Letter to the Editor

Letter to the Editor-"Reply-letter to the editor"

Huang et al. noted the use of SHAP and LIME methods in clinical applications in their reply [1]. However, they must acknowledge that these techniques depend on machine learning models, which inherently introduce biases in feature importance due to their model-specific nature [2–4]. As a result, different models can yield inconsistent feature importance rankings, even when genuine associations between the target and features exist. This reliance on machine learning for calculating feature importance can lead to erroneous and misleading conclusions, compromising the interpretability of results and undermining the overall validity of research findings [2–4]. A more effective approach would involve employing robust statistical methods [5] that accurately reveal true associations, thereby enhancing the integrity of the analysis.

Many researchers, including Huang et al., need to recognize the critical distinction between machine learning and feature importance. While the primary goal of machine learning is to accurately predict outcomes, feature importance aims to reveal the true associations between the target variable and its features. However, due to the model-specific nature of machine learning approaches, these models often fail to generate genuine associations [2–4]. This limitation can lead to misleading interpretations and potentially erroneous conclusions, emphasizing the necessity for researchers to employ robust statistical methods that can more faithfully uncover the underlying relationships in the data. By understanding this discrepancy, researchers can better align their analytical strategies with their research objectives, ensuring more reliable results.

Instead of using these potentially biased feature importances, it's essential to focus on establishing true associations through robust statistical methods, such as Chi-squared tests with p-values and/or Spearman's correlation with p-values [5]. These methods provide a clearer and more accurate representation of relationships within the data. This paper emphasizes the limitations of SHAP and LIME due to their reliance on biased machine learning models and advocates for the use of robust statistical techniques to derive meaningful and valid associations, ultimately enhancing the integrity of research findings. By prioritizing true statistical associations over biased feature importances, researchers can make more informed decisions and improve the reliability of their analyses.

Consent to participate

Not applicable.

Ethics approval

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Code availability

Not applicable.

Authors' contributions

Yoshiyasu Takefuji completed this research and wrote this article.

Funding

Conflict of interest

The author has no conflict of interest.

References

[1] Huang W, Wang C, Chen J. Letter to the editor, reply-letter to the editor. Clin Nutr 2024;43(9):2283–4. https://doi.org/10.1016/j.clnu.2024.07.046.
[2] Danesh T, Ouaret R, Floquet P, Negny S. Hybridization of model-specific and model-agnostic methods for interpretability of neural network predictions: application to a power plant. Comput Chem Eng 2023;176:108306. https://doi.org/10.1016/j.compchemeng.2023.108306.
[3] Ibrahim SA, Balzter H, Tansey K. Machine learning feature importance selection for predicting aboveground biomass in African savannah with Landsat 8 and ALOS PALSAR data. Machine Learn Appl 2024;16:100561. https://doi.org/10.1016/j.mlwa.2024.100561.

[4] Mandler H, Weigand B. Feature importance in neural networks as a means of interpretation for data-driven turbulence models. Comput Fluids 2023;265: 105993. https://doi.org/10.1016/j.compfluid.2023.105993.

[5] Chen Y, Liu S, Ren Z, Wang F, Liang Q, Jiang Y, et al. Cross-ancestry analysis of brain QTLs enhances interpretation of schizophrenia genome-wide association studies. Am J Hum Genet 2024. https://doi.org/10.1016/j.ajhg.2024.09.001.

Yoshiyasu Takefuji
Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku,
Tokyo 135-8181, Japan
E-mail address: takefuji@keio.jp.

30 October 2024