scientific journals and books, these data can be used as raw data for supervised learning. In here, 80 synthesis conditions and %yield of synthesis from the literature were assigned as input and output data, respectively. The data were trained by using neural network which was selected as supervised learning technique. The model obtained from the learning had the highest error of 2.44% when validated by eight new data synthesis conditions predictions. The learning model is then used to predict and find the synthesis conditions that are expected to produce the highest %yield with optimisation techniques. It was found that by searching for the expected synthesis conditions, the maximum %yield would not exceed the maximum value used in training machine due to the searching conditions were only within the domain of the training data. Therefore, if the search is to be opted out of the practice data, a method must be found to make the search outside the trained data. Such a problem is a challenging one for further research.

Keywords: biodiesel, data-driven technique, neural network, simulated annealing, optimisation

1. บทนำ

เมื่อเร็วๆ นี้เทคนิคการขับเคลื่อนด้วยข้อมูล (data-driven technique) ได้รับความสนใจในการนำไปประยุกต์ใช้กับกระบวนการทางเคมี อันเนื่องมาจากปัจจุบันมีข้อมูลจำนวนมหาศาลถูกบันทึกในทุก ๆ วินาที อย่างไรก็ตามด้วยข้อมูลดิบเพียงอย่างเดียวอาจไม่สามารถทำการค้นหาความรู้ที่ซ่อนอยู่ในข้อมูลได้ ดังนั้นจึงจำเป็นต้องอาศัยเทคนิคในการวิเคราะห์ข้อมูลเหล่านี้ เทคนิคทางสถิติเป็นเทคนิคพื้นฐานที่ใช้กันอย่างแพร่หลายในการวิเคราะห์ข้อมูล อย่างไรก็ตามด้วยข้อมูลที่มีขนาดใหญ่และมีความซับซ้อนเทคนิคทางสถิติพื้นฐานก็อาจไม่เพียงพอที่จะวิเคราะห์และสังเคราะห์ความรู้จากข้อมูลได้อย่างมีประสิทธิภาพ ในปัจจุบันการเรียนรู้เชิงลึก (deep learning) [1] ได้รับความสนใจเป็นอย่างมากในการวิเคราะห์และสกัดความรู้จากข้อมูลที่มีขนาดใหญ่และซับซ้อน การเรียนรู้เชิงลึกถูกพัฒนามาจากเทคนิคโครงข่ายประสาท (neural network; NN) โดยการเรียนรู้ข้อมูลที่มีอยู่ จากนั้นจึงสร้างแบบจำลองที่ได้จากการเรียนรู้ เทคนิคโครงข่ายประสาทอาจใช้เพื่ออธิบายพฤติกรรมของกระบวนการที่ซับซ้อนซึ่งไม่อาจสามารถอธิบายได้ด้วยแบบจำลองทางคณิตศาสตร์ เช่นกระบวนการสังเคราะห์สารทางเคมี

ด้วยข้อจำกัดของพฤติกรรมการสังเคราะห์ที่ไม่ทราบเงื่อนไขของสภาวะในการผลิตไบโอดีเซล บทความนี้จึงนำเสนอการใช้เทคนิคการขับเคลื่อนด้วยข้อมูลเป็นตัวขับเคลื่อนในการแก้ปัญหานี้ ไบโอดีเซลคือกรดไขมันเมทิลเอสเทอร์ (fatty acid methyl ester; FAME) ซึ่งสามารถสังเคราะห์ได้ภายใต้ปฏิกิริยาทรานเอสเทอรีฟิเคชัน (transesterification) โดยไตรกลีเซอไรด์ (น้ำมัน) จากสัตว์หรือพืชทำปฏิกิริยากับแอลกอฮอล์ภายใต้ตัวเร่งปฏิกิริยาจะแปรสภาพไปเป็นไบโอดีเซล (FAME) และกลีเซอรอล ซึ่งเป็นผลผลิตภัณฑ์ที่ได้จากการสังเคราะห์ เนื่องจากการเปลี่ยนไตรกลีเซอไรด์ไปเป็น FAME จะขึ้นอยู่กับเงื่อนไขสภาวะในการสังเคราะห์ เช่น อุณภูมิ ความเข้มข้น ชนิดตัวเร่งปฏิกิริยา เป็นต้น ซึ่งการปรับเปลี่ยนตัวแปรเหล่านี้จะส่งผลต่อร้อยละการผลิต FAME และแม้ว่าจะอาศัยเทคนิคการออกแบบการทดลอง (design of experiment; DOE) มาช่วยในการค้นหาสภาวะการสังเคราะห์ที่ให้ผลผลิตสูงที่สุดและทำการทดลองจำนวนน้อยที่สุด แต่การดำเนินการตามวิธีดังกล่าวยังคงมีค่าใช้จ่ายสูงในระหว่างที่ต้องทำการทดลองในห้องปฏิบัติการ

ดังนั้นในบทความนี้จะเสนอแนวทางในการค้นหาสภาวะที่เหมาะในการสังเคราะห์ไบโอดีเซลโดยใช้เทคนิคการเรียนรู้ของเครื่องจักร โดยข้อมูลที่ใช้เป็นตัวขับเคลื่อนนั้นจะอาศัยจากรายงานข้อมูลการสังเคราะห์ไบโอดีเซลซึ่งสามารถหาได้จาก

บทความทางวิชาการ เนื่องจากมีรายงานเงื่อนไขการสังเคราะห์อยู่เป็นจำนวนมากในวารสารวิชาการซึ่งสามารถนำมารวบรวมและใช้เป็นฐานข้อมูลในการสอนเครื่องจักร ตัวอย่างของการประยุกต์ใช้โครงข่ายประสาทกับการสังเคราะห์ไบโอดีเซล เช่น Baroutian และคณะ [2] ใช้โครงข่ายประสาทเทียม (artificial neural network; ANN) สำหรับทำนายความหนาแน่นของไบโอดีเซลจากน้ำมันปาล์ม เช่นเดียวกันกับกลุ่มของ Dorado [3-4] ใช้ NN เพื่อทำนายผลผลิตของไบโอดีเซลจากน้ำมันมะกอก นอกจากนี้ Bobadilla et. [5] ใช้เทคนิคอัลกอริทึมทางพันธุกรรม (genetic algorithm; GA) ร่วมกับเทคนิคเครื่องจักรสนับสนุนเวกเตอร์ (support vector machine; SVM) สำหรับการปรับสภาพไบโอดีเซลจากเสียน้ำมันปรุงอาหาร

อย่างไรก็ตาม จากการทบทวนวรรณกรรมจะพบว่างานวิจัยโดยส่วนใหญ่นั้นจะเน้นไปที่การสร้างแบบจำลองเพื่ออธิบายพฤติกรรมระหว่างเงื่อนไขสภาวะการสังเคราะห์ และการนำเอาแบบจำลองไปใช้ประโยชน์ก็เพียงเพื่อการทำนายว่าสภาวะที่สังเคราะห์จะให้ผลผลิตในปริมาณเท่าใด แต่ในบทความนี้จะเป็นนำเอาแบบจำลองที่สร้างขึ้นมานั้นเป็นแบบจำลองสำหรับการค้นหาสภาวะที่คาดว่าจะสามารถผลิตไบโอดีเซลได้ในปริมาณที่สูงที่สุดด้วยเทคนิคการหาค่าที่เหมาะสม (optimisation)

2. การเรียนรู้ของเครื่องจักรและการหาสภาวะที่เหมาะสม

ในหัวข้อนี้จะเป็นการเสนอแนวทางในการใช้เทคนิคการขับเคลื่อนด้วยข้อมูลเพื่อค้นหาสภาวะที่เหมาะสมในการสังเคราะห์ไบโอดีเซลด้วยปฏิกิริยาทรานส์เอสเทอริฟิเคชัน แนวทางที่นำเสนอแบ่งออกเป็น 3 ส่วน ได้แก่ การสร้างแบบจำลองด้วยโครงข่ายประสาท การทดสอบแบบจำลอง และการค้นหาสภาวะที่เหมาะสมด้วยเทคนิคการจำลองการอบเหนียว เทคนิคโครงข่ายประสาทถูกเลือกให้เป็นเทคนิคการเรียนรู้เครื่องจักรเพื่อสร้างแบบจำลองการสังเคราะห์ไบโอดีเซล โดยโครงข่ายประสาทจะได้รับการป้อนข้อมูลเป็นสภาวะการสังเคราะห์ไบโอดีเซล ซึ่งได้แก่ อุณหภูมิการสังเคราะห์ อัตราส่วนแอลกอฮอล์ต่อน้ำมัน ระยะเวลาในการทำปฏิกิริยา และความเข้มข้นของตัวเร่งปฏิกิริยา และร้อยละผลผลิต (%yield) เป็นข้อมูลผลผลิต จากนั้นแบบจำลองที่ได้จากการฝึกจะนำไปทดสอบความแม่นยำโดยการทดสอบการทำนายผลด้วยข้อมูลการสังเคราะห์ใหม่ และสุดท้ายแบบจำลองที่มีความแม่นยำจะถูกนำไปใช้ในการทำนายสภาวะที่เหมาะสมในการสังเคราะห์ไบโอดีเซล โดยการใช้เทคนิคการหาค่าที่เหมาะสมซึ่งการจำลองการอบเหนียว (simulated annealing; SA) จะถูกใช้เป็นเทคนิคในการหาสภาวะการสังเคราะห์ที่ให้ร้อยละผลผลิตสูงที่สุด

2.1 โครงข่ายประสาท

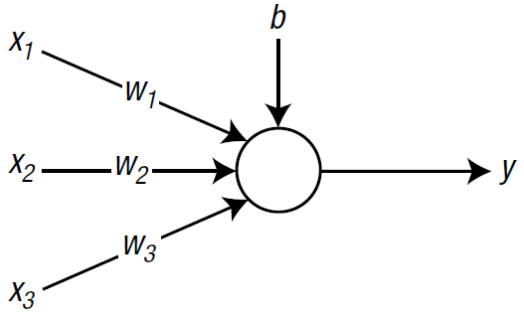
โครงข่ายประสาทถูกเลือกให้เป็นเทคนิคสำหรับการฝึกการเรียนรู้ของเครื่องจักรเพื่อสร้างแบบจำลองของการสังเคราะห์ไบโอดีเซล แบบจำลองที่โครงข่ายประสาทใช้ในการอธิบายความสัมพันธ์หรือรูปแบบระหว่างข้อมูลป้อนเข้าและข้อมูลผลลัพธ์ดังแสดงในรูปที่ 1 และสามารถแสดงความสัมพันธ์ในรูปของเมทริกซ์ได้ดังแสดงในสมการที่ (1) [6]

$$v_m = w_{m,n}x_{m,n}^T + b \tag{1}$$

โดยตัวแปร x คือ ข้อมูลป้อนเข้า v คือ ผลรวมถ่วงน้ำหนัก w คือ สัมประสิทธิ์ถ่วงน้ำหนัก b คือ ค่าเอนเอียง ตัวยก T คือ การสับเปลี่ยนแถวและหลักของเมทริกซ์ ตัวห้อย m และ n คือ ดัชนีชุดข้อมูลและดัชนีตัวแปรตามลำดับ จากนั้นผลรวมถ่วงน้ำหนักจะถูกเปลี่ยนไปเป็นข้อมูลผลลัพธ์ผ่านฟังก์ชันกระตุ้น (activated function) ดังสมการที่ (2)

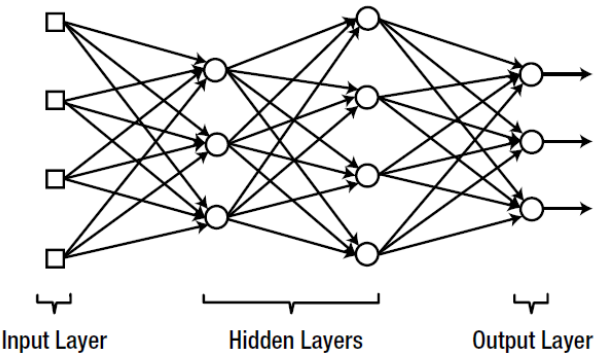
$$y_m = \varphi(v_m) \tag{2}$$

โดยที่ตัวแปร y คือ ข้อมูลผลลัพธ์ และ φ คือ ฟังก์ชันกระตุ้น



รูปที่ 1 ตัวอย่างระบบประสาทที่ประกอบไปด้วย 3 ตัวแปรป้อนเข้า (x) และ 1 ตัวแปรผลลัพธ์ (y) โดยตัวแปรป้อนเข้าแต่ละตัวจะมีการให้ค่าถ่วงน้ำหนัก (w) แตกต่างกันไป รวมไปจนถึงค่าเอนเอียง (b)

รูปที่ 1 นั้นเป็นการแสดงแผนภาพของแบบจำลองของเซลล์ประสาทในเชิงคณิตศาสตร์ โดยวงกลมหรือโหนด (node) เปรียบเสมือนเซลล์ประสาท (neuron cell) และลูกศรเปรียบเสมือนเส้นประสาทที่คอยส่งข้อมูลไปยังเซลล์ประสาท นอกจากนี้รูปที่ 1 ยังเป็นการแสดงเซลล์ประสาทเพียงเซลล์เดียว ซึ่งในระบบประสาทจริง ๆ นั้น แต่ละเซลล์ประสาทจะเชื่อมต่อกันเป็นระบบโครงข่ายประสาท ดังแสดงในรูปที่ 2 (หมายเหตุ สี่เหลี่ยม หมายถึงข้อมูลป้อนเข้า)



รูปที่ 2 โครงข่ายประสาท

ในรูปที่ 2 ระบบโครงข่ายประสาทจะประกอบด้วย 3 ชั้น (layer) หลัก ๆ ชั้นที่ 1 คือ ชั้นป้อนข้อมูล (Input layer) ชั้นที่ 2 คือ ชั้นซ่อนตัว (hidden layer) และชั้นที่ 3 ชั้นผลลัพธ์ (output layer) การดำเนินการของโครงข่ายประสาทนั้น ในชั้นซ่อนตัวจะเป็นการดำเนินการปรับค่าสัมประสิทธิ์ถ่วงน้ำหนักของสมการเชิงเส้นเพื่อให้ค่าที่ได้จากชั้นป้อนเข้าและชั้นผลลัพธ์มีค่าใกล้เคียงกันมากที่สุดตามสมการที่ (1) และ (2)

2.2 การจำลองการอบเหนียว

เทคนิคการจำลองการอบเหนียวเป็นเทคนิคสำหรับการหาค่าที่เหมาะสม โดยเป็นการเลียนแบบกระบวนการทางอุณหพลศาสตร์ในกระบวนการอบเหนียวโลหะซึ่งเป็นขั้นตอนการลดอุณหภูมิระหว่างการหลอมโลหะเพื่อให้ได้โลหะที่อยู่ในสภาวะที่เหมาะที่สุด [7]

จากกฎของอุณหพลศาสตร์ ความน่าจะเป็น $P(\Delta E)$ ที่จะเกิดการเพิ่มขึ้นของพลังงาน ΔE ณ อุณหภูมิ t สามารถประมาณได้จากสมการ

$$P(\Delta E) = e^{-\Delta E/kt} \quad (3)$$

โดยที่ k คือ ค่าคงที่ของโบลทซมาน (Boltzmann's constant) วิธีการอบเหนียวจะสุ่มการก่อกวน (random perturbation) ในระบบและคำนวณการเปลี่ยนแปลงของพลังงาน ΔE เมื่อมีการเปลี่ยนสถานะ (state) จากสถานะหนึ่งไปยังสถานะต่อไป ถ้ามีการลดลงของพลังงาน $(\Delta E < 0)$ สถานะใหม่จะถูกยอมรับ แต่หากไม่มีการลดลงของพลังงาน $(\Delta E > 0)$ สถานะใหม่จะถูกยอมรับด้วยความน่าจะเป็น $P(\Delta E)$ กระบวนการนี้จะถูกทำซ้ำตามลำดับการลดลงของอุณหภูมิ ซึ่งจะลดพลังงานลงจนกระทั่งเข้าสู่สภาวะสมดุลทางความร้อน (thermal equilibrium)

การประยุกต์ใช้การจำลองการอบเหนียวในการหาค่าที่เหมาะสมจะกำหนดฟังก์ชันเป้าหมาย (cost function) เป็นพลังงาน E พารามิเตอร์ควบคุมคืออุณหภูมิ t และผลเฉลยที่ดีที่สุดคือสถานะ วิธีการจำลองการอบเหนียวจะเริ่มจากการสุ่มค่าผลเฉลยหรือสถานะและตรวจสอบค่าของฟังก์ชันเป้าหมายจากค่าสถานะและพารามิเตอร์ควบคุม จากนั้นก็จะทำซ้ำต่อโดยใช้การสุ่มเลือกค่าใกล้เคียงผลเฉลยปัจจุบัน และนำค่าฟังก์ชันเป้าหมายจากสถานะเดิมและสถานะใหม่มาเปรียบเทียบกัน

สถานะปัจจุบันนั้นถือเป็นสถานะที่ดีที่สุด ณ ขณะนั้น และถ้าสถานะใหม่มีการเปลี่ยนแปลงทำให้พลังงานลดลง ก็จะยอมรับค่าสถานะใหม่เป็นค่าสถานะที่ดีที่สุดแทน แต่ถ้าไม่ก็จะยอมรับสถานะใหม่เป็นสถานะที่ดีที่สุดแทนด้วยความน่าจะเป็นในสมการ (3)

3. วัสดุและวิธีการ

3.1 ชุดข้อมูลการสังเคราะห์ไบโอดีเซล

ชุดข้อมูลการสังเคราะห์ไบโอดีเซลถูกรวบรวมจากรายงานการวิจัยในฐานข้อมูลทางวิชาการ ได้แก่ ScienceDirect, Springer, ACS และ TCI จำนวน 88 เงื่อนไขการสังเคราะห์ โดยข้อมูลที่รวบรวมนั้นจะเน้นที่ตัวแปรควบคุม 4 ตัวแปร (ตัวแปรป้อนเข้า) ได้แก่ อุณหภูมิ อัตราส่วนแอลกอฮอล์ต่อน้ำมัน น้ำหนักตัวเร่งปฏิกิริยา และเวลาในการทำปฏิกิริยา (retention time) และร้อยละผลผลิตเป็นตัวแปรผลลัพธ์ ชุดข้อมูลดังกล่าวจะถูกแบ่งออกเป็น 2 ชุด โดยชุดแรกจำนวน 80 รายการจะใช้ในการฝึกการ

เรียนรู้ของเครื่องจักร และชุดที่ 2 อีก 8 รายการที่เหลือจะใช้สำหรับการตรวจสอบแบบจำลองที่ได้จากการฝึกการเรียนรู้

นอกจากนี้ข้อมูลที่ใช้สำหรับการจำลองนั้นจะเฉพาะเจาะจงไปที่ประเภทน้ำมันตั้งต้น 3 ประเภท คือ น้ำมันปาล์ม น้ำมันถั่วเหลือง และน้ำมันที่เหลือทิ้งจากการปรุงอาหาร ส่วนตัวเร่งปฏิกิริยาจะเน้นเฉพาะโปแตสเซียมไฮดรอกไซด์ (KOH) สัดส่วนของแอลกอฮอล์ต่อน้ำมันนั้นจะเน้นเฉพาะแอลกอฮอล์ประเภทเมทานอล (CH_3OH) เท่านั้น

3.2 การทดสอบแบบจำลองของเครื่องจักร

เมื่อทำการฝึกการเรียนรู้ของเครื่องจักรผ่านวิธีการโครงข่ายประสาทแล้ว ผลลัพธ์ที่ได้ คือ แบบจำลองที่เป็นตัวแทนความสัมพันธ์ระหว่างตัวแปรป้อนเข้า (เงื่อนไขสภาวะการสังเคราะห์) และตัวแปรผลลัพธ์ (ร้อยละของผลผลิต) อย่างไรก็ตาม เพื่อให้แน่ใจได้ว่าแบบจำลองที่สร้างขึ้นมานั้นมีความแม่นยำเพียงใด จึงต้องใช้วิธีการพยากรณ์ผลลัพธ์โดยการป้อนข้อมูลสภาวะการสังเคราะห์ของข้อมูลชุดที่ 2 แล้วให้แบบจำลองทำนายผลลัพธ์ออกมาเพื่อเปรียบเทียบกับข้อมูลผลลัพธ์ของข้อมูลชุดที่ 2 ในการตัดสินความแม่นยำของแบบจำลองจะใช้ร้อยละความคลาดเคลื่อนดังแสดงในสมการที่ (4) เป็นเกณฑ์ตัดสิน

$$\%error = \left|\frac{y_{exact} - y_{predict}}{y_{exact}}\right| \times 100 \qquad (4)$$

เมื่อ $\%error$ คือ ร้อยละความคลาดเคลื่อน y_{exact} คือ ข้อมูลผลลัพธ์จากชุดข้อมูลชุดที่ 2 และ $y_{predict}$ คือ ผลลัพธ์ที่เกิดจากการทำนายโดยแบบจำลอง

4. ผลการทดลองและอภิปรายผล

4.1 แบบจำลองการเรียนรู้

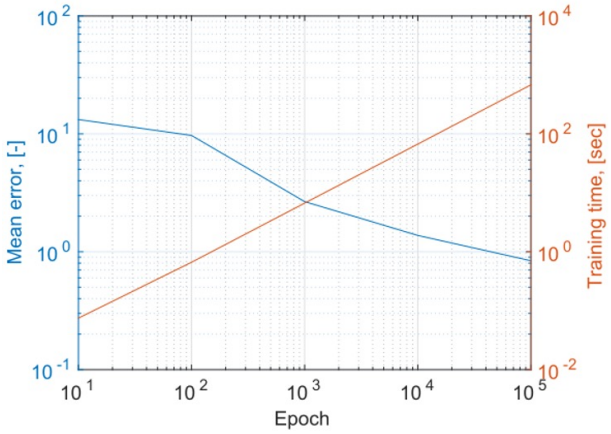
ชุดข้อมูล $(X_{80\times4})$ ของสภาวะการสังเคราะห์ไบโอดีเซลที่มีตัวแปรควบคุม (n) จำนวน 4 ตัวแปร ได้แก่ อุณหภูมิ อัตราส่วนแอลกอฮอล์ต่อน้ำมัน ร้อยละโดยน้ำหนักตัวเร่งปฏิกิริยา และเวลาในการทำปฏิกิริยา จำนวน 80 รายการ (m) ถูกป้อนเข้าในแบบจำลองโครงข่ายประสาท โดยก่อนที่จะป้อนข้อมูลเข้าไปนั้นชุดข้อมูลถูกปรับขนาดให้อยู่ในช่วง [0, 1] โดยให้ค่าต่ำสุดของแต่ละตัวแปรมีค่าเป็น 0 และค่าสูงสุดของแต่ละตัวแปรมีค่าเป็น 1 ด้วยสมการที่ (5)

$$x_{m,n}^{\text{scale}} = \frac{x_{m,n} - x_{m,n}^{\min}}{x_{m,n}^{\max} - x_{m,n}^{\min}} \qquad (5)$$

โดยที่ $x_{m,n}^{\text{scale}}$ คือ ตัวแปรที่ถูกปรับค่าให้อยู่ในช่วง [0, 1] $x_{m,n}^{\max}$ และ $x_{m,n}^{\min}$ คือ ค่าสูงสุดและค่าต่ำสุดของตัวแปร x_n จากทั้งหมด m รายการตามลำดับ ฟังก์ชันซิกมอยด์ (sigmoid) ถูกเลือกเป็นฟังก์ชันกระตุ้น และกำหนดชั้นซ่อนตัวจำนวน 2 ชั้น

จากการฝึกการเรียนรู้เครื่องจักรตามรูปแบบและข้อมูลที่กำหนด พบว่าหากจำนวนรอบของการเรียนรู้ (epoch) เพิ่มขึ้น ร้อยละความคลาดเคลื่อนในการทำนายผลลัพธ์จะลดลง รูปที่ 3 แสดงให้เห็นว่าค่าเฉลี่ยร้อยละความคลาดเคลื่อนของการทำนายผลลัพธ์ของแบบจำลองจะลดลงในขณะที่เพิ่มจำนวนรอบการเรียนรู้ ถ้าแบบจำลองได้รับการฝึกฝนเป็นเวลา 100,000 รอบ ร้อยละค่าความคลาดเคลื่อนเฉลี่ยคือ 0.83 และใช้เวลาฝึกประมาณ 10 นาที (ขึ้นอยู่กับคุณสมบัติของเครื่อง ในที่นี้ใช้ Intel (R) Core (TM) i3-2310M CPU @ 2.10GHz หน่วยความจำ 8 GB) เมื่อจำนวนรอบการเรียนรู้ลดลงทุก ๆ 10 เท่า เวลาในการคำนวณจะลดลงในสัดส่วนเดียวกัน อย่างไรก็ตามความคลาดเคลื่อนจะเพิ่มขึ้น 2 – 3 เท่าในขณะที่ลดจำนวนรอบของการเรียนรู้ลง

จะเห็นได้ว่าความคลาดเคลื่อนขึ้นอยู่กับจำนวนรอบของการเรียนรู้ซึ่งมีผลต่อเวลาในการคำนวณด้วย ดังนั้นการตัดสินใจเลือกรูปแบบการฝึกการเรียนรู้อาจขึ้นอยู่กับเวลาของการฝึกการเรียนรู้ อย่างไรก็ตาม ในปัจจุบันเวลาในการคำนวณอาจถูกยกเว้นได้หากการฝึกการเรียนรู้ทำบนคอมพิวเตอร์ที่มีประสิทธิภาพสูง การฝึกการเรียนรู้จะสามารถทำได้รวดเร็วกว่านี้ ดังนั้นจึงควรให้ความสนใจไปที่ร้อยละค่าความคลาดเคลื่อนเป็นหลัก ในบทความนี้จึงเลือกแบบจำลองที่เกิดจากการเรียนรู้จำนวน 100,000 รอบ ซึ่งให้ร้อยละความคลาดเคลื่อนเฉลี่ยที่น้อยที่สุดในการทดลอง ในหัวข้อถัดไปจะเป็นการนำเอาแบบจำลองดังกล่าวไปใช้สำหรับการทำนายข้อมูลชุดใหม่ 8 รายการ



รูปที่ 3 แสดงความสัมพันธ์ระหว่างจำนวนรอบของการฝึกการเรียนรู้ที่ส่งผลต่อร้อยละความคลาดเคลื่อนและระยะเวลาในการฝึกการเรียนรู้ด้วยเทคนิคโครงข่ายประสาทจากข้อมูลและรูปแบบที่กำหนด

4.2 การทดสอบแบบจำลอง

แบบจำลองที่ได้สร้างขึ้นจากข้อมูลที่กำหนดและฝึกการเรียนรู้จำนวน 100,000 รอบ ถูกนำมาตรวจสอบความแม่นยำกับข้อมูลชุดที่ 2 จำนวน 8 รายการ ตารางที่ 1 แสดงชุดข้อมูลใหม่ที่ป้อนสภาวะการสังเคราะห์ให้กับแบบจำลองเพื่อทำนายร้อยละผลผลิต จะพบได้ว่าผลการทำนายด้วยแบบจำลองมีร้อยละความคลาดเคลื่อนจากข้อมูลเดิมน้อย (สูงสุดคือ 2.44 %) ดังนั้นจึงสามารถตรวจสอบได้ว่าแบบจำลองดังกล่าวสามารถนำไปใช้ทำนายร้อยละผลผลิตของการสังเคราะห์ไบโอดีเซลต่อไปได้ ถัดไปจะเป็นการนำเอาแบบจำลองดังกล่าวมาใช้สำหรับการค้นหาสภาวะการสังเคราะห์ที่สามารถให้ผลผลิตไบโอดีเซลสูงสุด

4.3 การทำนายการสังเคราะห์ไบโอดีเซล

การค้นหาสภาวะการสังเคราะห์เพื่อให้ได้ร้อยละผลผลิตสูงสุดจะใช้เทคนิคการจำลองการอบเหนียว โดยการปรับเปลี่ยนค่าของตัวแปรควบคุมทั้ง 4 ตัวตามกลยุทธ์ของการจำลองการอบเหนียว และกำหนดให้ y จากแบบจำลองที่ได้จากการฝึกเครื่องจักรเป็นฟังก์ชันเป้าหมาย ผลการค้นหาสภาวะที่ให้ร้อยละผลผลิตสูงสุดจะอยู่ที่ร้อยละ 97 โดยมีเงื่อนไขการสังเคราะห์ดังนี้ 55.05 °C, 4.71 (MeOH: oil), 4.87 (Cat.%wt) และ 54.45 นาที ตามลำดับ

แม้ว่าจุดมุ่งหมายของการวิจัยนี้ต้องการสภาวะการสังเคราะห์ที่สามารถให้ผลผลิตที่ 100% แต่ผลการทำนายแสดงเพียง 97.00% นี่เป็นเพราะชุดข้อมูลที่ใช้สำหรับฝึกการเรียนรู้บันทึกร้อยละผลผลิตสูงสุดที่ 97 ดังนั้นผลการทำนายคือค่าสูงสุดของชุดข้อมูลและจะมีค่าไม่มากเกินกว่านี้

อย่างไรก็ตามสภาวะการสังเคราะห์ที่ได้จากการทำนายแตกต่างจากรายการข้อมูลที่บันทึกค่าสูงสุดในชุดข้อมูล ซึ่งก็คือ 60°C, 1.47 (MeOH: oil), 2 (Cat.%wt) และ 15 นาที เมื่อเปรียบเทียบกับผลการทำนายเงื่อนไขที่บันทึกไว้แสดงให้เห็นว่าดีกว่าเงื่อนไขการทำนายเนื่องจากใช้เวลาและวัสดุน้อยกว่า

แม้ว่าผลการทำนายจะแสดงให้เห็นว่าสภาวะการสังเคราะห์นั้นใช้ทรัพยากรในการสังเคราะห์ที่สูงกว่าข้อมูลที่

ตารางที่ 1 แสดงผลการทดสอบแบบจำลองกับชุดข้อมูลใหม่ 8 รายการ

ชนิดของน้ำมัน	อุณหภูมิ (°C)	อัตรส่วน MeOH:Oil	ความเข้มข้นตัวเร่งปฏิกิริยา (Cat. %wt)	ระยะเวลาการเกิดปฏิกิริยา (min)	ร้อยละผลผลิต (%yield)	ผลการทำนายร้อยละผลผลิต (%yield)	ร้อยละความผิดพลาด
ถั่วเหลือง	60	9.00	1.0	60	93.20	91.73	1.58
ทอดอาหาร	60	1.47	1.5	40	94.00	95.82	1.94
ปาล์มดิบ	75	8.00	10.0	120	60.77	60.55	0.36
ถั่วเหลือง	60	3.00	1.5	60	89.30	87.12	2.44
ปาล์มดิบ	75	8.00	5.0	90	61.23	60.91	0.52
ทอดอาหาร	60	1.47	1.0	30	79.00	79.06	0.08
ถั่วเหลือง	60	3.00	2.0	60	78.00	78.90	1.15
ปาล์มดิบ	75	6.00	10.0	90	63.40	63.12	0.44

บันทึก แต่หากสามารถยืนยันผลการทำนายในเชิงการทดลองในห้องปฏิบัติการได้ว่าสภาวะที่ทำนายนั้นให้ร้อยละผลผลิตเท่ากับที่ได้ทำนายไว้ แสดงให้เห็นว่าแบบจำลองที่สร้างขึ้นนั้นสามารถใช้ในการทำนายได้

นอกจากนี้หากสมมุติว่าแบบจำลองที่สร้างขึ้นสามารถใช้เป็นตัวแทนในการสังเคราะห์ได้ จะต้องมีการปรับวิธีการค้นหาสภาวะที่เหมาะสมใหม่ เนื่องจากการค้นหานั้นจะทำการค้นเฉพาะภายในข้อมูลที่ทำการฝึกการเรียนรู้ ดังนั้นการค้นหาสภาวะที่ทำให้การสังเคราะห์ได้ผลผลิตร้อยละ 100 โดยที่ไม่เคยรู้ข้อมูลการสังเคราะห์มาก่อนเลย จึงเป็นโจทย์ที่ท้าทายสำหรับการพัฒนาศาสตร์ทางการเรียนรู้ของเครื่องจักรซึ่งจะส่งผลกระทบต่ออุตสาหกรรมเคมีเป็นอย่างสูง

5. สรุป

สภาวะการสังเคราะห์ไบโอดีเซลที่สูงที่สุดผ่านกระบวนการเอสเทอริฟิเคชันได้ถูกทำนายโดยใช้เทคนิคการขับเคลื่อนด้วยข้อมูล ชุดข้อมูลสภาพการสังเคราะห์รวบรวมจากวรรณกรรมและใช้เป็นข้อมูลป้อนเข้าและข้อมูลผลลัพธ์สำหรับการฝึกการเรียนรู้ในเทคนิคโครงข่ายประสาท แบบจำลองที่สร้างขึ้นจากการฝึกอบรมด้วยโครงข่ายประสาท แสดงให้เห็นว่ามีความแม่นยำในการนำไปใช้เมื่อทำการตรวจสอบด้วยชุดข้อมูลใหม่ จากนั้นแบบจำลองได้ถูกนำมาใช้สำหรับการค้นหาสภาวะการสังเคราะห์ที่เหมาะสมด้วยเทคนิคการจำลองการอบเหนียว พบว่าการทำนายการให้ผลผลิตสูงสุดจะมีค่าเช่นเดียวกับข้อมูลที่บันทึก แต่แตกต่างกันในเงื่อนไขการสังเคราะห์ ซึ่งแสดงให้เห็นว่าการค้นหาสภาวะที่เหมาะสมนั้นจะไม่สามารถทำนายได้เกินกว่าข้อมูลที่บันทึกไว้ อย่างไรก็ตามหากสภาวะที่ทำนายได้สามารถสังเคราะห์ในเชิงห้องปฏิบัติการได้ตามที่ทำนายไว้ ก็แสดงให้เห็นว่าแบบจำลองที่สร้างขึ้นมานั้นสามารถนำมาประยุกต์ใช้ในกระบวนการสังเคราะห์ไบโอดีเซลได้

6. เอกสารอ้างอิง

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-44, May 28 2015.

[2] S. Baroutian, M. K. Aroua, A. A. A. Raman, and N. M. N. Sulaiman, "Prediction of Palm Oil-Based Methyl Ester Biodiesel Density Using Artificial Neural Networks," Journal of Applied Sciences, vol. 8, no. 10, pp. 1938-1943, 2008.

[3] M. P. Dorado, E. Ballesteros, M. Mittelbach, and F. J. López, "Kinetic Parameters Affecting the Alkali-Catalyzed Transesterification Process of Used Olive Oil," Energy & Fuels, vol. 18, no. 5, pp. 1457-1462, 2004.

[4] A. J. Yuste and M. P. Dorado, "A Neural Network Approach to Simulate Biodiesel Production from Waste Olive Oil," Energy & Fuels, vol. 20, no. 1, pp. 399-402, 2006.

[5] M. Corral Bobadilla, R. Fernández Martínez, R. Lostado Lorza, F. Somovilla Gómez, and E. Vergara González, "Optimizing Biodiesel Production from Waste Cooking Oil Using Genetic Algorithm-Based Support Vector Machines," Energies, vol. 11, no. 11, 2018.

[6] P. Kim, MATLAB Deep Learning. APress, 2017.

[7] P. J. M. van Laarhoven and E. H. L. Aarts, Simulated Annealing: Theory and Applications. Springer, 1987.