

การออกแบบและพัฒนาระบบบริหารข้อมูลความมั่นคงทางอาหารสำหรับการสร้าง โมเดลพยากรณ์ตัวบ่งชี้เอฟซีเอสโดยใช้การเรียนรู้ของเครื่อง

A Design and Development of Food Security Management System for Household-level FCS Prediction using Machine Learning

ธงชัย เจือจันทร์¹, กฤษณณัฐ หนูन्छู², สุวัฒน์ กล้วยทอง^{1,*} และ ปาจารย์ ผลประเสริฐ³

Thongchai Chuachan¹, Kritsananut Nunchoo², Suwat Gluaythong^{1,*} and Pajaree Polprasert³

¹ สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสุรินทร์ ประเทศไทย; Department of Computer Science, Faculty of Science and Technology, Surindra Rajabhat University, Thailand

² สาขาวิชาเทคโนโลยีดิจิทัลเพื่อการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสุรินทร์ ประเทศไทย; Department of Information Computer Technology, Faculty of Education, Surindra Rajabhat University, Thailand

³ คณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏกำแพงเพชร ประเทศไทย; Faculty of Management Science, Kamphaeng Phet Rajabhat University, Thailand

* Corresponding author email: suwat_gl@srru.ac.th

บทคัดย่อ

วัตถุประสงค์: งานวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบและพัฒนาระบบสำรวจและวิเคราะห์ข้อมูลที่สามารถใช้พยากรณ์ตัวบ่งชี้ Food Consumption Score (FCS) ได้อย่างแม่นยำในบริบทของประเทศไทย เพื่อสนับสนุนการจัดการปัญหาความมั่นคงทางอาหารอย่างมีประสิทธิภาพ

วิธีการศึกษา: ดำเนินการพัฒนาระบบสำรวจและวิเคราะห์ข้อมูลที่ถูกใช้งานจริงโดยเครือข่ายนักวิจัยด้านความมั่นคงทางอาหารของมหาวิทยาลัยราชภัฏ 10 แห่ง พร้อมนำข้อมูลที่ได้มาประยุกต์ใช้กับอัลกอริทึมเรียนรู้ด้วยเครื่อง ได้แก่ Naïve Bayes, Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbors (kNN) และ Extreme Gradient Boosting (XGBoost) เพื่อสร้างโมเดลพยากรณ์ FCS

ข้อค้นพบ: ผลการศึกษาพบว่าอัลกอริทึม XGBoost มีประสิทธิภาพสูงสุดในการพยากรณ์ FCS ด้วยความแม่นยำที่ไม่น้อยกว่าร้อยละ 99 ซึ่งแสดงถึงศักยภาพในการใช้เป็นเครื่องมือสำหรับการพยากรณ์ความมั่นคงทางอาหารในประเทศไทย

การประยุกต์ใช้จากการศึกษา: ระบบและโมเดลที่พัฒนาขึ้นสามารถใช้เป็นเครื่องมือสำคัญในการติดตามและพยากรณ์ระดับความมั่นคงทางอาหารในระดับครัวเรือน เพื่อสนับสนุนการตัดสินใจและการวางแผนนโยบายที่เกี่ยวข้อง ตลอดจนการจัดการปัญหาความมั่นคงทางอาหารได้อย่างรวดเร็วและแม่นยำ

คำสำคัญ: อัลกอริทึม, ความมั่นคงทางอาหาร, การเรียนรู้ของเครื่อง

Abstract

Purpose: This study aims to design and develop a data collection and analysis system capable of accurately predicting the Food Consumption Score (FCS) within the context of Thailand, supporting effective food security management.

Methodology: A data collection and analysis system was developed and utilized by a network of food security researchers from 10 Rajabhat Universities. The collected data were applied to machine learning algorithms, including Naïve Bayes, Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbors (kNN), and Extreme Gradient Boosting (XGBoost), to build an FCS prediction model.

Findings: The study found that the XGBoost algorithm demonstrated the highest accuracy in predicting FCS, with a precision rate of at least 99%. This highlights its potential as a predictive tool for food security in Thailand.

Application of this study: The developed system and model can serve as critical tools for monitoring and forecasting household-level food security. They can support decision-making, policy planning, and the efficient management of food security challenges with speed and precision.

Keywords: Algorithms, Food security, Machine learning

1. บทนำ

ประเทศไทยได้รับผลกระทบในหลายมิติจากสถานการณ์การแพร่ระบาดของไวรัสโคโรนาสายพันธุ์ใหม่ (COVID-19) โดยเฉพาะครัวเรือนที่อยู่ในภาวะเปราะบางและมีรายได้น้อย ซึ่งต้องเผชิญกับความยากลำบากและการสูญเสียรายได้อย่างฉับพลัน (Eksiri & Charoenratana, 2020) สถานการณ์ดังกล่าวสะท้อนให้เห็นถึงปัญหาที่คล้ายคลึงกับประเทศเวียดนาม ซึ่งประชากรบางกลุ่มต้องเผชิญกับความไม่มั่นคงทางรายได้ในช่วงวิกฤติ (Vu et al., 2022) ทั้งนี้ ความมั่นคงทางอาหารจึงกลายเป็นปัจจัยสำคัญที่ส่งผลกระทบโดยตรงต่อความเป็นอยู่ของครัวเรือน

จากรายงานของดัชนีความมั่นคงทางอาหารโลก (Global Food Security Index: GFSI) (Economist Impact, 2022) ประเทศไทยมีคะแนนความมั่นคงทางอาหาร 65.1 คะแนนอยู่อันดับที่ 52 จาก 113 ประเทศ ในปี พ.ศ. 2562 และตกลงมาที่ 60.1 คะแนนอยู่อันดับที่ 63 ในปี พ.ศ. 2565 ส่วนการประเมินของ Food and Agriculture Organization (FAO) (United Nations, 2022) ช่วงปี พ.ศ. 2560 - 2562 ประเทศไทยมีประชากรร้อยละ 7.8 ขาดสารอาหาร (Eksiri & Charoenratana, 2020) ทำให้การสำรวจ วิเคราะห์ และ

สร้างแบบจำลองจากข้อมูล เพื่อใช้สำหรับสร้างฉากทัศน์วางแผนแก้ปัญหาความมั่นคงทางอาหารที่แม่นยำ จึงเป็นเครื่องมือที่สำคัญ

โดยก่อนหน้านี้มีหลากหลายงานวิจัย (Deléglise et al., 2020, 2022; Martini et al., 2022; Nica-Avram et al., 2021; Razzaq et al., 2021) ได้พยายามใช้การเรียนรู้ของเครื่องสร้างโมเดลสำหรับการพยากรณ์ตัวบ่งชี้ด้านความมั่นคงทางอาหาร เช่น Food Consumption Score (FCS) (World Food Program, 2009), Household Dietary Diversity Score (HDDS), Copping Strategies Index (CSI) (Maxwell & Caldwell, 2008) และ Household Food Insecurity Access Score (HFIAS) (INDDEx Project, 2018) แต่กลไกเหล่านี้ใช้ข้อมูลที่เป็นบริบทและปัจจัยของแต่ละประเทศ ซึ่งการนำโมเดลเหล่านี้มาใช้ในบริบทของประเทศไทยอาจทำให้การพยากรณ์ FCS ได้ผลที่แม่นยำน้อยลงได้

งานวิจัยนี้จึงออกแบบและพัฒนาระบบสำรวจ รายงานผล วิเคราะห์ข้อมูล และสร้างโมเดลสำหรับพยากรณ์ตัวบ่งชี้ FCS ในระดับครัวเรือน พร้อมทั้งประเมินประสิทธิภาพการใช้ Support Vector Machine (SVM) (James et al., 2021), Naïve Bayes (Korb & Nicholson, 2011), Random Forest (RF) (Breiman, 2001), k-Nearest Neighbors (kNN) (Rezaei & Jabbari, 2022) และ Extreme Gradient Boosting (XGBoost) (Chen & Guestrin, 2016) กับข้อมูลที่เป็นบริบทของประเทศไทย ก่อนจะนำระบบดังกล่าวไปใช้ประโยชน์จริงโดยเป็นเครื่องมือในการสำรวจข้อมูลจากกลุ่มเครือข่ายวิจัยของมหาวิทยาลัยราชภัฏ 10 แห่ง ได้แก่ มหาวิทยาลัยราชภัฏสงขลา มหาวิทยาลัยราชภัฏราชนครินทร์ มหาวิทยาลัยราชภัฏสุรินทร์ มหาวิทยาลัยราชภัฏกาญจนบุรี มหาวิทยาลัยราชภัฏภูเก็ต มหาวิทยาลัยราชภัฏสกลนคร มหาวิทยาลัยราชภัฏอุดรดิตถ์ มหาวิทยาลัยราชภัฏกำแพงเพชร มหาวิทยาลัยราชภัฏลำปาง และมหาวิทยาลัยราชภัฏชัยภูมิ ซึ่งมีการสำรวจข้อมูลจำนวน 4,465 ครัวเรือน แล้วประมวลผลตัวบ่งชี้ FCS แต่ละครัวเรือนก่อนที่จะแบ่งข้อมูลออกเป็น 3 ประเภท แล้วนำมาเรียนรู้ด้วย SVM, Naïve Bayes, RF, kNN และ XGBoost ผลการประเมินความแม่นยำของโมเดลพบว่า XGBoost สามารถพยากรณ์ตัวบ่งชี้ FCS ได้แม่นยำสูงที่ร้อยละ 99 จากข้อมูลทั้ง 3 ประเภทได้

2. ทบทวนวรรณกรรมที่เกี่ยวข้อง

2.1 ตัวบ่งชี้ที่ใช้ประเมินระดับความมั่นคงทางอาหารในระดับครัวเรือน (Food Consumption Score)

Food Consumption Score (FCS) เป็นตัวบ่งชี้สำหรับการประเมินความมั่นคงทางอาหารที่ได้รับความนิยมเชื่อถือ และถูกนำไปใช้ประเมินในระดับครัวเรือน โดย FCS จะศึกษาพฤติกรรมการรับประทานอาหารย้อนหลัง 7 วัน มีทั้งหมด 8 หมวดประกอบด้วย 1) หมวดข้าว/ธัญพืช/แป้ง/หัวมัน 2) หมวดพืชตระกูลถั่ว 3) หมวดพืชผัก 4) หมวดผลไม้ 5) เนื้อสัตว์ 6) ผลิตภัณฑ์จากนม 7) น้ำตาล/น้ำผึ้ง และ 8) น้ำมัน/ไขมัน โดยแต่ละหมวดจะถูกแบ่งย่อยลักษณะการกินที่กำกับด้วยค่าน้ำหนัก ก่อนที่จะนำมาคำนวณเป็นคะแนนตัว

บ่งชี้ FCS ตั้งแต่ 0 ถึง 100 คะแนน แล้วนำมาแบ่งเป็น 3 ช่วงชั้น คือ ขัดสน (Poor) อยู่ระหว่าง 0-28 คะแนน คาบเส้น (Borderline) อยู่ระหว่าง 29 ถึง 42 คะแนน และยอมรับได้ (Acceptable) มีคะแนนมากกว่า 42 ซึ่งตัวบ่งชี้ FCS นี้จะช่วยให้เข้าใจสถานการณ์ความมั่นคงทางอาหารของครัวเรือนได้

2.2 การพยากรณ์ความมั่นคงทางอาหารโดยใช้การเรียนรู้ของเครื่อง

มีงานวิจัยหลายฉบับที่ใช้การเรียนรู้ของเครื่องในการประเมินความมั่นคงทางอาหาร ตัวอย่างเช่น Westerveld และคณะ (Westerveld et al., 2021) ซึ่งได้ใช้ข้อมูลจากครัวเรือนและปัจจัยอื่น ๆ ที่อาจส่งผลต่อความมั่นคงทางอาหาร เช่น ความขัดแย้งระหว่างประเทศ การเปลี่ยนแปลงสภาพอากาศ และปัญหาทางเศรษฐกิจ โดยได้ใช้ชุดข้อมูลทั้งหมด 19 กลุ่ม จำนวน 130 ตัวแปร ซึ่งการจัดเก็บข้อมูลถูกแบ่งออกเป็นระยะและรวบรวมจากแหล่งข้อมูลที่หลากหลาย จากนั้นทำการทำความสะอาดข้อมูลและสกัดคุณสมบัติข้อมูล (Feature Extraction) ก่อนที่จะนำข้อมูลที่ได้นำมาฝึกด้วยอัลกอริธึม XGBoost, Random Forest (RF) และ Categorical Boosting (CatBoost) ผลการทดลองพบว่า RF ให้ความแม่นยำสูงสุดในการประเมินความมั่นคงทางอาหาร

การใช้วิธีการเรียนรู้ของเครื่องพยากรณ์ความไม่มั่นคงทางอาหาร เมื่อเกิดการขาดรายได้อย่างกะทันหัน (Income Shock) ในประเทศเวียดนาม โดย Vu และคณะ (Vu et al., 2022) ได้ใช้ข้อมูลระดับครัวเรือนจาก Vietnam Household Living Standard Survey (VHLSS) และ Labor Force Survey (LFS) สํารวจเมื่อ ค.ศ. 2010 - 2018 มาใช้กับการเรียนรู้ของเครื่องที่ออกแบบเฉพาะ และผลการประเมินผลโมเดลของ Vu และคณะ พบว่าหากรายได้ครัวเรือนลดลงร้อยละ 10 ทำให้ครัวเรือนมีความไม่มั่นคงทางอาหารเพิ่มขึ้นร้อยละ 3.5 ซึ่งโมเดลของ Vu และคณะอาจถูกนำไปใช้เพื่อทำนายความไม่มั่นคงทางอาหารของครัวเรือน เมื่อเกิดสถานการณ์ขาดรายได้อย่างกะทันหันได้

Martini et al. (2022) ได้ใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เพื่อพยากรณ์ค่าตัวบ่งชี้ Food Consumption Score (FCS) และ Coping Strategies Index (CSI) ในสถานการณ์ที่ข้อมูลมีจำกัด โดยใช้ข้อมูลจากหลากหลายแหล่ง เช่น ข้อมูลการใช้โทรศัพท์มือถือ ภาพถ่ายดาวเทียม ข้อมูลการเดินเงินโทรศัพท์มือถือ ตำแหน่งทางภูมิศาสตร์ ข้อความจากทวีตในระบบ Twitter และการโฆษณาผ่านสื่อสังคมออนไลน์ ข้อมูลดังกล่าวถูกนำมาเรียนรู้ด้วยเทคนิค SHapley Additive exPlanations (SHAP) เพื่อสร้างโมเดลที่สามารถอธิบายการตัดสินใจของระบบได้อย่างโปร่งใส และถูกนำไปใช้เป็นเครื่องมือในการกำกับติดตามความยากจนและสถานการณ์ความไม่มั่นคงทางอาหารในรูปแบบเรียลไทม์อย่างมีประสิทธิภาพ

Nica-Avram et al. (2021) เล็งเห็นถึงปัญหาของระบบการแลกเปลี่ยนอาหาร (Food Sharing System: FSS) ที่อาจส่งผลให้เกิดการทิ้งอาหารอย่างสูญเปล่าในสหราชอาณาจักร จึงอาศัยข้อมูลการแลกเปลี่ยนอาหารมาใช้ในการเรียนรู้ของเครื่อง ประกอบด้วย RF, Adaptive Boosting (AdaBoost), SVM และ kNN สำหรับการประเมินความไม่มั่นคงทางอาหาร เพื่อใช้สำหรับปรับปรุงวิธีการแลกเปลี่ยนอาหารให้มีประสิทธิภาพมากขึ้น

Razzaq et al. (2021) ใช้ข้อมูลครัวเรือนเรียนรู้ด้วย SVM, Naïve Bayes, kNN, RF และ Logistic Regression สร้างโมเดลพยากรณ์ความมั่นคงทางอาหาร ซึ่ง Razzaq และคณะ พบว่า RF เป็นอัลกอริทึมที่สามารถทำนายความมั่นคงทางอาหารได้แม่นยำที่สุด และเป็นอัลกอริทึมที่น่าสนใจที่จะสามารถนำมาประยุกต์ใช้เพื่อประเมินความมั่นคงทางอาหารได้

จากหลายงานวิจัยที่ใช้การเรียนรู้ของเครื่องข้างต้น เน้นการสร้างโมเดลเพื่อใช้พยากรณ์ตัวบ่งชี้ด้านความมั่นคงทางอาหาร แต่โมเดลเหล่านี้ใช้เครื่องมือที่ถูกออกแบบสำหรับการนำไปใช้ในบริบทเฉพาะของแต่ละประเทศ เพื่อศึกษาบริบทของประเทศไทย งานวิจัยนี้จึงเลือกทดสอบโดยใช้อัลกอริทึม SVM, Naïve Bayes, kNN, RF และ XGBoost ซึ่งสอดคล้องกับงานวิจัยก่อนหน้ามาเปรียบเทียบประสิทธิภาพการพยากรณ์ตัวบ่งชี้ FCS นอกจากนี้ยังมีการพัฒนาระบบเพื่อการสำรวจ วิเคราะห์ และรายผลที่สามารถนำไปใช้ประโยชน์จริงได้

3. วิธีดำเนินการวิจัย

3.1 การออกแบบและพัฒนาระบบนำเข้าข้อมูล

การสร้างแบบฟอร์มสำหรับใช้สำรวจข้อมูลในงานวิจัยนี้ได้พัฒนาขึ้น และถูกใช้สำหรับการนำเข้าข้อมูล แสดงผลข้อมูล วิเคราะห์ข้อมูลสถิติพื้นฐาน ประเมินตัวบ่งชี้ FCS, CSI และ HFIAS พร้อมกับข้อมูลทางภูมิศาสตร์ เพื่อใช้ระบุพื้นที่ที่ประสบปัญหาด้านความมั่นคงทางอาหาร โดยการพัฒนาระบบเริ่มจากการวิเคราะห์แบบสอบถามของกลุ่มเครือข่ายนักวิจัย ซึ่งมี 6 ด้านคือ 1) ข้อมูลพื้นฐานของสมาชิกทุกคนในครัวเรือน 2) ข้อมูลเศรษฐกิจของครัวเรือน 3) การผลิตและการบริโภคอาหารของครัวเรือน 4) สถานการณ์ความมั่นคงทางอาหารของครัวเรือน 5) วิธีการแก้ไขปัญหาความมั่นคงทางอาหารของครัวเรือน และ 6) ความวิตกกังวลในอนาคตต่อการแพร่ระบาดของโรคอุบัติใหม่ด้านความมั่นคงทางอาหารของครัวเรือน รวมทั้งสิ้น 362 ตัวแปร จากนั้นได้วางโครงสร้างระบบออกเป็น 3 ส่วน ประกอบด้วย

1) ฐานข้อมูล ออกแบบและสร้างฐานข้อมูลโดยใช้ MySQL (Oracle, 2022) และได้จัดการโครงสร้างแบบ (Configuration) ให้ทำงานร่วมกับ Laravel Framework (Laravel LLC, 2022) สำหรับการพัฒนา Application Programming Interface (API)

2) API โดยเป็นส่วนที่ใช้เชื่อมต่อระหว่างฐานข้อมูลและคอยให้บริการเข้าถึงข้อมูลโดยใช้สถาปัตยกรรม RESTful

3) ส่วนติดต่อกับผู้ใช้ซึ่งเป็นส่วน นำเข้า วิเคราะห์ รายงานผล และบริหารจัดการข้อมูล โดยการพัฒนาได้จัดลำดับการเข้าถึงข้อมูลตั้งแต่ผู้ใช้ระดับประเทศจนถึงผู้ใช้ระดับหมู่บ้าน และส่วนของผู้ดูแลระบบซึ่งคอยจัดการข้อมูลและสิทธิการเข้าถึงข้อมูล และข้อมูลทั้งหมดจะถูกเชื่อมผ่าน API มายังหน้าตาผู้ใช้ซึ่งถูกพัฒนาโดยใช้ Vue.js (You, 2022)

3.2 พื้นที่การศึกษาและรวบรวมข้อมูล

การศึกษาและรวบรวมข้อมูลในระดับครัวเรือนของงานวิจัยนี้ใช้เครือข่ายมหาวิทยาลัยราชภัฏที่ครอบคลุมพื้นที่จากทุกภูมิภาคในประเทศไทย ประกอบด้วยจังหวัดกาญจนบุรี กำแพงเพชร ฉะเชิงเทรา ชัยภูมิ ภูเก็ต ลำปาง สกลนคร สงขลา สุรินทร์ และอุดรดิตถ์ โดยแบ่งการสำรวจข้อมูลออกเป็น 2 ช่วง คือ ช่วงแรกเดือนธันวาคม พ.ศ. 2565 - มกราคม พ.ศ. 2566 และช่วงที่ 2 เดือนมีนาคม พ.ศ. 2566 ถึงเมษายน พ.ศ. 2566 ด้วยการสุ่มตัวอย่างแบบ Simple Random Sampling (SRS) อย่างน้อยพื้นที่ละ 400 ชุด ผ่านระบบที่พัฒนาขึ้นจากงานวิจัยนี้ โดยแบบสำรวจจะมีรอบข้อมูลรายครัวเรือนที่ประกอบด้วย

- 1) ข้อมูลเบื้องต้น เป็นข้อมูลพื้นฐานของครัวเรือน เช่น อาชีพ รายรับรายจ่าย อายุของหัวหน้าครัวเรือน และอื่น ๆ ที่เป็นหนึ่งในส่วนประกอบสำหรับการประเมินตัวบ่งชี้ FCS ได้
- 2) ข้อมูลพื้นฐานสมาชิกในครัวเรือน เป็นข้อมูลของสมาชิกในครัวเรือนเบื้องต้น เช่น อาชีพหลัก อาชีพเสริม รายได้ ผลกระทบจากโรคระบาด COVID-19
- 3) ข้อมูลทางเศรษฐกิจของครัวเรือน เป็นข้อมูลสถานะทางเศรษฐกิจของครัวเรือน เช่น รายรับรวม รายจ่าย การกู้ยืมหนี้ และการถือครองทรัพย์สิน ที่พักอาศัย และเครื่องมือในการประกอบอาชีพ
- 4) ข้อมูลด้านความมั่นคงทางอาหารในช่วงสถานการณ์ COVID-19 ถูกใช้สำหรับการประเมินตัวบ่งชี้ HFIAS ซึ่งประกอบด้วยประเด็นความเพียงพอของอาหาร การไม่ได้รับประทานอาหาร การจัดการเมื่ออาหารไม่เพียงพอต่อสมาชิกในครัวเรือน
- 5) ข้อมูลสถานการณ์ความไม่มั่นคงทางอาหาร ถูกใช้สำหรับการประเมินตัวบ่งชี้ CSI ซึ่งมุ่งเน้นสอบถามประเด็นการตอบสนองต่อเมื่อเกิดสถานการณ์ความไม่มั่นคงทางอาหารในครัวเรือน
- 6) ข้อมูลด้านความวิตกกังวลในอนาคตต่อการแพร่ระบาดของเชื้อโรค ซึ่งเป็นการสำรวจข้อมูลที่มุ่งเน้นสอบถามประเด็น ความเพียงพอ การเข้าถึง การใช้ประโยชน์ และเสถียรภาพของการเข้าถึงอาหาร ถ้าหากมีการระบาดของโรคอุบัติใหม่

3.3 การเตรียมข้อมูล

หลังจากการลงพื้นที่มีครัวเรือนถูกสำรวจทั้งสิ้น 4,465 ครัวเรือน และงานวิจัยนี้เรียกว่า Household Survey Data (HSD) จากนั้นได้ทำความสะอาดข้อมูล HSD ด้วยการตรวจสอบการป้อนข้อมูลผิดพลาด โดยใช้วิธี Exploratory Data Analysis (EDA), Anomaly Detection (AD) และตรวจสอบค่าสุดโต่ง (Outlier) ก่อนที่จะแจ้งกลับไปยังผู้สำรวจข้อมูลเพื่อตรวจสอบและแก้ไข จากนั้นจึงประมวลผลตัวบ่งชี้ FCS, CSI และ HFIAS ของแต่ละครัวเรือน รวมตัวแปรทั้งหมด 362 ตัวแปร ก่อนที่จะนำ HSD มาประมวลผลเพื่อให้ได้ตัวแปรที่เกี่ยวข้องกับความมั่นคงทางอาหารตามแนวทางของ WFP ตามตารางที่ 1 ซึ่งเรียกข้อมูลกลุ่มนี้ว่า Food Security Related Data (FSRD)

หลังจากได้กลุ่มของตัวแปร FSRD แล้วได้คัดเลือกตัวแปรที่เกี่ยวข้องกับตัวบ่งชี้ FCS ด้วยวิธีวิเคราะห์การถดถอยโลจิสติก (Logistic Regression) โดยใช้ตัวแปรตาม FCS กลุ่ม Poor และ Borderline

มีค่าเท่ากับ 0 และ Acceptable มีค่าเท่ากับ 1 แล้วใช้ Python statsmodels (Perktold, Seabold & Taylor, 2023) ในการวิเคราะห์ข้อมูล และได้ผลการประเมินดังตารางที่ 2 ก่อนที่จะเลือกเฉพาะตัวแปรที่มีความสัมพันธ์กับตัวบ่งชี้ FCS ที่ช่วงความเชื่อมั่นระดับร้อยละ 95 และมีนัยสำคัญทางสถิติที่ $p < 0.05$ ประกอบด้วย WELFARE_TOTAL, EXPENSE_MONTHLY, INCOME_MONTHLY, ASSETS_LAND, WATER_CONSUME และ COVID_AFFECTS ซึ่งเรียกตัวแปรในกลุ่มนี้ว่า Feature Selection of Food Security Related Data (FSofSRD) แล้วนำตัวแปร HSD, FSRD และ FSofSRD สร้างโมเดลสำหรับทำนาย FCS ตามที่อธิบายในหัวข้อขั้นตอนการสร้างโมเดล

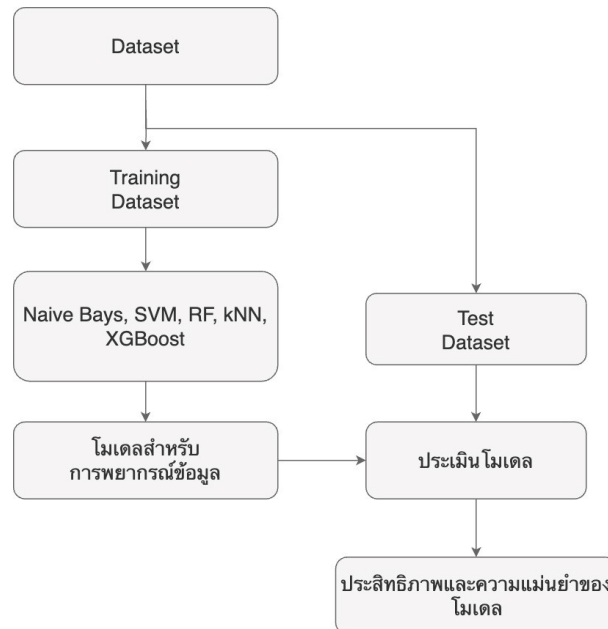
ตารางที่ 1 ชื่อและรายละเอียดของตัวแปร FSRD

ชื่อตัวแปร	รายละเอียด
HOH_GENDER	เพศของหัวหน้าครัวเรือน
HH_SIZE	จำนวนสมาชิกในครัวเรือน
HOH_AGE	อายุหัวหน้าครัวเรือน
HOH_EDU	ระดับการศึกษาของหัวหน้าครัวเรือน
DEPEND_RATIO	สัดส่วนประชากรวัยพึ่งพิงในครัวเรือน
GROUPS_MEMBER	จำนวนกลุ่มสมาชิกในชุมชนที่ครัวเรือนเข้าร่วม
WELFARE_TOTAL	จำนวนสวัสดิการของครัวเรือนที่ได้รับ
EXPENSE_MONTHLY	รายจ่ายของครัวเรือน/เดือน
DEBT_TOTAL	หนี้สินทั้งหมดของครัวเรือน
INCOME_MONTHLY	รายรับของครัวเรือน/เดือน
ASSETS_LAND	ขนาดพื้นที่ที่ดินที่ถือครอง
ASSETS_TOTAL	จำนวนทรัพย์สินที่ถือครอง
WATER_CROP	จำนวนแหล่งน้ำสำหรับการทำการเกษตร
WATER_CONSUME	จำนวนแหล่งน้ำสะอาดที่เข้าถึง
COVID_AFFECTS	จำนวนด้านที่ได้รับผลกระทบจากปัญหา COVID-19
FOOD_STOCKS	จำนวนที่จัดเก็บอาหารเพื่อการบริโภคในครัวเรือน
FOODSRC_DIST	ระยะทางการเข้าถึงแหล่งอาหาร

3.4 ขั้นตอนการสร้างโมเดล

ขั้นตอนการสร้างโมเดลความมั่นคงทางอาหารจะเป็นไปตามภาพที่ 1 จากข้อมูล HSD, FSRD และ FSofSRD ในแต่ละกลุ่มจะถูกนำมาแบ่งออกเป็น 2 ส่วน คือ ชุดข้อมูลสำหรับการเรียนรู้ (Training Set) และชุดข้อมูลสำหรับการประเมิน (Test Set) โดย Training Set เป็นข้อมูลจากการสำรวจรอบแรกมีจำนวน 3,876 ชุด คิดเป็นร้อยละ 86.8 และ Test Set มีจำนวน 589 ชุดคิดเป็นร้อยละ 13.1 ซึ่งเป็นข้อมูลจากการสำรวจครั้งที่สอง โดยข้อมูลทั้ง 2 กลุ่มจะมีสัดส่วนของตัวบ่งชี้ FCS ทั้ง Acceptable, Borderline และ

Poor ใกล้เคียงกัน จากนั้นนำ HSD, FSRD และ FSoFSRD มาเรียนรู้ด้วยอัลกอริทึม SVM, Naïve Bayes, RF, kNN และ XGBoost โดยใช้ไลบรารี Scikit-learn (scikit-learn, 2023) ในภาษาไพทอน (Python) เมื่อได้โมเดลสำหรับพยากรณ์ FCS แล้วจึงประเมินความแม่นยำด้วย Test Set



ภาพที่ 1 ขั้นตอนการสร้างโมเดลสำหรับพยากรณ์ตัวบ่งชี้ FCS

4. ผลการทดลองและประเมินผล

4.1 การกำหนดพารามิเตอร์

ในการทดลองประเมินประสิทธิภาพของโมเดล เพื่อพยากรณ์ตัวบ่งชี้ FCS จากข้อมูล HSD, FSRD และ FSoFSRD ด้วยอัลกอริทึม SVM, Naïve Bayes, RF, kNN และ XGBoost ใช้ค่าพารามิเตอร์ตั้งต้น (Default) ของไลบรารี Scikit-learn

4.2 วิธีประเมินประสิทธิภาพของโมเดล

การประเมินประสิทธิภาพของโมเดลจากงานวิจัยนี้ ใช้ข้อมูลกลุ่ม Test Set ทดสอบด้วย Confusion Matrix เพื่อประเมินความแม่นยำในการพยากรณ์ตัวบ่งชี้ FCS ของครีวรีออนและมีผลลัพธ์ คือ True Positive (TP), True Negative (TN), False Positive (FP) และ False Negative (FN) โดยทั้ง TP, TN, FP และ FN จะถูกนำมาประเมินเพื่อหาค่าความแม่นยำ (Accuracy) ความเที่ยง (Precision) ค่าเรียกคืน (Recall) และ F1-Score ด้วยไลบรารี Scikit-learn ดังต่อไปนี้

1) ความแม่นยำ (Accuracy) ใช้สำหรับประเมินความแม่นยำของโมเดลในการพยากรณ์ค่า FCS ซึ่งค่า Accuracy สูงหมายถึงโมเดลมีความแม่นยำสูง

2) ความเที่ยง (Precision) ใช้สำหรับประเมินความเที่ยงตรงของการจำแนกเฉพาะกลุ่ม TP ซึ่งถ้าหาก Precision สูง จะหมายถึง FP มีสัดส่วนน้อย บ่งบอกถึงโมเดลมีความเที่ยงตรงต่อการพยากรณ์ข้อมูล

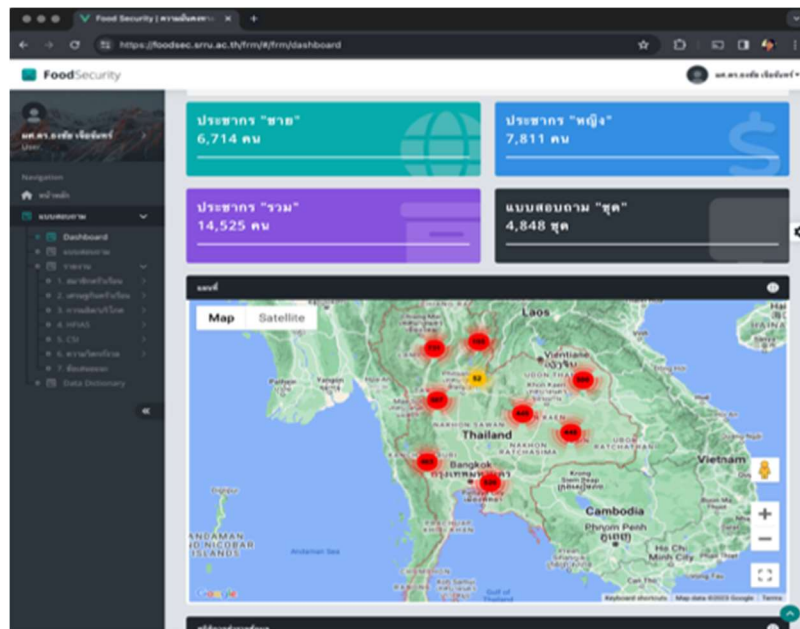
3) ค่าเรียกคืน (Recall) คือการประเมินความแม่นยำของโมเดลในกรณีที่ข้อมูลอยู่ในกลุ่มตัวอย่างที่ถูกประเมินเป็น TP และสามารถจำแนกข้อมูลใน Test Set ได้เป็น TP ได้ถูกต้อง

4) F1-Score ใช้สำหรับประเมินประสิทธิภาพของโมเดลในภาพรวมโดยใช้ Precision และ Recall มาร่วมประมวลผล ซึ่ง F1-Score สูงหมายถึงโมเดลมีประสิทธิภาพสูง

5. ผลการพัฒนาระบบและประเมินผล

5.1 ผลการพัฒนาระบบบริหารข้อมูลความมั่นคงทางอาหาร

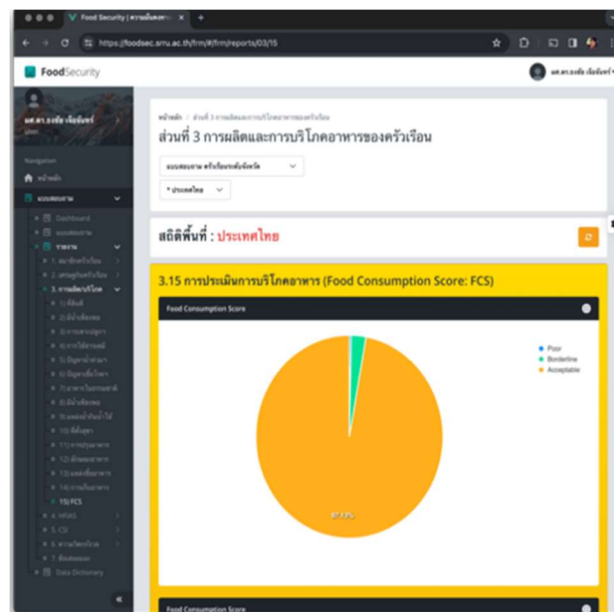
ผลการพัฒนาระบบเป็นดังภาพที่ 2 และภาพที่ 3 ซึ่งถูกนำไปใช้สำรวจ ประมวลผล และรายงานผลข้อมูล พร้อมกับผลการประเมินตัวบ่งชี้ FCS, CSI และ HFIAS ซึ่งระบบนี้จะสามารถนำไปใช้บริหารจัดการเพื่อตอบสนองต่อปัญหาความมั่นคงทางอาหารได้อย่างรวดเร็วดังนี้



ภาพที่ 2 หน้าต่างหลักของระบบบริหารจัดการข้อมูล

1) ระบบการรายงานผลในภาพที่ 2 และภาพที่ 3 ประกอบไปด้วยการรายงานเชิงตัวเลข ตาราง กราฟ และแผนที่ทั้งสิ้น 60 รายการ ซึ่งมีรายงานที่สำคัญ เช่น ข้อมูลครัวเรือน การผลิตและบริโภคอาหาร การเข้าถึงและแหล่งอาหาร ตัวบ่งชี้ FCS, CSI, HFIAS และข้อมูลทางเศรษฐกิจของครัวเรือน ซึ่งสามารถนำไปสู่การตัดสินใจได้อย่างมีประสิทธิภาพ

2) เครื่องมือนำเข้าข้อมูลในภาพที่ 4 เป็นส่วนการนำเข้าข้อมูลซึ่งถูกนำไปใช้สำรวจข้อมูลผ่านเว็บแอปพลิเคชัน ทำให้สามารถนำข้อมูลเข้าสู่ระบบได้จากพื้นที่ที่ทำการสำรวจ และสามารถรับตำแหน่งละติจูดและลองจิจูดจากระบบ Global Positioning System (GPS) ได้อย่างแม่นยำ



ภาพที่ 3 หน้าต่างระบบรายงานผลทางสถิติของข้อมูลครัวเรือนและตัวบ่งชี้ด้านความมั่นคงทางอาหาร

The screenshot shows the 'Add New Household' form in the Food Security system. The form is titled 'แบบสอบถามเรื่องความมั่นคงทางอาหารในช่วงการแพร่ระบาดของโรคโควิด-19 (เก็บข้อมูลระหว่างวันที่ 21 พฤศจิกายน - 30 ธันวาคม 2565)'. It includes fields for household ID, date, and various demographic and economic indicators. The form is divided into sections for 'Household Information', 'Demographic Information', and 'Economic Information'. Each section contains multiple input fields with dropdown menus and checkboxes.

ภาพที่ 4 หน้าต่างการนำเข้าข้อมูลผ่านระบบเว็บแอปพลิเคชัน

ตารางที่ 2 ผลการคัดเลือกคุณสมบัติข้อมูลด้วย Logistic Regression

ตัวแปร	Coef	Std err	z	P> z	CL Lower	CL Upper
HOH_GENDER	0.0556	0.202	0.276	0.783	-0.340	0.451
HH_SIZE	-0.0312	0.11	-0.284	0.776	-0.247	0.184
HOH_AGE	0.0185	0.01	1.771	0.077	-0.002	0.039
HOH_EDU	0.0679	0.073	0.935	0.350	-0.074	0.210

ตัวแปร	Coef	Std err	z	P> z	CL Lower	CL Upper
DEPEND_RATIO	-0.1057	0.519	-0.204	0.839	-1.122	0.911
GROUPS_MEMBER	-0.0146	0.082	-0.178	0.859	-0.175	0.146
WELFARE_TOTAL	0.3439-	0.084	4.111-	0.000*	-0.508	-0.180
EXPENSE_MONTHLY	0.0219	0.0000	3.196	0.001*	0.0000	0.0001
DEBT_TOTAL	0.0207	0.0000	1.905	0.057	-0.0000	0.0000
INCOME_MONTHLY	0.0287-	0.0000	-3.7	0*	-0.0000	0.0000
ASSETS_LAND	-0.7507	0.376	-1.995	0.046*	-1.488	-0.013
ASSETS_TOTAL	0.0136	0.056	0.243	0.808	0.096-	0.123
WATER_CROP	0.3603-	0.249	1.446-	0.148	0.849-	0.128
WATER_CONSUME	0.4483	0.139	3.233	0.001*	0.177	0.72
COVID_AFFECTS	-0.2037	0.068	-3.008	0.003*	-0.336	-0.071
FOOD_STOCKS	0.0437	0.227	0.193	0.847	0.401-	0.488
FOODSRC_DIST	-0.0054	0.037	-0.145	0.884	-0.079	0.068

* หมายถึง $p < 0.05$ ที่ช่วงความเชื่อมั่นร้อยละ 95

5.2 ผลการประเมินความแม่นยำของโมเดล

ตารางที่ 3 ประสิทธิภาพของโมเดลความมั่นคงทางอาหาร

Algorithms	Data Types	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	FSofSRD	0.97	0.96	0.97	0.97
	FSRD	0.97	0.96	0.97	0.97
	HSD	0.28	0.95	0.28	0.42
SVM	FSofSRD	0.98	0.96	0.98	0.97
	FSRD	0.98	0.96	0.98	0.97
	HSD	0.98	0.96	0.98	0.97
RF	FSofSRD	0.97	0.96	0.97	0.97
	FSRD	0.98	0.96	0.98	0.97
	HSD	0.99	0.98	0.99	0.98
kNN	FSofSRD	0.98	0.96	0.98	0.97
	FSRD	0.98	0.96	0.98	0.97
	HSD	0.98	0.96	0.98	0.97
XGBoost	FSofSRD	0.97	0.96	0.97	0.97
	FSRD	0.98	0.97	0.98	0.97
	HSD	0.99	0.99	0.99	0.99

หลังจากนำ HSD, FSRD และ FSoFSRD เรียนรู้ด้วย SVM, Naïve Bayes, RF, kNN และ XGBoost แล้วจะได้โมเดลความมั่นคงทางอาหาร และเมื่อนำมาประเมิน Accuracy, Precision, Recall และ F1-Score ด้วย Test Set แล้วได้ผลดังตารางที่ 1 ซึ่งจะเห็นว่า XGBoost มีประสิทธิภาพสูงกว่าอัลกอริทึมอื่น ๆ ตามรายละเอียดต่อไปนี้

1) Accuracy เมื่อตรวจสอบความแม่นยำในการประเมินตัวบ่งชี้ FCS จะพบว่า XGBoost ให้ความแม่นยำสูงสุดที่ระดับร้อยละ 99 กับข้อมูล HSD และ HSRD ส่วน RF จะให้ความแม่นยำที่ระดับเดียวกันเฉพาะข้อมูล HSD สำหรับอัลกอริทึมที่ให้ความแม่นยำรองลงมาที่ระดับร้อยละ 98 มี SVM และ kNN แต่สามารถนำไปใช้ได้กับข้อมูลทุกประเภท จึงสรุปได้ว่า XGBoost และ RF มีศักยภาพที่จะพยากรณ์ตัวบ่งชี้ FCS ได้แม่นยำเมื่อมีจำนวนตัวแปรมากขึ้น

2) Precision เมื่อประเมินความเที่ยงของการพยากรณ์ FCS พบว่า XGBoost ยังคงให้ความเที่ยงสูงสุดที่ร้อยละ 99 ในกรณีใช้ข้อมูล HSD ในขณะที่อัลกอริทึมที่เหลือสามารถให้ความเที่ยงได้ที่ร้อยละ 96 หรือมากกว่ากับข้อมูลทุกประเภท ยกเว้น Naïve Bayes ที่สามารถทำได้เพียงร้อยละ 95 กับข้อมูล HSD

3) Recall เป็นผลที่เกิดจากข้อมูลใน Test Set ที่เป็น TP และโมเดลสามารถจำแนกเป็น TP ได้ถูกต้อง โดยค่า Recall พบว่า XGBoost ทำได้ที่ระดับร้อยละ 99 กรณีข้อมูล HSD ในขณะที่ Naïve Bayes มีค่า Recall ต่ำสุดเพียงร้อยละ 28 กับข้อมูล HSD ส่วนอัลกอริทึม SVM, RF, kNN ค่อนข้างมีเสถียรภาพที่ดีโดยมีค่า Recall ที่ระดับร้อยละ 97 ขึ้นไปกับข้อมูลทั้ง 3 ประเภท

4) F1-Score เป็นการประเมินประสิทธิภาพในภาพรวมของอัลกอริทึม ซึ่งจะเห็นว่า XGBoost มีประสิทธิภาพสูงสุดที่ร้อยละ 99 และ Naïve Bayes มีประสิทธิภาพต่ำสุดโดยทำได้เพียงร้อยละ 42 เมื่อใช้ข้อมูล HSD สำหรับอัลกอริทึม SVM, RF, kNN และ XGBoost มีประสิทธิภาพที่ร้อยละ 97 ขึ้นไปเมื่อใช้กับข้อมูลทุกประเภท ดังนั้นในบริบทของประเทศไทยนั้น XGBoost มีประสิทธิภาพดีกว่า Naïve Bayes, SVM, RF และ kNN

6. สรุปผลการศึกษา

ความมั่นคงทางอาหารเป็นประเด็นสำคัญที่จำเป็นต้องมีเครื่องมือที่สามารถติดตาม วิเคราะห์ และรายงานผลได้อย่างแม่นยำ เพื่อเป็นแนวทางในการแก้ไขปัญหาได้อย่างมีประสิทธิภาพ ตัวบ่งชี้ Food Consumption Score (FCS) เป็นหนึ่งในเครื่องมือสำคัญที่ถูกนำมาใช้ในการประเมินสถานการณ์ความมั่นคงทางอาหาร อย่างไรก็ตาม งานวิจัยก่อนหน้านี้ที่พยายามประยุกต์ใช้การเรียนรู้โดยเครื่อง (Machine Learning) ในการพยากรณ์ตัวบ่งชี้ FCS ยังขาดการปรับให้เหมาะสมกับบริบทของประเทศไทย และไม่สามารถนำไปใช้แก้ปัญหาในระดับครัวเรือนได้อย่างมีประสิทธิภาพ งานวิจัยนี้มีวัตถุประสงค์หลักสองประการ ได้แก่ 1) การออกแบบและพัฒนาระบบสำหรับสำรวจ วิเคราะห์ และรายงานข้อมูล ความมั่นคงทางอาหาร และรายได้ของครัวเรือน โดยระบบดังกล่าวได้รับการนำไปใช้จริงในการสำรวจข้อมูลร่วมกับกลุ่มเครือข่ายนักวิจัยจากมหาวิทยาลัยราชภัฏ 10 แห่งทั่วประเทศ 2) การสร้างและเปรียบเทียบโมเดลการพยากรณ์ตัวบ่งชี้ FCS โดยใช้อัลกอริทึม Machine Learning ได้แก่ Random Forest (RF) และ Extreme Gradient

Boosting (XGBoost) จากการศึกษพบว่า RF และ XGBoost มีประสิทธิภาพในการพยากรณ์ข้อมูลประเภท Household Dietary Diversity (HSD) ได้อย่างแม่นยำ โดยให้ค่าความแม่นยำไม่น้อยกว่าร้อยละ 99 ผลการวิจัยนี้จึงแสดงถึงศักยภาพของการใช้เทคโนโลยี Machine Learning ในการแก้ปัญหาความมั่นคงทางอาหารของประเทศไทย และสามารถนำไปประยุกต์ใช้ในระดับครัวเรือนได้อย่างมีประสิทธิภาพ

เอกสารอ้างอิง

- Breiman, L. (2001). Random Forests. **Machine Learning**, 45(1), 5–32. <https://doi.org/10.1023/a:1010933404324>
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16**, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Deléglise, H., Bégué, A., Interdonato, R., d'Hôtel, E. M., Roche, M., & Teisseire, M. (2020). Linking Heterogeneous Data for Food Security Prediction. **Communications in Computer and Information Science**, 335–344. https://doi.org/10.1007/978-3-030-65965-3_22
- Deléglise, H., Interdonato, R., Bégué, A., Maître d'Hôtel, E., Teisseire, M., & Roche, M. (2022). Food security prediction from heterogeneous data combining machine and deep learning methods. **Expert Systems with Applications**, 190, 116189. <https://doi.org/10.1016/j.eswa.2021.116189>
- Economist Impact. (2022). **Global Food Security Index (GFSI)**. Retrieved from <https://impact.economist.com/sustainability/project/food-security-index/>
- Eksiri, R. & Charoenratana, S. (2020). The Alternative and Survival of Food Workers in Food Security Amid Covid-19. **Journal of Social Research and Review**, 43(2), 23–54.
- INDDEx Project. (2018). **Household Food Insecurity Access Scale (HFIAS) | INDDEx Project**. Retrieved from <https://inddex.nutrition.tufts.edu/data4diets/indicator/household-food-insecurity-access-scale-hfias>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). **An Introduction to Statistical Learning with Applications in R. In Support Vector Machines**, 2, 597. Springer.
- Korb, K. B., & Nicholson, A. E. (2011). **Bayesian artificial intelligence**. Boca Raton, FL: Crc Press.
- Martini, G., Bracci, A., Riches, L., Jaiswal, S., Corea, M., Rivers, J., Husain, A., & Omodei, E. (2022). Machine learning can guide food security efforts when primary data are not available. **Nature Food**, 3(9), 716–728. <https://doi.org/10.1038/s43016-022-00587-8>
- Maxwell, D., & Caldwell, R. (2008). **The Coping Strategies Index** (Field Methods Manual Second Edition). World Food Programme.
- Nica-Avram, G., Harvey, J., Smith, G., Smith, A., & Goulding, J. (2021). Identifying food insecurity in food sharing networks via machine learning. **Journal of Business Research**, 131, 469–484. <https://doi.org/10.1016/j.jbusres.2020.09.028>

Oracle. (2024). **MySQL**. Retrieved from <https://www.mysql.com/>

Perktold, J., Seabold, S., & Taylor, J. (2023). **Statistical models, hypothesis tests, and data exploration**. Retrieved from <https://www.statsmodels.org/>

Razzaq, A., Ahmed, U. I., Hashim, S., Hussain, A., Qadri, S., Ullah, S., Shah, A. N., Imran, A., & Asghar, A. (2021). An Automatic Determining Food Security Status: Machine Learning based Analysis of Household Survey Data. **International Journal of Food Properties**, 24(1), 726–736. <https://doi.org/10.1080/10942912.2021.1919703>

Rezaei, N., & Jabbari, P. (2022). K-nearest neighbors in R. **Elsevier EBooks**, 181–190. <https://doi.org/10.1016/b978-0-12-822400-7.00006-3>

Scikit-learn. (2019). **Scikit-learn: machine learning in Python**. Retrieved from <https://scikit-learn.org/stable/>

United Nations. (2022). **Food and Agriculture Organization of the United Nations**. Retrieved from <https://www.fao.org/>

Vu, K., Vuong, N. D. T., Vu-Thanh, T.-A., & Nguyen, A. N. (2022). Income shock and food insecurity prediction Vietnam under the pandemic. **World Development**, 153, 105838. <https://doi.org/10.1016/j.worlddev.2022.105838>

Westerveld, J. J. L., Homberg, M. J. C. van den, Nobre, G. G., Berg, D. L. J. van den, Teklesadik, A. D., & Stuit, S. M. (2021). Forecasting transitions in the state of food security with machine learning using transferable features. **Science of The Total Environment**, 786, 147366. <https://doi.org/10.1016/j.scitotenv.2021.147366>

World Food Program. (2009). **Comprehensive Food Security & Vulnerability Analysis Guidelines (Food Security Analysis Service)**. WFP.

You, E. (2022). **The Progressive JavaScript Framework**. Retrieved from <https://vuejs.org/>