Food and Humanity 6 (2026) 101006

Contents lists available at ScienceDirect

Food and Humanity

journal homepage: www.editorialmanager.com/foohum/journal_overview.html

Methodological pitfalls in food security analytics: Addressing biases in regression, SHAP, and feature importance analyses

Yoshiyasu Takefuji

Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan

ARTICLE INFO

Keywords:
Food security analytics
Methodological bias
Feature importance stability
Model-specific limitations
Explainable artificial intelligence

ABSTRACT

This paper critically examines methodological challenges in food security research through analysis of 576 articles published in the Global Food Security journal, introducing a novel "leave-top1-out" validation approach to assess feature importance reliability. Using a public dataset of 12,564 instances with 195 features, we demonstrate how supervised models achieve high prediction accuracy while producing unstable feature importance rankings. Our empirical findings reveal that unsupervised methods and non-target-prediction approaches maintain remarkable consistency in feature rankings despite perturbation, while supervised models and their SHapley Additive exPlanations (SHAP) variants exhibit dramatic shifts in feature importance hierarchies when top predictors are removed. These results substantiate our identification of primary methodological misapplications in ground truth challenges in model interpretation. The supervised models' focus on demographic characteristics rather than direct food security indicators further demonstrates how optimization for prediction accuracy can compromise reliable feature identification without consistency and dose-response relationships validation. To address these limitations, we advocate for a multifaceted analytical framework combining unsupervised techniques with nonlinear nonparametric methods to achieve greater stability in identifying reliable food security determinants. Our complete methodology and implementation code are publicly available on GitHub to promote reproducibility and methodological transparency in food security research.

1. Introduction

The Global Food Security journal has published a substantial body of research employing various analytical methodologies to investigate food security challenges worldwide, including 56 articles using statistical methods and 25 articles applying machine learning approaches (with 5 published in 2025), alongside 19 articles on artificial intelligence (6 in 2025). This extensive collection reflects the multifaceted nature of food security issues that demand diverse analytical approaches. Specifically, the journal contains 27 articles utilizing logistic regression, a statistical method particularly valuable for predicting binary outcomes such as food secure versus insecure households. Logistic regression is favored in this context because many food security classifications are inherently binary in nature, allowing researchers to calculate probabilities of households falling into specific categories based on multiple predictors. Additionally, 48 articles have implemented linear regression techniques, which allow researchers to model continuous dependent variables and quantify relationships between predictors and food security metrics. Linear regression is particularly useful because it enables quantification of the magnitude of effects that different factors have on continuous food security indicators such as caloric intake or food expenditure.

More recently, only 3 articles have employed SHapley Additive exPlanations (SHAP), an emerging machine learning interpretability approach that helps researchers understand complex model predictions by attributing feature importance values. The limited adoption of SHAP is notable because it represents the cutting edge of explainable AI methods that could potentially offer deeper insights into complex food security dynamics that traditional statistical methods might miss.

The prevalence of association-focused methodologies is particularly noteworthy, with 317 articles conducting association analysis and 181 articles performing correlation analysis, collectively demonstrating researchers' strong interest in identifying and quantifying true relationships between variables in food security contexts. This dominance of association methods exists because understanding interconnections between socioeconomic, environmental, and policy factors is fundamental to addressing systemic food security challenges. These association methods enable scholars to uncover patterns and interdependencies among diverse factors affecting global food systems without necessarily

E-mail address: takefuji@keio.jp.

https://doi.org/10.1016/j.foohum.2026.101006
Received 29 July 2025; Received in revised form 7 January 2026; Accepted 8 January 2026
Available online 9 January 2026

implying causation, which is critical because conflating correlation with causation could lead to ineffective policy interventions in vulnerable food systems.

This paper critically examines the methodological pitfalls and challenges researchers face when applying logistic regression, linear regression, and SHAP analysis in food security research, providing a rigorous assessment of common errors, assumption violations, and interpretation limitations that may compromise research validity and reliability in this vital field. This critical examination is essential because methodological flaws can lead to incorrect conclusions that might misdirect limited resources in addressing food insecurity, potentially harming vulnerable populations. The paper offers detailed guidance on proper application of these methods because researchers often lack specialized statistical training despite using increasingly complex analytical tools. Furthermore, the consequences of methodological errors are particularly severe in food security research because findings directly inform interventions affecting human wellbeing and survival, making methodological rigor not merely an academic concern but an ethical imperative.

There are three types of machine learning misapplications: violating fundamental assumptions of data analysis tools, ground truth challenges in model interpretation, and other critical misapplications such as preprocessing including scaling, normalization and transformation. This classification is important because each category represents a distinct pathway through which research validity can be compromised.

When linear methods such as linear regression are applied to nonlinear data, the outcomes are potentially distorted, leading to erroneous interpretations on feature importance analysis against nonlinear data (Anandhi & Nathiya, 2023; Bansal & Singh, 2023; Chen et al., 2023; Janse et al., 2021; Jarantow et al., 2023; Kumar, 2024; Moon et al., 2019; Owoeye et al., 2023; Sahu et al., 2020; Zuur et al., 2009). Linear methods assume that the relationship between variables follows a straight line pattern where the change in the dependent variable is consistently proportional to the change in independent variables across all values. In contrast, nonlinear relationships involve complex patterns where the response variable may change disproportionately or in varying directions depending on the values of predictors, often following curved, threshold-based, or irregular patterns that cannot be represented by simple straight lines. This misapplication is particularly problematic because food security systems often exhibit complex nonlinear relationships that linear models fundamentally cannot capture, resulting in oversimplified understanding of critical dynamics.

When parametric methods such as logistic regression are applied to nonparametric data, similarly the outcomes are inherently skewed, leading to erroneous interpretations (Dey et al., 2025; Pinheiro-Guedes et al., 2024; Rifada et al., 2022; Wang et al., 2023; Work et al., 1989; van Maanen et al., 2019). Parametric methods make specific assumptions about the probability distribution underlying the data (such as normal, binomial, or Poisson distributions) and estimate a fixed number of parameters within these predefined distributional structures. Nonparametric data, however, does not conform to standard probability distributions and requires methods that adapt to the data's structure without imposing distributional constraints, allowing for greater flexibility in modeling irregular patterns or unexpected relationships. This mismatch occurs because parametric methods impose specific distributional assumptions that may not reflect the true nature of food security data, which often exhibits irregular patterns due to complex socio-ecological interactions and threshold effects that parametric models cannot adequately represent. Violating assumptions of data analysis tools inherently distorts the outcomes including feature importance, odd ratio, p-values and other measured metric scores, which can lead researchers to draw fundamentally flawed conclusions about the drivers and dynamics of food security systems despite seemingly robust statistical results.

While supervised machine learning models such as linear regression and logistic regression possess ground truth values for target prediction accuracy validation, feature importances derived from models lack its ground truth for accuracy validation. This distinction is critical because researchers often implicitly assume that a model with high prediction accuracy must also yield reliable feature importance rankings, which is not necessarily true, which is called ground truth challenges in model interpretation. Due to the absence of ground truth, different models generate distinct feature importances, which are called model specific nature, potentially leading to biased feature importances. This model specificity is problematic because food security researchers often seek to identify key drivers of food insecurity to inform policy interventions, but may unknowingly base recommendations on model-dependent artifacts rather than robust relationships.

Supervised machine learning models have two types of accuracy: target prediction accuracy and feature importance reliability, which are distinct issues. This dual nature of model evaluation is frequently overlooked because conventional model assessment focuses primarily on prediction metrics while neglecting feature importance validity. Feature importance refers to contributions of prediction rather than true associations between variables. This conceptual distinction matters because policy makers need to understand actual causal drivers of food insecurity, not merely statistical contributors to model predictions. Consequently, high target prediction accuracy does not guarantee reliable feature importances (Parr et al., 2024; Molnar et al., 2022; Fisher et al., 2019; Lenhof et al., 2024; Mandler & Weigand, 2024; Potharlanka & Bhat, 2024), which is concerning because research conclusions about key factors affecting food security might be based on misleading feature importance rankings despite seemingly robust models.

The implementation of explain=SHAP(model) implies that SHAP solely relies on given model, inherits and inherently amplifies biases in feature importances derived from the model (Bilodeau et al., 2024; Hooshyar & Yang, 2024; Huang & Marques-Silva, 2024; Kumar et al., 2021; Lones, 2024; Létoffé et al., 2025; Molnar et al., 2022; Wu, 2025). This dependency is often overlooked because SHAP is frequently treated as an objective explanation method rather than a reflection of the underlying model's biases and limitations. Therefore, explanation with SHAP(model) propagates and may amplify biases in feature importances, leading to erroneous interpretations. This amplification effect is particularly problematic in food security research because SHAP visualizations carry persuasive power that may lead stakeholders to place unwarranted confidence in flawed feature importance rankings, potentially misdirecting resources and policy attention away from truly important determinants of food security outcomes. The seductive clarity of SHAP visualizations can mask fundamental issues in the underlying model, creating a false sense of understanding that may be more dangerous than acknowledged uncertainty about complex food security dynamics.

This paper makes significant and novel contributions to the methodological literature by providing the first comprehensive framework for identifying, categorizing, and addressing critical statistical misapplications in food security research. By systematically documenting how violations of fundamental assumptions, ground truth challenges, and preprocessing errors compromise research validity, we establish a new methodological standard for the field. Our empirical demonstrations using public datasets (Ogundari, 2023) conclusively reveal how seemingly robust models can generate fundamentally misleading feature importance rankings—a finding with profound implications for resource allocation and policy design. Unlike previous work that merely identifies isolated statistical issues, our research provides an integrated analytical approach that enables researchers to distinguish between reliable and unreliable model interpretations. This breakthrough has immediate practical applications, offering researchers clear pathways to improve methodological rigor and enhance the reliability of feature importance analyses in line with recent statistical advances (Létoffé et al., 2025; Molnar et al., 2022; Parr et al., 2024). By establishing this methodological foundation, our work directly strengthens the evidence base for food security interventions, potentially improving outcomes for millions of vulnerable people worldwide through more accurately targeted policies and programs based on methodologically sound research.

2. Methods

The public dataset consists of 12,564 instances and 195 features (Ogundari, 2023). To rigorously validate true feature associations, we implemented a "leave-top1-out" approach that examines both consistency and dose-response relationships. Our methodology follows a structured process: first, we identify the top 10 features from the complete dataset (CV10). Next, we remove the highest-ranked feature to create a reduced dataset, from which we select the top 9 features (CV9). By comparing feature importance rankings between these two sets, we can assess the stability of feature selection algorithms under perturbation. 5-fold cross-validation is conducted to examine prediction accuracy for both feature sets, providing dual metrics of performance: prediction accuracy and feature importance consistency. This approach reveals how feature importance hierarchies reorganize when the dominant predictor is removed—a critical test of underlying consistency in feature-outcome relationships. Our analysis employs diverse feature selection methods, including supervised algorithms (Random Forest, XGBoost), unsupervised approaches (Highly Variable Gene Selection), and correlation-based techniques (Spearman correlation), providing comprehensive insights into feature importance from multiple analytical perspectives.

3. Results

For purposes of reproducibility and transparency, Python code, shapanalysis.py is publicly available at GitHub (GitHub, 2025).

Table 1 reveals striking patterns in feature selection stability across different algorithms. Notably, the unsupervised method (HVGS) and non-target-prediction approach (Spearman correlation) demonstrate remarkable consistency in feature importance rankings compared to supervised models. Despite achieving lower overall cross-validation accuracy (HVGS: 0.7229 for CV10, 0.6996 for CV9), the HVGS model maintained perfect stability in feature rankings when the top feature was removed, preserving the exact same order for the remaining features. Similarly, Spearman correlation showed exceptional consistency in feature importance, with identical top 4 features between CV9 and the corresponding subset of CV10, while still delivering strong predictive performance (1.0 for CV10, 0.7919 for CV9).

Table 1
cross-validation accuracy and feature rankings per algorithm.

Method	CV10	CV9	top5 feature rankings of CV10	top4 feature rankings of CV9
RF	1.0	0.7941	curfoodsuf, childfood, expns_dif, foodrsnrv4, foodrsnrv1	childfood, expns_dif, **foodrsnrv1, foodrsnrv4**
XGB	1.0	0.8166	curfoodsuf, ms, agenid_birth, genid_describe, sexual_orientation	**childfood, expns_dif, foodrsnrv4, foodrsnrv1**
HVGS	0.7229	0.6996	pweight, est_msa, hweight, foodrsnrv1, kids_5_11y	est_msa, hweight, foodrsnrv1, kids_5_11y
Spearman	1.0	0.7919	curfoodsuf, childfood, expns_dif, foodrsnrv1, foodrsnrv4	childfood, expns_dif, foodrsnrv1, foodrsnrv4
RF-SHAP	1.0	0.7941	curfoodsuf, childfood, expns_dif, foodrsnrv4, foodrsnrv1	childfood, expns_dif, **foodrsnrv1, foodrsnrv4**
XGB-SHAP	1.0	0.8166	curfoodsuf, ms, agenid_birth, genid_describe, sexual_orientation	**childfood, expns_dif, foodrsnrv4, foodrsnrv1**

In contrast, as shifted variables with bold fonts, supervised models like RF, XGB, and their SHAP variants achieved higher accuracy but demonstrated considerable instability in feature rankings. XGB and XGB-SHAP, while attaining perfect accuracy in CV10 and the highest CV9 accuracy (0.8166), showed dramatic shifts in important features after removing the top predictor. RF and RF-SHAP models exhibited the same pattern of ranking instability despite their high accuracy. Particularly revealing is how XGB models identified entirely different feature sets (focused on demographic characteristics like 'ms', 'agenid_birth') compared to other models that prioritized direct food security indicators.

These findings suggest that while supervised models may optimize for prediction accuracy, unsupervised and correlation-based approaches offer superior stability in feature identification—a critical consideration for research applications where consistent feature importance is essential for establishing reliable associations.

4. Discussion

Our findings illuminate critical methodological concerns when applying machine learning approaches to food security research. The implementation of explain=SHAP(model) reveals a fundamental dependency often overlooked in current research: SHAP inherently relies on the underlying model it explains, thereby inheriting and potentially amplifying any biases in feature importance rankings derived from that model (Bilodeau et al., 2024; Hooshyar & Yang, 2024; Huang & Marques-Silva, 2024; Kumar et al., 2021; Lones, 2024; Létoffé et al., 2025; Molnar et al., 2022; Wu, 2025). This dependency undermines the perception of SHAP as an objective explanation method, when it is more accurately characterized as a reflection of the underlying model's biases and limitations.

Our empirical analysis using the Ogundari (2023) dataset demonstrates that supervised models—both with and without SHAP explanations—produce inherently unstable feature importance rankings. When the top feature is removed, these models dramatically reorganize their feature hierarchies, suggesting they capture circumstantial correlations rather than fundamental relationships. This instability persists regardless of high cross-validation accuracy, highlighting the dangerous disconnect between predictive performance and explanatory reliability.

Due to the absence of methods for calculating true associations between variables, we advocate for multifaceted approaches combining unsupervised machine learning models such as highly variable gene selection (HVGS), followed by nonlinear nonparametric statistical methods such as Spearman's correlation with p-values for monotonic relationships. This multifaceted approach addresses the fundamental limitations of single-method analyses that fail to capture the complex realities of food security determinants.

Though originally developed for genomic research, HVGS offers powerful applications in food security analysis by identifying features with the highest variability across observations, effectively highlighting factors demonstrating meaningful differences between food secure and insecure populations. This technique prioritizes features based on inherent variability rather than fit to a particular model structure, making it less susceptible to model-specific biases.

Spearman's correlation provides robust assessments of monotonic relationships without requiring restrictive assumptions of linearity or normal distribution that plague conventional methods, making it ideally suited for complex, nonlinear relationships in food security systems. As our results demonstrate, unsupervised and correlation-based approaches maintain substantially greater stability in feature rankings compared to conventional supervised machine learning models. This stability is crucial for policy-making because unstable rankings can lead to dramatically different conclusions about intervention priorities depending on model selection. Feature ranking stability testing through our "leave-top1-out" approach provides a straightforward yet powerful validation method, revealing whether identified relationships are robust or merely artifacts of particular model specifications.

Our study has several limitations that warrant consideration. While we demonstrate the instability of supervised models and SHAP explanations in feature importance rankings, we do not fully explore the mathematical mechanisms behind this instability. Future research should investigate the theoretical underpinnings of feature importance instability across different model architectures. Our analysis uses a single dataset, albeit a comprehensive one. The generalizability of our findings should be tested across multiple food security datasets from different regions and contexts to establish broader validity. Additionally, while we explored several common machine learning approaches, future work should expand to include other emerging methods such as causal machine learning and structural equation modeling.

While we identify the limitations of current approaches, our proposed multifaceted methodology requires further validation through simulation studies with known ground-truth associations. Such studies would provide clearer evidence of which methods most accurately recover true feature relationships under different data conditions. The practical implementation challenges of our proposed approach in resource-constrained settings deserve attention. Future work should develop simplified frameworks and accessible tools that enable food security researchers and policymakers with varying technical backgrounds to implement these more robust analytical approaches.

In conclusion, our findings call for a fundamental shift in how feature importance is assessed in food security research—moving from overreliance on supervised models and their explanatory tools toward more robust, multi-method approaches that emphasize stability and consistency in feature identification. By implementing this comprehensive methodology, researchers can develop more reliable insights into the true determinants of food security, ultimately enabling more effective and targeted interventions that address genuine causal factors rather than statistical artifacts or model-dependent relationships.

CRediT authorship contribution statement

Takefuji Yoshiyasu: Writing – review & editing, Writing – original draft, Validation, Investigation, Conceptualization.

Consent to participate

Not applicable

Consent for publication

Not applicable

Ethics approval

Not applicable

Declaration of Generative AI and AI-assisted technologies in the writing process

Not applicable

Code availability

Not applicable.

According to ScholarGPS

Yoshiyasu Takefuji holds notable global rankings in several fields. He ranks 25th out of 1287,415 scholars in life sciences, 22nd out of 805,705 in COVID-19, and 1st out of 109,919 in environmental sciences.

Funding

This research has no fund.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Not applicable.

References

Anandhi, P., & Nathiya, E. (2023). Application of linear regression with their advantages, disadvantages, assumption and limitations. *International Journal of Statistics and Applied Mathematics, 8*(6), 133–137. https://doi.org/10.22271/maths.2023.v8.i6b.1463

Bansal, S., & Singh, G. (2023). Multiple linear regression based analysis of weather data: Assumptions and limitations. ICACIS 2023. In R. N. Shaw, M. Paprzycki, & A. Ghosh (Eds.), *Advanced communication and intelligent systems, 1920*. Cham: Springer. https://doi.org/10.1007/978-3-031-45121-8_19.

Bilodeau, B., Jaques, N., Koh, P. W., & Kim, B. (2024). Impossibility theorems for feature attribution. *Proceedings of the National Academy of Sciences, 121*(2), Article e2304406120. https://doi.org/10.1073/pnas.2304406120

Chen, M., Papadikis, K., Jun, C., & Macdonald, N. (2023). Linear, nonlinear, parametric, and nonparametric regression models for nonstationary flood frequency analysis. *Journal of Hydrology, 616*, Article 128772. https://doi.org/10.1016/j.jhydrol.2022.128772

Dey, D., Haque, M. S., Islam, M. M., et al. (2025). The proper application of logistic regression model in complex survey data: A systematic review. *BMC Medical Research Methodology, 25*, 15. https://doi.org/10.1186/s12874-024-02454-5

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research, 20*, 177.

GitHub. (2025). sshanalysis.py. https://github.com/y-takefuji/foodsecurity.

Hooshyar, D., & Yang, Y. (2024). Problems with SHAP and LIME in interpretable AI for education: A comparative study of post-hoc explanations and neural-symbolic rule extraction. *IEEE Access, 12*, 137472–137490. https://doi.org/10.1109/ACCESS.2024.3463948

Huang, X., & Marques-Silva, J. (2024). On the failings of Shapley values for explainability. *International Journal of Approximate Reasoning, 171*, Article 109112. https://doi.org/10.1016/j.ijar.2023.109112

Janse, R. J., Hoekstra, T., Jager, K. J., et al. (2021). Conducting correlation analysis: Important limitations and pitfalls. *Clinical Kidney Journal, 14*(11), 2332–2337. https://doi.org/10.1093/ckj/sfab085

Jarantow, S. W., Pisors, E. D., & Chiu, M. L. (2023). Introduction to the use of linear and nonlinear regression analysis in quantitative biological assays. *Current Protocols, 3*(6), Article e801. https://doi.org/10.1002/cpz1.801

Kumar, A. (2024). Comparative evaluation of linear and nonlinear regression models in predicting VLC channel response and BER performance. *Journal of Optics*. https://doi.org/10.1007/s12596-024-02361-4

Kumar, I., Scheidegger, C., Venkatasubramanian, S., & Friedler, S. (2021). Shapley residuals: Quantifying the limits of the Shapley value for explanations. *Advances in Neural Information Processing Systems, 34*, 26598–26608.

Létoffé, O., Huang, X., & Marques-Silva, J. (2025). Towards trustable SHAP scores. *Proceedings of the AAAI Conference on Artificial Intelligence, 39*(17), 18198–18208. https://doi.org/10.1609/aaai.v39i17.34002

Lenhof, K., Eckhart, L., Rolli, L. M., & Lenhof, H. P. (2024). Trust me if you can: A survey on reliability and interpretability of machine learning approaches for drug sensitivity prediction in cancer. *Briefings in Bioinformatics, 25*(5), Article bbae379. https://doi.org/10.1093/bib/bbae379

Lones, M. A. (2024). Avoiding common machine learning pitfalls. *Patterns, 5*(10), Article 101046. https://doi.org/10.1016/j.patter.2024.101046

Mandler, H., & Weigand, B. (2024). A review and benchmark of feature importance methods for neural networks. *ACM Computing Surveys, 56*(12), 318. https://doi.org/10.1145/3679012

Molnar, C., König, G., Herbinger, J., Freiesleben, T., Dandl, S., Scholbeck, C. A., et al. (2022). In A. Holzinger, R. Goebel, R. Fong, T. Moon, K. R. Müller, & W. Samek (Eds.), *General pitfalls of model-agnostic interpretation methods for machine learning models, 13200* p. 4). Springer. https://doi.org/10.1007/978-3-031-04083-2_4.

Moon, K. R., van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D. B., Chen, W. S., et al. (2019). Visualizing structure and transitions in high-dimensional biological data. *Nature Biotechnology, 37*(12), 1482–1492. https://doi.org/10.1038/s41587-019-0336-3

Ogundari, K. (2023). Food assistance and child food security (Version 1) [Data set]. *Mendeley Data*. https://doi.org/10.17632/r63rcnddts.1

Owoeye, O. R., Oluwole, A. M., Jolayemi, O. S., et al. (2023). Linear and nonlinear regression modeling of the chemical, physical and quality variations in Cardaba banana (Musa acuminata x balbisiana – ABB) during ripening. *Food Measure, 17*, 12–23. https://doi.org/10.1007/s11694-022-01570-4

Parr, T., Hamrick, J., & Wilson, J. D. (2024). Nonparametric feature impact and importance. *Information Sciences, 653*, Article 119563. https://doi.org/10.1016/j.ins.2023.119563

Pinheiro-Guedes, L., Martinho, C., & Martins, M. R. (2024). Logistic regression: Limitations in the estimation of measures of association with binary health outcomes. *Acta Megyeédica Portuguesa, 37*(10), 697–705. https://doi.org/10.20344/amp.21435

Potharlanka, J. L., & Bhat, M., N. (2024). Feature importance feedback with Deep Q process in ensemble-based metaheuristic feature selection algorithms. *Scientific Reports, 14*(1), 2923. https://doi.org/10.1038/s41598-024-53141-w

Rifada, M., Chamidah, N., & Ningrum, R. A. (2022). Estimation of nonparametric ordinal logistic regression model using generalized additive models (GAM) method based on local scoring algorithm. *AIP Conference Proceedings, 2668*(1), Article 070013. https://doi.org/10.1063/5.0111771

Sahu, P., Kang, J., Erdemci-Tandogan, G., & Manning, M. L. (2020). Linear and nonlinear mechanical responses can be quite different in models for biological tissues. *Soft Matter, 16*(7), 1850–1856. https://doi.org/10.1039/c9sm01068h

van Maanen, L., Katsimpokis, D., & van Campen, A. D. (2019). Fast and slow errors: Logistic regression to identify patterns in accuracy–response time relationships. *Behavior Research Methods, 51*, 2378–2389. https://doi.org/10.3758/s13428-018-1110-z

Wang, T., Tang, W., Lin, Y., & Su, W. (2023). Semi-supervised inference for nonparametric logistic regression. *Statistics in Medicine, 42*(15), 2573–2589. https://doi.org/10.1002/sim.9737

Work, J. W., Ferguson, J. G., & Diamond, G. A. (1989). Limitations of a conventional logistic regression model based on left ventricular ejection fraction in predicting coronary events after myocardial infarction. *American Journal of Cardiology, 64*(12), 702–707. https://doi.org/10.1016/0002-9149(89)90751-0

Wu, L. (2025). A review of the transition from Shapley values and SHAP values to RGE. *Statistics*, 1–23. https://doi.org/10.1080/02331888.2025.2487853

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). Limitations of linear regression applied on ecological data. *Mixed effects models and extensions in ecology with R. Statistics for biology and health*. New York, NY: Springer. https://doi.org/10.1007/978-0-387-87458-6_2