European Journal of Internal Medicine xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Internal Medicine

journal homepage: www.elsevier.com/locate/ejim

Letter to the Editor

Evaluating feature importance biases in logistic regression: Recommendations for robust statistical methods

ARTICLE INFO

Keywords:
Logistic regression
Feature importance
Statistical bias
Spearman's correlation
Chi-squared tests
Mortality risk factors

Dear Editor,

Paiva et al. examined the patterns of intra-abdominal infections in critically ill, immunocompromised patients and aimed to identify mortality risk factors [1]. They employed logistic regression analysis with a logit link function to evaluate the relationship between various factors and mortality. The performance of their models was assessed by plotting the receiver operating characteristic (ROC) curve, calculating sensitivity and specificity, and determining the area under the curve (AUC). Two distinct regression models were developed: one for the entire cohort to investigate the link between 'immunosuppression' and mortality, and another specifically for the subgroup of immunocompromised patients to identify specific risk factors associated with mortality.

However, a significant issue arises from the biased nature of feature importances generated by logistic regression and other models [2–7]. This bias stems from the inherent model-specific characteristics, which lead to different models producing varying feature importance scores that consistently fails to accurately represent true associations between the target variable and its predictors, leading to erroneous conclusions [2–7]. While ROC curves and AUC are valuable metrics for assessing prediction accuracy, they do not provide insight into the validity of the associations calculated between features and the target variable. This paper highlights the reasons why logistic regression generates biased feature importances, leading to erroneous conclusions [2–7] and strongly advocates for the use of robust statistical methods [8–10], such as Spearman's correlation with accompanying p-values and Chi-squared tests with p-values, to uncover genuine associations without bias. Furthermore, Paiva et al. should reevaluate their studies using these unbiased association measures, rather than relying on biased feature importances, to ensure that their conclusions are valid and reliable.

Logistic regression is a widely used method for binary outcomes, especially in fields like epidemiology, social sciences, and machine learning. However, it can produce biased estimates of feature importance under certain conditions [2–7]. There are several reasons, from algorithmic perspectives, that can help us understand why logistic regression induces biased feature importances.

Firstly, logistic regression operates under linear assumptions regarding the relationship between the log-odds of the outcome and the independent variables. If the true relationship deviates from linearity or if significant interaction effects among features are not accounted for, the estimated coefficients—and consequently the feature importances—may be biased. When interaction terms are omitted, the impacts of individual features are misestimated. For example, if two features interact, treating them as independent can lead to incorrect conclusions about their importance.

Multicollinearity presents another major issue. When two or more features are highly correlated, it becomes challenging to isolate their individual effects on the outcome. This correlation can lead to unstable coefficient estimates in logistic regression, making feature importance assessments either inflated or deflated. Moreover, large standard errors for the coefficients of correlated variables result in misleading significance tests, causing some features to appear important purely due to their correlation with other significant variables, rather than their actual contribution.

Sample bias and representation issues also contribute to biased feature importances. Logistic regression can inherit biases from the training data; if certain classes are over- or under-represented, the learned coefficients might reflect this imbalance, skewing the perceived importance of features. Consequently, a model that performs well on a biased sample may struggle when applied to a more representative dataset, leading to distorted feature importance conclusions.

Non-monotonic relationships present another challenge. Logistic regression models may assign importance to features characterized by such relationships—where a feature could increase the risk of an outcome only until a certain point, after which it decreases the risk beyond that point. Without appropriate transformations in the model, logistic regression fails to capture these complexities, thus introducing bias in feature imports.

Regularization techniques, such as Lasso and Ridge regression, can introduce additional bias into coefficient estimates. For instance, Lasso (L1 regularization) can shrink some coefficients to zero, leading to omitted feature importance and a loss of information about a feature's actual impact on the outcome. Although regularization helps control

https://doi.org/10.1016/j.ejim.2024.11.022
Received 11 November 2024; Accepted 26 November 2024

Please cite this article as: Yoshiyasu Takefuji, *European Journal of Internal Medicine,* https://doi.org/10.1016/j.ejim.2024.11.022

overfitting, it can distort the true importance of features.

Feature scaling also plays a crucial role. Logistic regression is sensitive to the scale of features, and if the features are not standardized or normalized, those with larger ranges may disproportionately influence the coefficients. This can lead to biased perceptions of feature importance.

Finally, the interpretation of logistic regression coefficients must be approached with caution. The coefficients are interpreted as the change in log-odds of the outcome for a one-unit increase in the predictor, which can be misleading. For example, a small coefficient may not necessarily signify low importance if the variable is measured on a large scale.

In conclusion, while logistic regression is a widely used method, caution is necessary when interpreting feature importances due to potential biases rooted in its assumptions and the nature of the data. To obtain more reliable estimates of feature importance, researchers should consider employing robust statistical methods, such as Spearman's rank correlation coefficient or chi-squared tests, which can provide unbiased insights into the associations between features and outcomes. These methods can help circumvent some of the pitfalls associated with model-specific feature importance interpretation, thus leading to more valid conclusions regarding the relevance of different features in the predictive context.

Funding

This research has no fund.

Ethics approval

Not applicable.

Consent to participate

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.

Declaration of competing interest

The author has no conflict of interest.

References

[1] Paiva J-A, Rello J, Eckmann C, et al. Intra-abdominal infection and sepsis in immunocompromised intensive care unit patients: disease expression, microbial aetiology, and clinical outcomes. Eur J Intern Med 2024;129:100–10. https://doi.org/10.1016/j.ejim.2024.07.019.

[2] Steyerberg EW, Eijkemans MJ, Habbema JD. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. J Clin Epidemiol 1999;52 (10):935–42. https://doi.org/10.1016/S0895-4356(99)00103-1.

[3] Cornish RP, Bartlett JW, Macleod J, Tilling K. Complete case logistic regression with a dichotomized continuous outcome led to biased estimates. J Clin Epidemiol 2023;154:33–41. https://doi.org/10.1016/j.jclinepi.2022.11.022.

[4] Das U, Maiti T, Pradhan V. Bias correction in logistic regression with missing categorical covariates. J Stat Plan Inference 2010;140(9):2478–85. https://doi.org/10.1016/j.jspi.2010.02.018.

[5] Kost S, Rheinbach O, Schaeben H. Using logistic regression model selection towards interpretable machine learning in mineral prospectivity modeling. Geochemistry 2021;81(4):125826. https://doi.org/10.1016/j.chemer.2021.125826.

[6] Levin B, Paik MC. The unreasonable effectiveness of a biased logistic regression procedure in the analysis of pair-matched case-control studies. J Stat Plan Inference 2001;96(2):371–85. https://doi.org/10.1016/S0378-3758(00)00217-2.

[7] Gorelick MH. Bias arising from missing data in predictive models. J Clin Epidemiol 2006;59(10):1115–23. https://doi.org/10.1016/j.jclinepi.2004.11.029.

[8] Jiang J, Zhang X, Yuan Z. Feature selection for classification with Spearman's rank correlation coefficient-based self-information in divergence-based fuzzy rough sets. Expert Syst Appl 2024;249(Pt B):123633. https://doi.org/10.1016/j.eswa.2024.123633.

[9] Li X, Ma Y, Zhou Q, Zhang X. Sparse large-scale high-order fuzzy cognitive maps guided by Spearman correlation coefficient. Appl Soft Comput 2024;167(Pt A):112253. https://doi.org/10.1016/j.asoc.2024.112253.

[10] Hollman JH, Krause DA. Machine learning in admissions?: Use of Chi-Square Automatic Interaction Detection (CHAID) to predict matriculants to physical therapy school. J Allied Health 2023;52(3):e93–8.

Yoshiyasu Takefuji*

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

* Corresponding author at: Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan.

E-mail address: takefuji@keio.jp.