

IMPROVING MACHINE LEARNING MODEL PERFORMANCE FOR NILE TILAPIA FCR PREDICTION VIA FUNCTIONAL ADDITIVE CATEGORIZATION: A STEP-WISE FEATURE ADDITION ANALYSIS

Jeerawat BANCHACHAI^{1*}, Hien Van DOAN² and Kriangsak PROMPAK¹

1 Department of Physics, Faculty of Science, Naresuan University, Thailand; jeerawatb66@nu.ac.th
(Corresponding Author)

2 Department of Animal and Aquatic Sciences, Faculty of Agriculture, Chiang Mai University, Thailand

ARTICLE HISTORY

Received: 10 April 2026

Revised: 24 April 2026

Published: 16 May 2026

ABSTRACT

The purpose of this research is to compare the performance of machine learning models and identify the optimal input dataset using step-wise feature addition analysis to predict feed conversion ratio (FCR) in Nile tilapia. We collected a total of 171 diet formulations from studying the effect of dietary supplementation with functional additives on the growth performance of Nile tilapia. The functional additives were classified in 3 groups, i.e., probiotics, prebiotics, and plant extracts. In this research, we studied 8 machine learning models, including Multiple Linear Regression, Ridge, Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost. This study found that the CatBoost achieved the best accurate FCR predicted performance (RMSE = 0.0889 and $R^2 = 0.9011$). Additionally, the step-wise feature addition analysis found that identifying types and dosages of functional additives significantly improved FCR prediction performance. This research shows the enhancing performance of machine learning to support precision aquaculture.

Keywords: Machine Learning, Step-wise Feature Addition Analysis, Feed Conversion Ratio, Functional Additives

CITATION INFORMATION: Banchachai, J., Doan, H. V., & Prompak, K. (2026). Improving Machine Learning Model Performance for Nile Tilapia FCR Prediction via Functional Additive Categorization: A Step-Wise Feature Addition Analysis. *Procedia of Multidisciplinary Research*, 4(5), 32

การปรับปรุงประสิทธิภาพโมเดลการเรียนรู้ของเครื่องสำหรับการทำนาย อัตราการเปลี่ยนอาหารเป็นเนื้อของปลาไหลด้วยการจำแนกประเภทสาร เสริมเชิงหน้าที่: การวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น

จิรวัดน์ ปัญชาชัย^{1*}, Hien Van Doan² และ เกรียงศักดิ์ พรหมภักดี¹

1 ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร; jeerawatb66@nu.ac.th (ผู้ประพันธ์บรรณกิจ)

2 ภาควิชาสัตวศาสตร์และสัตว์น้ำ คณะเกษตรศาสตร์ มหาวิทยาลัยเชียงใหม่

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องและระบุชุดข้อมูลนำเข้าที่เหมาะสมที่สุด โดยศึกษาการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น (Step-wise Feature Addition Analysis) ในการทำนายค่าอัตราการเปลี่ยนอาหารเป็นเนื้อ (FCR) ของปลาไหล ผู้วิจัยรวบรวมข้อมูลสารเสริมเชิงหน้าที่ 171 สูตรอาหาร แบ่งเป็น 3 กลุ่ม คือ โปรไบโอติกส์ 프리ไบโอติกส์ และสารสกัดจากพืช และศึกษา 8 แบบจำลอง ได้แก่ Multiple Linear Regression, Ridge Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM และ CatBoost ผลการศึกษาพบว่า CatBoost ให้ประสิทธิภาพแม่นยำที่สุด (RMSE = 0.0889, R^2 = 0.9011) นอกจากนี้ การวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น พบว่า การระบุชนิดและปริมาณสารเสริมช่วยเพิ่มความแม่นยำในการทำนายอย่างชัดเจน งานวิจัยนี้แสดงถึงแนวทางการเพิ่มประสิทธิภาพแบบจำลองเพื่อส่งเสริมการเพาะเลี้ยงสัตว์น้ำอย่างแม่นยำ

คำสำคัญ: การเรียนรู้ของเครื่อง การวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น อัตราการเปลี่ยนอาหารเป็นเนื้อ สารเสริมเชิงหน้าที่

ข้อมูลการอ้างอิง: จิรวัดน์ ปัญชาชัย, Doan, H. V. และ เกรียงศักดิ์ พรหมภักดี. (2569). การปรับปรุงประสิทธิภาพโมเดลการเรียนรู้ของเครื่องสำหรับการทำนายอัตราการเปลี่ยนอาหารเป็นเนื้อของปลาไหลด้วยการจำแนกประเภทสารเสริมเชิงหน้าที่: การวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น. *Procedia of Multidisciplinary Research*, 4(5), 32

บทนำ

ปลานิล (Nile Tilapia) เป็นสัตว์น้ำที่มีการเพาะเลี้ยงกันอย่างแพร่หลายซึ่งมีบทบาทความสำคัญในด้านความมั่นคงทางอาหารและเพิ่มมูลค่าทางเศรษฐกิจ เนื่องจากมีความต้องการทางตลาดสูง มีรสชาติและเนื้อสัมผัสที่ดี รวมถึงมีการเจริญเติบโตที่รวดเร็ว ทว่าความท้าทายหลักในการเพาะเลี้ยง คือ “ต้นทุนค่าอาหาร” ซึ่งเป็นปัจจัยสำคัญที่ส่งผลโดยตรงต่อต้นทุนการผลิต ดังนั้น ปัจจุบันจึงมีหลายการศึกษาที่พัฒนาสูตรอาหารร่วมกับการประยุกต์ใช้สารเสริมเชิงหน้าที่ (Functional Additives) เช่น โพรไบโอติกส์ 프리ไบโอติกส์ และสารสกัดจากพืช เพื่อช่วยเพิ่มประสิทธิภาพการเจริญเติบโต กระตุ้นภูมิคุ้มกัน และเพิ่มความต้านทานโรคในปลานิล เพื่อลดการใช้จ่ายปุ๋ยชีวณะ ซึ่งดัชนีวัดความคุ้มค่าและประสิทธิภาพของสูตรอาหารคือค่าอัตราการเปลี่ยนอาหารเป็นเนื้อ (Feed Conversion Ratio: FCR) ซึ่งคำนวณจากสัดส่วนปริมาณอาหารที่กินต่อน้ำหนักตัวที่เพิ่มขึ้น โดยค่า FCR ที่ต่ำแสดงประสิทธิภาพการเปลี่ยนอาหารเป็นเนื้อที่ดี อย่างไรก็ตาม การทดลองเพื่อหาสูตรอาหารที่เหมาะสมด้วยวิธีดั้งเดิม มีข้อจำกัดสำคัญด้านต้นทุนและระยะเวลาในการทดลองสูง

ดังนั้น ในงานวิจัยนี้ เพื่อแก้ปัญหาดังกล่าวผู้วิจัยจึงมีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องจำนวน 8 แบบจำลอง ในการทำนายค่า FCR ในปลานิลโดยใช้ข้อมูลที่รวบรวมมาจากการศึกษาผลของสารเสริมเชิงหน้าที่ และศึกษาการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น (Step-wise Feature Addition Analysis) รวมถึงระบุชุดข้อมูลและตัวแปรนำเข้าที่เหมาะสมที่สุดที่ส่งผลต่อความแม่นยำของแบบจำลอง เพื่อเป็นแนวทางในการคัดเลือกสูตรอาหารและลดค่า FCR ได้อย่างมีประสิทธิภาพ ซึ่งผู้วิจัยหวังว่าผลลัพธ์จากงานวิจัยนี้จะเป็นแนวทางสำคัญในการพัฒนาการจัดการอาหารและเพิ่มประสิทธิภาพการผลิตในระบบการเพาะเลี้ยงสัตว์น้ำอย่างแม่นยำ

การทบทวนวรรณกรรม

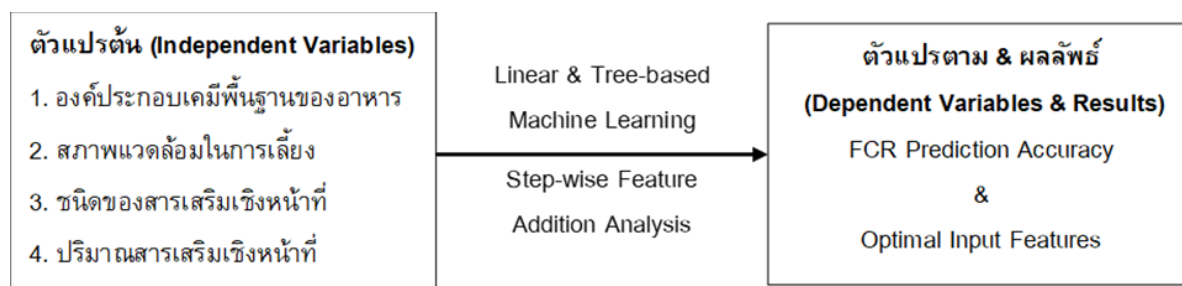
ในปัจจุบัน เพื่อเพิ่มประสิทธิภาพในอุตสาหกรรมปศุสัตว์และสัตว์น้ำ ได้มีการนำแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) มาประยุกต์ใช้เพื่อทำนายค่า FCR โดย Yang et al. (2025) ได้ศึกษาการใช้ข้อมูล FCR ระยะสั้นเพื่อทำนายค่า FCR ระยะยาวในสุกรขุน Davison et al. (2023) ได้ทำการสร้างแบบจำลองการเรียนรู้ของเครื่องโดยใช้ข้อมูลปริมาณอาหารที่ให้ ในการทำนายค่า FCR Johansen et al. (2019) แสดงให้เห็นถึงประสิทธิภาพในการทำนายค่า FCR ของไก่โดยการประยุกต์ใช้โครงข่ายประสาทเทียมแบบพลวัต (Dynamic Neural Network; DNN) Li et al. (2024) ได้สร้างแบบจำลองการทำนาย FCR ในไก่โดยใช้ตัวแปรความหลากหลายของสภาพแวดล้อม Huang et al. (2025) ได้ใช้ประโยชน์จาก Meta-analysis ร่วมกับ Gradient Boosting Machine ในการวิเคราะห์ความถี่ในการให้อาหารที่เหมาะสมต่ออัตราการเปลี่ยนอาหารเป็นเนื้อและน้ำหนักที่เพิ่มขึ้นต่อวันในปลาหลายสายพันธุ์ และ Serradell et al. (2023) ได้นำเสนอการใช้ Multiple Linear Regression เพื่อแสดงให้เห็นถึงความหลากหลายของสารเสริมเชิงหน้าที่ส่งผลต่อการทำนายประสิทธิภาพการเจริญเติบโตของปลากะพงขาว

แต่อย่างไรก็ตาม จากงานวิจัยที่ผ่านมา ยังไม่มีการศึกษาที่ประยุกต์ใช้แบบจำลองการเรียนรู้ของเครื่องร่วมกับการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น (Step-wise Feature Addition Analysis) เพื่อศึกษาอิทธิพลของตัวแปรนำเข้าและระบุชุดข้อมูลที่เหมาะสมที่ส่งผลต่อความแม่นยำของแบบจำลองในการทำนายค่า FCR ในปลานิล ดังนั้นงานวิจัยนี้จึงทำการศึกษาส่วนนี้เพื่อเป็นแนวทางในการเพาะเลี้ยงสัตว์น้ำอย่างแม่นยำ

สมมติฐานการวิจัย

- 1) แบบจำลองในกลุ่ม Tree-based มีประสิทธิภาพความแม่นยำในการทำนายค่า FCR สูงกว่ากลุ่ม Linear-based เนื่องจากความซับซ้อนและความสัมพันธ์ที่ไม่เป็นเชิงเส้นของข้อมูล
- 2) การพิจารณาชนิดและปริมาณของสารเสริมเชิงหน้าที่ร่วมกับข้อมูลองค์ประกอบทางเคมีพื้นฐานของอาหารและสภาพแวดล้อม จะช่วยเพิ่มประสิทธิภาพความแม่นยำในการทำนายค่า FCR ในปลานิล

กรอบแนวคิดการวิจัย



ภาพที่ 1 กรอบแนวคิด

วิธีดำเนินการวิจัย

ขั้นตอนและวิธีการรวบรวมข้อมูล (Data Collection)

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลสูตรอาหารจากงานวิจัยเกี่ยวกับการศึกษาผลของสารเสริมเชิงหน้าที่ที่ส่งผลต่อประสิทธิภาพการเจริญเติบโตของปลานิล ในฐานข้อมูล Google Scholar, Scopus และ ScienceDirect ที่ได้รับการตีพิมพ์ ตั้งแต่ปี ค.ศ.2005-2025 รวมทั้งหมด 171 สูตรอาหาร จาก 34 การศึกษา โดยผู้วิจัยได้ทำการเลือกตามเงื่อนไขดังนี้

- 1) การศึกษาต้องเกี่ยวข้องกับปลานิลซึ่งประกอบไปด้วยกลุ่มควบคุม (Control Group) ที่ไม่ได้รับสารเสริมเชิงหน้าที่และกลุ่ม ทรีตเมนต์ (Treatment Group) ที่ได้รับสารเสริมเชิงหน้าที่
- 2) การศึกษาต้องสนใจเกี่ยวกับสารเสริมเชิงหน้าที่ ซึ่งประกอบไปด้วยโพรไบโอติกส์ 프리ไบโอติกส์ และสารสกัดจากพืช
- 3) การศึกษาต้องมีการรายงานค่าพารามิเตอร์การเจริญเติบโต ซึ่งประกอบไปด้วยน้ำหนักเริ่มต้นและอัตราการเปลี่ยนอาหารเป็นเนื้อ (FCR)
- 4) การศึกษาต้องมีการรายงานปริมาณส่วนผสมทางเคมีของอาหาร (Proximate Composition) เช่น โปรตีน ไขมัน และพลังงาน
- 5) การศึกษาต้องมีการรายงานระยะเวลาในการทดลองอาหาร
- 6) การศึกษาต้องมีการรายงานค่าคุณภาพน้ำในการเพาะเลี้ยง เช่น อุณหภูมิ pH และออกซิเจนละลายน้ำ

ตารางที่ 1 แสดงผลค่าต่ำสุด ค่าสูงสุด ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของตัวแปรนำเข้าและตัวแปรเป้าหมาย

ตัวแปร	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย	S.D.
จำนวนวันเลี้ยง	30	213	69.5	32.9
น้ำหนักเริ่มต้น	1.23	44.6	14.5	11.7
จำนวนปลา/บ่อ	10	63	21.7	11.3
ปริมาตรบ่อเลี้ยง	30	4500	400.6	945.6
อุณหภูมิ	24.24	32	27.2	1.7
ออกซิเจนละลายในน้ำ	4.2	9.61	6	1.2
ค่าความเป็นกรด-ด่างเฉลี่ยของน้ำ	6.4	8.37	7.5	0.5
แอมโมเนีย	0.01	4.2	0.3	0.8
ค่าโปรตีนในอาหาร	272.32	393.7	319.1	24.2
ค่าไขมันในอาหาร	16.5	99.95	64.7	21.7
ค่าเถ้าในอาหาร	40.5	152.38	81.3	21.1
ค่ากากใยในอาหาร	26.2	99.3	52.9	12.2

ตัวแปร	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย	S.D.
ค่าความชื้นของอาหาร	4.1	185.4	84.7	42.2
ค่าพลังงานของอาหาร	14.29	20.65	17.8	1.4
ปริมาณสารเสริมเชิงหน้าที่	0	160	9.5	18.6
อัตราการเปลี่ยนอาหารเป็นเนื้อ	0.83	3.91	1.40	0.43

การจัดเตรียมข้อมูล (Data Preprocessing)

ก่อนเริ่มทำการวิเคราะห์ข้อมูล ผู้วิจัยได้นำวิธีการการจัดเตรียมข้อมูลมาใช้เพื่อปรับปรุงและเพิ่มคุณภาพกับข้อมูล โดยมีขั้นตอนดังนี้

ขั้นตอนที่ 1 การแปลงหน่วยพารามิเตอร์ให้เป็นมาตรฐาน (Standardizing Units) เนื่องจากข้อมูลถูกรวบรวมมาจากหลากหลายแหล่งการศึกษาวิจัย ดังนั้นข้อมูลทั้งหมดจะถูกนำมาแปลงหน่วยให้เป็นหน่วยมาตรฐาน โดยข้อมูลตัวแปรและหน่วยมาตรฐานทั้งหมดที่นำมาใช้ในการสร้างแบบจำลองการเรียนรู้ของเครื่องประกอบด้วย ระยะเวลาในการทดลองอาหาร (วัน) น้ำหนักเริ่มต้นเฉลี่ย (กรัม) ความหนาแน่นของปลาต่อบ่อ (ตัว/บ่อ) ปริมาตรของบ่อเลี้ยง (ลิตร) อุณหภูมิเฉลี่ยของน้ำ (องศาเซลเซียส) ค่าออกซิเจนละลายในน้ำเฉลี่ย (มิลลิกรัม/ลิตร) ค่าความเป็นกรด-ด่างเฉลี่ยของน้ำ (-) ค่าแอมโมเนียเฉลี่ยของน้ำ (มิลลิกรัม/ลิตร) ค่าโปรตีนในอาหาร (กรัม/กิโลกรัม) ค่าไขมันในอาหาร (กรัม/กิโลกรัม) ค่าเถ้าในอาหาร (กรัม/กิโลกรัม) ค่ากากใยในอาหาร (กรัม/กิโลกรัม) ค่าความชื้นของอาหาร (กรัม/กิโลกรัม) ค่าพลังงานของอาหาร (เมกะจูล/กิโลกรัม) ปริมาณสารเสริมเชิงหน้าที่ (กรัม/กิโลกรัม) โปรไบโอติกส์ (ใส่ = 1, ไม่ใส่ = 0) 프리ไบโอติกส์ (ใส่ = 1, ไม่ใส่ = 0) และสารสกัดจากพืช (ใส่ = 1, ไม่ใส่ = 0)

ขั้นตอนที่ 2 การทำความสะอาดข้อมูล (Data Cleaning) ขั้นตอนนี้เป็นการเตรียมและปรับปรุงคุณภาพข้อมูลให้สมบูรณ์ก่อนนำไปวิเคราะห์และสร้างแบบจำลองการเรียนรู้ของเครื่อง ดังนั้นผู้วิจัยจึงได้ทำความสะอาดข้อมูลตามขั้นตอนดังนี้

- 1) การตรวจสอบและกำจัดค่าซ้ำ (Duplicates) ในการบันทึกข้อมูลอาจเกิดการบันทึกข้อมูลซ้ำ การตรวจสอบและลบข้อมูลที่เป็นค่าซ้ำจะช่วยลดความเอนเอียงของโมเดล
- 2) การจัดการกับข้อมูลสูญหาย (Missing Values) ในการเก็บรวบรวมข้อมูลมักจะพบกับข้อมูลสูญหายซึ่งจะถูกบันทึกเป็นค่าว่างในชุดข้อมูล (Dataset) ซึ่งการจัดการกับข้อมูลสูญหายจะช่วยเพิ่มประสิทธิภาพให้กับชุดข้อมูล ในงานนี้ผู้วิจัยได้เลือกวิธีการ Random Forest Imputation (หรือ Miss Forest) มาใช้ในการจัดการข้อมูลสูญหาย
- 3) การจัดการค่าผิดปกติ (Outliers) เพื่อไม่ให้ข้อมูลที่ผิดปกติส่งผลทำให้ผลลัพธ์คลาดเคลื่อน โดยผู้วิจัยได้ใช้วิธีการแปลงลอการิทึม (log-transformation) สำหรับพารามิเตอร์ปริมาณของสารเสริมเชิงหน้าที่ ที่มีการกระจายตัวแบบเบ้ (Skewed Distribution) สูงและแก้ปัญหาค่าศูนย์ที่มากเกินไป (Zero-inflated) เพื่อลดผลกระทบของค่าผิดปกติต่อการสร้างแบบจำลอง และสำหรับพารามิเตอร์อื่น ผู้วิจัยได้ใช้วิธีการกำหนดขอบเขต (Capping) เพื่อปรับค่าที่อยู่นอกช่วงให้อยู่ในขอบเขตที่กำหนดโดยอาศัยหลักการพิสัยระหว่าง ควอร์ไทล์ (Interquartile Range; IQR) ซึ่งใช้ช่วงควอร์ไทล์ในการระบุค่าผิดปกติ จุดข้อมูลที่ต่ำกว่าขอบเขตล่างหรือสูงกว่าขอบเขตบนจะถือว่าเป็นค่าผิดปกติ
- 4) การทำให้เป็นมาตรฐาน (Normalization) เนื่องจากค่าพารามิเตอร์แต่ละตัวมีขนาดข้อมูลที่แตกต่างกัน จึงจำเป็นต้องปรับขนาดข้อมูลนั้นให้มีขนาดเท่าๆกัน โดยผู้วิจัยได้ใช้วิธีการ Standardization ดังสมการที่ (1)

$$X_{Scaled} = \frac{x - \bar{x}}{S.D.} \quad (1)$$

กำหนดให้ X คือ ค่าข้อมูล ณ จุดนั้น $\bar{x}$ คือ ค่าเฉลี่ยของข้อมูล X_{Scaled} คือ ค่าข้อมูลที่ผ่านการปรับขนาด และ $S.D.$ คือ ค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูล ทั้งนี้ ขั้นตอนที่ 2 ถึง 4 จะถูกดำเนินการหลังจากการแบ่งข้อมูล (Data Splitting) ซึ่งจะถูกอธิบายในลำดับถัดไป โดยค่าพารามิเตอร์ทางสถิติที่ใช้ในการแปลงข้อมูลทั้งหมด จะถูกคำนวณและ

เรียนรู้จากชุดข้อมูลฝึกฝน (Training set) เท่านั้น ก่อนจะนำพารามิเตอร์ดังกล่าวไปประยุกต์ใช้เพื่อแปลงค่าในชุดข้อมูลทดสอบ (Testing set) เพื่อป้องกันปัญหาการรั่วไหลของข้อมูล (Data Leakage)

การแบ่งข้อมูล (Data Splitting)

ในขั้นตอนนี้ผู้วิจัยได้แบ่งข้อมูลแบบสุ่มออกเป็น 2 ชุด โดยแบ่งเป็นชุดข้อมูลฝึกฝน (Training set) 85 เปอร์เซ็นต์ และชุดข้อมูลทดสอบ (Test set) 15 เปอร์เซ็นต์ และนำชุดข้อมูลฝึกฝนไปสร้างแบบจำลอง ซึ่งผู้วิจัยได้เลือกใช้วิธีตรวจสอบความถูกต้องไขว้แบบ 5 ส่วน (5-Fold Cross-Validation) ภายในชุดข้อมูลฝึกฝนเพื่อลดโอกาสการเกิดภาวะเรียนรู้ข้อมูลเกิน (Overfitting) และนำชุดข้อมูลทดสอบที่แบบจำลองไม่เคยเห็นมาก่อนมาใช้ทดสอบความแม่นยำของแบบจำลอง เปรียบเสมือนการนำเอาแบบจำลองไปใช้กับสถานการณ์จริง

อัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning Algorithms)

งานวิจัยนี้ได้คัดเลือกอัลกอริทึมแบบจำลองการเรียนรู้ของเครื่องจำนวน 8 แบบจำลอง เพื่อเปรียบเทียบประสิทธิภาพในการทำนายค่า FCR โดยแบ่งออกเป็น 2 กลุ่มหลักตามสถาปัตยกรรมพื้นฐาน ได้แก่ 1) กลุ่มแบบจำลองเชิงเส้น (Linear-based Models) ประกอบด้วย Multiple Linear Regression ซึ่งเป็นแบบจำลองพื้นฐานที่ใช้ในการทดสอบความสัมพันธ์เชิงเส้น และ Ridge Regression เป็นแบบจำลองที่ใช้เพื่อลดปัญหาความสัมพันธ์ซ้อนทับกันของตัวแปร (Multicollinearity) และป้องกันปัญหาการเรียนรู้ข้อมูลเกิน (Overfitting) 2) กลุ่มแบบจำลองโครงสร้างต้นไม้ (Tree-based Models) ประกอบด้วย แบบจำลองพื้นฐานอย่าง Decision Tree ไปจนถึงเทคนิคขั้นสูงกลุ่ม Ensemble Learning ได้แก่ Random Forest (อิงหลักการ Bagging) รวมถึงอัลกอริทึมกลุ่ม Boosting (Gradient Boosting, XGBoost, LightGBM และ CatBoost) ซึ่งถูกพัฒนามาเพื่อแก้ปัญหาความคลาดเคลื่อนสะสม เพิ่มความเร็วในการประมวลผล มีความสามารถในการจัดการตัวแปรเชิงกลุ่มและข้อมูลสูญหาย รวมถึงมีกลไกป้องกันปัญหาการเรียนรู้ข้อมูลเกิน (Overfitting) อย่างมีประสิทธิภาพ

การประเมินประสิทธิภาพแบบจำลอง (Model Performance Evaluation)

ในการตรวจสอบความแม่นยำของแบบจำลอง ผู้วิจัยได้เลือกใช้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Root Mean Square Error; RMSE) และค่าสัมประสิทธิ์การตัดสินใจ (R-Squared; R^2) ดังสมการที่ (2) และ (3) ตามลำดับ จากนั้นผู้วิจัยเลือกแบบจำลองที่มีค่า RMSE ต่ำที่สุดและค่า R^2 สูงที่สุด ซึ่งเป็นแบบจำลองที่ดีที่สุดและนำไปใช้ในการทำนายต่อไป

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

$$R^2 = 1 - \left(\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \right) \quad (3)$$

กำหนดให้ y_i คือ ค่าจริง (Actual Values) $\hat{y}_i$ คือ ค่าทำนาย (Predicted Values) $\bar{y}$ คือ ค่าเฉลี่ย N คือ จำนวนข้อมูลทั้งหมด

การศึกษาการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น (Step-wise Feature Addition Analysis)

ในงานวิจัยนี้ทำการประยุกต์ใช้เทคนิคการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น เพื่อประเมินระดับความสำคัญและศึกษาผลกระทบของปัจจัยนำเข้าแต่ละกลุ่มต่อประสิทธิภาพความแม่นยำของแบบจำลองนั้น ผู้วิจัยได้จำแนกปัจจัยนำเข้าออกเป็น 4 กลุ่มหลัก คือ 1) องค์ประกอบทางเคมีพื้นฐานของอาหาร 2) สภาพแวดล้อมในการเพาะเลี้ยง 3) ชนิดสารเสริมเชิงหน้าที่ และ 4) ปริมาณของสารเสริมเชิงหน้าที่ ในขั้นตอนการทดสอบ ผู้วิจัยได้นำแบบจำลองที่ดีที่สุดมาฝึกฝนใหม่โดยเริ่มจากกลุ่มปัจจัยพื้นฐานและทำการเพิ่มกลุ่มปัจจัยเข้าสู่แบบจำลองทีละลำดับและประเมินประสิทธิภาพความแม่นยำของแต่ละส่วนใหม่ เพื่อศึกษาผลกระทบและวิเคราะห์ส่วนต่างของความแม่นยำที่เพิ่มขึ้นจากการรวมปัจจัยนำเข้าทีละกลุ่ม

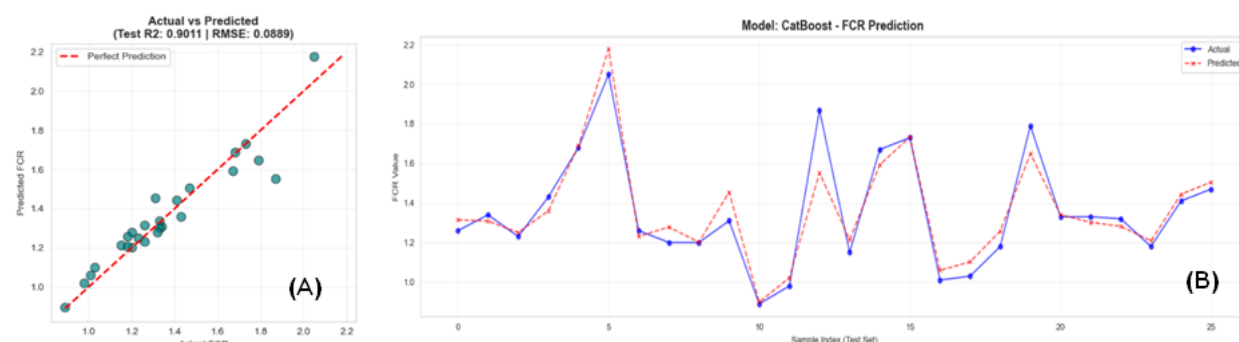
ผลการวิจัย

ผลการประเมินประสิทธิภาพความแม่นยำของแบบจำลอง

ตารางที่ 2 แสดงให้เห็นถึงผลการศึกษาการเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองทั้ง 8 แบบจำลอง โดยเรียงลำดับจากแบบจำลองที่มีค่าความแม่นยำในการทำนายจากมากไปน้อย พบว่าแบบจำลองในกลุ่ม Tree-based ให้ประสิทธิภาพสูงกว่ากลุ่ม Linear-based อย่างชัดเจน โดยแบบจำลอง CatBoost ให้ประสิทธิภาพความแม่นยำสูงที่สุด ($RMSE = 0.0889$ และ $R^2 = 0.9011$) และผลการทำนายออกมาใกล้เคียงค่าจริงที่สุด ดังแสดงในภาพที่ 2 ในขณะที่แบบจำลอง Multiple Linear Regression ให้ประสิทธิภาพความแม่นยำต่ำที่สุด ($RMSE = 0.2919$ และ $R^2 = -0.0674$) ดังนั้นผู้วิจัยจึงเลือกแบบจำลอง CatBoost เป็นแบบจำลองต้นแบบในการศึกษาการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น (Step-wise Feature Addition Analysis) เพื่อศึกษาผลกระทบของกลุ่มปัจจัยนำเข้า เป็นลำดับถัดไป

ตารางที่ 2 Machine Learning Model Performance Evaluation

Type	Model	CV R^2	Test R^2	RMSE
Tree-based	CatBoost	0.8336	0.9011	0.0889
	XGBoost	0.7954	0.8729	0.1007
	Gradient Boosting	0.7857	0.8703	0.1017
	Random Forest	0.8273	0.8631	0.1045
Linear-based	LightGBM	0.7901	0.7819	0.1320
	Decision Tree	0.7276	0.7628	0.1376
	Ridge	0.3004	-0.0090	0.2838
	Multiple Linear Regression	0.2097	-0.0674	0.2919

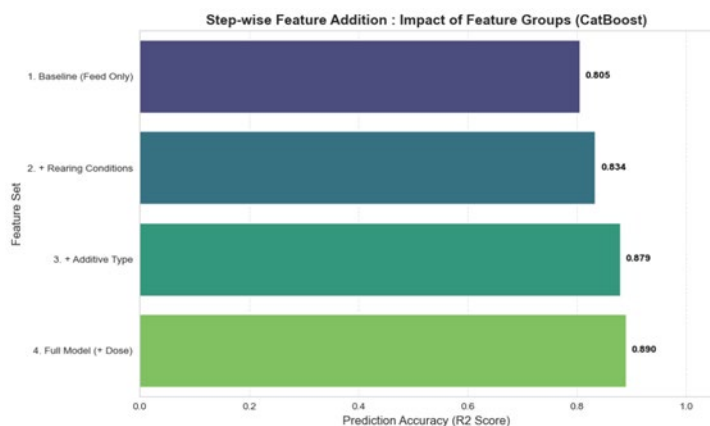


ภาพที่ 2 (A) แผนภาพการกระจาย ซึ่งแสดงความสัมพันธ์ระหว่างค่าจริงและค่าทำนายของ FCR ในพลาเนียล และ (B) กราฟเปรียบเทียบผลลัพธ์การทำนายของ FCR (โดยที่เส้นสีน้ำเงินแสดงถึงค่าจริง และเส้นประสีแดงแสดงถึงค่าทำนายของ FCR)

ผลการศึกษาการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น (Step-wise Feature Addition Analysis)

จากผลการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้นด้วยแบบจำลอง CatBoost ซึ่งให้ความแม่นยำสูงสุด เพื่อระบุชุดข้อมูลนำเข้าที่เหมาะสมที่สุด ดังแสดงในภาพที่ 3 พบว่า เมื่อมีการเพิ่มข้อมูลกลุ่มสภาพแวดล้อมในการเพาะเลี้ยงเข้ากับองค์ประกอบทางเคมีพื้นฐานทำให้ค่าสัมประสิทธิ์การตัดสินใจ (R^2) เพิ่มขึ้นเพียง 0.029 แต่เมื่อระบุกลุ่มข้อมูลชนิดสารเสริมเชิงหน้าที่เข้าไปในแบบจำลองส่งผลให้ค่า R^2 มีการเพิ่มขึ้นอย่างเด่นชัดถึง 0.045 และเมื่อมีการระบุปริมาณการใช้สารเสริมเชิงหน้าที่ทำให้ค่า R^2 เพิ่มขึ้นอีก 0.011 จากผลการศึกษาชี้ให้เห็นว่ากลุ่มตัวแปรสารเสริมเชิงหน้าที่มี

อิทธิพลต่อประสิทธิภาพการทำนายอย่างมาก เมื่อเทียบกับการใช้เพียงข้อมูลองค์ประกอบทางเคมีพื้นฐานของอาหาร และสภาพแวดล้อม



ภาพที่ 3 ผลการศึกษาอิทธิพลปัจจัยนำเข้าต่อประสิทธิภาพการทำนายของแบบจำลอง CatBoost

สรุปและอภิปรายผลการวิจัย

ในการศึกษานี้ ผู้วิจัยได้ใช้แบบจำลองการเรียนรู้ของเครื่องในการศึกษาการทำนายค่า FCR ในปลานิล ร่วมกับการศึกษาอิทธิพลของปัจจัยนำเข้าโดยการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้น ที่ส่งผลต่อประสิทธิภาพในการทำนายค่า FCR เพื่อระบุชุดข้อมูลที่เหมาะสมที่สุด จากผลการวิจัยพบว่า การเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องทั้งหมด 8 แบบจำลอง พบว่า แบบจำลองในกลุ่ม Tree-based โดยเฉพาะ CatBoost มีความแม่นยำสูงกว่ากลุ่ม Linear-based อย่าง Multiple Linear Regression ที่ให้ค่า R^2 ต่ำลง ซึ่งบ่งชี้ให้เห็นว่าความสัมพันธ์ระหว่างข้อมูลปัจจัยการเลี้ยงกับค่า FCR เป็นแบบไม่เป็นเชิงเส้น (Non-linear Relationship) จึงทำให้แบบจำลองในกลุ่ม Tree-based นั้นเหมาะสมในการนำมาวิเคราะห์ข้อมูลมากกว่ากลุ่ม Linear-based ซึ่งสอดคล้องกับงานวิจัยของ Saad et al. (2024) ที่ศึกษาการจัดการการให้อาหารที่เหมาะสมในการเพาะเลี้ยงสัตว์น้ำโดยใช้แบบจำลองการเรียนรู้ของเครื่อง และพบว่าแบบจำลองในกลุ่ม Tree-based อย่าง Gradient Boosting Regressor มีความเหมาะสมในการจัดการกับความซับซ้อนของข้อมูลในการเลี้ยงสัตว์น้ำและสามารถใช้ได้ดีกับข้อมูลที่มีรูปแบบความสำคัญที่ไม่เป็นเชิงเส้นได้ดีกว่าโมเดลพื้นฐานในกลุ่ม Linear-based นอกจากนี้ แบบจำลอง CatBoost ที่ได้รับการคัดเลือกกว่าเป็นแบบจำลองที่มีความแม่นยำในการทำนายที่ดีที่สุดในงานวิจัยนี้ ยังสอดคล้องกับงานวิจัยของ Yang et al. (2025) ที่ทำการศึกษาการทำนายอัตราการเปลี่ยนอาหารเป็นเนื้อในสุกรขุนโดยใช้แบบจำลองการเรียนรู้ของเครื่อง โดยเปรียบเทียบแบบจำลองทั้งหมด 19 แบบจำลอง และผลการศึกษพบว่าแบบจำลอง Gradient Boosting, LightGBM และ CatBoost ให้ประสิทธิภาพความแม่นยำในการทำนายค่า FCR สูงสุดในสุกรขุน เนื่องจากสามารถเรียนรู้ข้อมูลที่มีความซับซ้อนได้ดีกว่าแบบจำลองอื่นๆ นอกจากนี้ ในการศึกษาผลของปัจจัยนำเข้าโดยการวิเคราะห์แบบเพิ่มตัวแปรเป็นลำดับขั้นชี้ให้เห็นว่าชนิดของสารเสริมเชิงหน้าที่ เป็นปัจจัยที่มีอิทธิพลต่อประสิทธิภาพความแม่นยำสูงสุด ยิ่งกว่าการพิจารณาเพียงองค์ประกอบทางเคมีพื้นฐานของอาหารหรือสภาพแวดล้อมเพียงอย่างเดียว ซึ่งโดยทั่วไป องค์ประกอบเคมีพื้นฐานทางอาหารเป็นโภชนาการพื้นฐานที่เป็นแหล่งพลังงานที่ปลานิลใช้ในการสร้างกล้ามเนื้อและดำรงชีวิตแบบจำลองจึงได้แสดงให้เห็นถึงค่าความแม่นยำระดับหนึ่ง แต่อย่างไรก็ตามการที่ปลานิลได้รับอาหารที่มีโภชนาการเหมาะสม แต่สภาพแวดล้อมในการเพาะเลี้ยงไม่เหมาะสมย่อมส่งผลต่อประสิทธิภาพการเจริญเติบโตของปลานิล การที่เพิ่มตัวแปรกลุ่มสภาพแวดล้อมเข้าไปในแบบจำลองแล้วทำให้ความแม่นยำในการทำนายสูงขึ้นเพราะว่าตัวแปรกลุ่มนี้เป็นปัจจัยจำกัดทางสรีรวิทยา เช่น อุณหภูมิมีผลต่อการทำงานของเอนไซม์ในการย่อยอาหาร น้ำหนักเริ่มต้นและวันเลี้ยงมีผลต่อระยะการเจริญเติบโต โดยที่ปลานิลขนาดเล็กจะเจริญเติบโตเร็วกว่าปลานิลขนาดใหญ่เนื่องจากใช้พลังงานจากโภชนาการพื้นฐานในการสร้างกล้ามเนื้อ ส่วนปลานิลขนาดใหญ่จะใช้พลังงานส่วนใหญ่มากกว่าปัจจัยอื่นในการดำรงชีวิต

เมื่อมีการระบุชนิดสารเสริมเชิงหน้าที่เข้าไปในแบบจำลอง พบว่า ส่งผลความแม่นยำเพิ่มขึ้นอย่างเด่นชัดเนื่องจากสารเสริมเชิงหน้าที่แต่ละชนิดมีกลไกการออกฤทธิ์ทางชีวภาพ (Biological Mechanism) ที่ส่งผลต่อตัวปลาแตกต่างกัน เช่น โปรไบโอติกส์มีหน้าที่ช่วยสร้างสมดุลของจุลินทรีย์ในลำไส้ 프리ไบโอติกส์มีหน้าที่ช่วยในการเพิ่มความยาวและความหนาแน่นของวิลไล (Villi) ซึ่งส่งผลต่อการดูดซึมสารอาหาร และสารสกัดจากพืชที่มีสารออกฤทธิ์ทางชีวภาพที่ช่วยกระตุ้นภูมิคุ้มกันและต้านอนุมูลอิสระเพื่อต้านทานโรคและลดความเครียด และสุดท้ายเมื่อมีการระบุปริมาณของสารเสริมร่วมด้วยช่วยให้เพิ่มประสิทธิภาพในการทำนาย FCR ได้อย่างสมบูรณ์แบบมากขึ้น เนื่องจากสารเสริมเชิงหน้าที่ควรใส่ในปริมาณที่เหมาะสมเพื่อให้ได้ประสิทธิภาพการเจริญเติบโตสูงสุด ถ้าใส่ในปริมาณที่น้อยเกินไปอาจไม่ออกฤทธิ์ แต่หากใส่ในปริมาณที่มากเกินไปก็อาจส่งผลเป็นพิษต่อปลา ในงานวิจัยนี้จึงแสดงให้เห็นว่าการระบุชนิดพร้อมกับปริมาณของสารเสริมเชิงหน้าที่ช่วยให้แบบจำลองมีประสิทธิภาพในการทำนายแม่นยำขึ้น ซึ่งสอดคล้องกับงานวิจัยของ Serradell et al. (2023) ที่ทำการศึกษาการสร้างแบบจำลองผลกระทบของ โปรไบโอติกส์ โปรไบโอติกส์ และสารเสริมอื่นๆ ต่อการเจริญเติบโต การให้อาหาร และอัตราการเปลี่ยนอาหารเป็นเนื้อของปลากะพงขาว พบว่าการระบุกลุ่มของสารเสริมเชิงหน้าที่ เป็นตัวแปรสำคัญที่ช่วยเพิ่มความแม่นยำของแบบจำลองได้อย่างมีนัยสำคัญ ดังนั้น การศึกษานี้แสดงให้เห็นว่า การใช้แบบจำลองการเรียนรู้ของเครื่องในกลุ่ม Tree-based โดยเฉพาะ CatBoost มีประสิทธิภาพสูงที่สุดในการทำนายค่า FCR ในปลานิล เนื่องจากสามารถเรียนรู้ความสัมพันธ์ที่ซับซ้อนและไม่เป็นเชิงเส้นระหว่างตัวแปรนำเข้า ได้แก่ องค์ประกอบทางเคมีพื้นฐานของอาหาร สภาพแวดล้อมในการเพาะเลี้ยง ชนิดสารเสริมเชิงหน้าที่ และปริมาณของสารเสริมได้อย่างดี โดยให้ค่าความคลาดเคลื่อนต่ำที่สุด (RMSE ต่ำที่สุด) และค่าสัมประสิทธิ์การตัดสินใจสูงสุด (R^2 สูงที่สุด) นอกจากนี้ ผลการศึกษายังแสดงให้เห็นถึงอิทธิพลของปัจจัยนำเข้าที่ส่งผลต่อการทำนายค่า FCR โดยพบว่า การระบุชนิดร่วมกับปริมาณสารเสริมเชิงหน้าที่ช่วยเพิ่มประสิทธิภาพในการทำนายค่า FCR ให้แม่นยำขึ้น ซึ่งงานวิจัยนี้สามารถนำไปประยุกต์ใช้เพื่อเป็นแนวทางในการจัดการอาหารและเพิ่มประสิทธิภาพในการเพาะเลี้ยงสัตว์น้ำได้อย่างแม่นยำ

ข้อเสนอแนะที่ได้รับจากการวิจัย

จากผลการศึกษาพบว่า การระบุชนิดและปริมาณสารเสริมเชิงหน้าที่มีอิทธิพลสูงต่อความแม่นยำในการทำนายค่า FCR ในปลานิล ดังนั้น ผู้เพาะเลี้ยงหรือผู้ผลิตอาหารสัตว์น้ำสามารถนำชุดข้อมูลที่เหมาะสมที่สุด โดยเฉพาะการพิจารณาชนิดและปริมาณสารเสริมเชิงหน้าที่ร่วมกับสูตรอาหารพื้นฐานและสภาพแวดล้อม ไปเป็นเกณฑ์ในการประยุกต์ใช้ปรับปรุงสูตรอาหารเพื่อเพิ่มประสิทธิภาพของอาหารและลดค่า FCR ได้อย่างแม่นยำ นอกจากนี้สำหรับงานวิจัยในอนาคต ควรเพิ่มปริมาณข้อมูลให้ครอบคลุมข้อมูลสัตว์น้ำเศรษฐกิจชนิดอื่นๆ เพื่อยกระดับประสิทธิภาพของแบบจำลอง และสามารถนำไปประยุกต์ใช้ในการจัดการฟาร์มเพาะเลี้ยงสัตว์น้ำได้อย่างแพร่หลายและแม่นยำมากยิ่งขึ้น

เอกสารอ้างอิง

- Davison, C., Michie, C., Tachtatzis, C., Andonovic, I., Bowen, J., & Duthie, C.-A. (2023). Feed Conversion Ratio (FCR) and Performance Group Estimation Based on Predicted Feed Intake for the Optimisation of Beef Production. *Sensors*, 23(10), 4621.
- Huang, M., Zhou, Y.-G., Yang, X.-G., Gao, Q.-F., Chen, Y.-N., Ren, Y.-C., & Dong, S.-L. (2025). Optimizing feeding frequencies in fish: A meta-analysis and machine learning approach. *Aquaculture*, 595, 741678.
- Johansen, S. V. (2019). *Automated Optimization of Broiler Production*. Doctoral dissertation, Aalborg Universitetsforlag.

- Li, Y., Ma, R., Qi, R., Li, h., Li, J., Liu, W., Wan, Y., Li, S., Sun, Z., Xu, J., & Zhan, K. (2024). Novel insight into the feed conversion ratio in laying hens and construction of its prediction model. *Poultry Science*, 103(10), 104013.
- Saad, A., Baikas, A., Remen, M., & Bjørnson, F. O. (2024). Optimizing Feeding Strategies in Aquaculture Using Machine Learning: Ensuring Sustainable and Economically Viable Fish Farming Practices. *Procedia Computer Science*, 246, 4712-4721.
- Serradell, A., Torrecillas, S., Soares, F., Silva, T., & Montero, D. (2023). Modelling the effect of prebiotics, probiotics and other functional additives on the growth, feed intake and feed conversion of European sea bass (*Dicentrarchus labrax*) juveniles. *Aquaculture Reports*, 32, 101729.
- Yang, X., Zhu, L., Jiang, W., Yang, Y., Gan, M., Shen, L., & Zhu, L. (2025). Machine Learning-Based Prediction of Feed Conversion Ratio: A Feasibility Study of Using Short-Term FCR Data for Long-Term Feed Conversion Ratio (FCR) Prediction. *Animals*, 15(12), 1773.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



Copyright: © 2026 by the authors. This is a fully open-access article distributed under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).