Atherosclerosis xxx (xxxx) xxx

Contents lists available at ScienceDirect

Atherosclerosis

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

Correspondence

# Unveiling feature importance biases in linear regression: Implications for protein-centric cardiovascular research

ARTICLE INFO

*Keywords:*
Feature importance
Linear regression
Cardiovascular health
Statistical techniques
Protein levels
Multicollinearity

*To the Editor,*

Garcia et al. conducted a comprehensive analysis of the impact of 27 covariates on protein levels, taking into account variables such as blood counts, cardiovascular risk factors, and lifestyle-related parameters [1]. In the methodology, they log-transformed laboratory parameters and pulse wave velocity (PWV) into more closely approximate normal distributions. They then investigated the associations between specific proteins (designated as the dependent variable) and the identified clinical and lifestyle parameters by employing both univariable and multivariable linear regression models with the lm function [1].

However, it is important to recognize the potential biases inherent in feature importances derived from linear regression models, leading to incorrect conclusions [2–6]. To accurately capture true associations, Garcia et al. should consider robust statistical methods, such as Chi-squared tests with *p*-values [7] and/or Spearman's correlation with *p*-values [8], which provide a more reliable framework to analyze relationships between variables. Common pitfalls of machine learning models, including linear regression, often stem from their specific nature, which can lead to skewed interpretations of feature importance and misrepresentations of associations between the target variable and the features.

This paper not only highlights the discrepancies in feature importance derived from linear regression [2–6] but also emphasizes the urgent need for more robust statistical techniques to establish valid relationships [7,8]. By advocating for true associations over biased feature importances, Garcia et al. should reconsider their conclusions to ensure more reliable outcomes. Ultimately, this approach enriches our understanding of the complex interplay between proteins and cardiovascular health, providing deeper insights into the mechanisms underlying cardiovascular conditions and directing future research efforts in this vital area.

Linear regression, while a widely used method for modeling relationships between variables, can induce biases in feature importance assessments for several reasons. Firstly, linear regression relies on the assumption that relationships between features and the target variable are linear. This can lead to misleading feature importance scores when the actual relationships are nonlinear. Nonlinearity can cause significant interactions between features that linear models fail to capture accurately, resulting in an overestimation or underestimation of a feature's significance.

Secondly, the presence of multicollinearity—where two or more features are highly correlated—can distort feature importance in linear regression. When multicollinearity is present, it becomes challenging to determine the individual contribution of correlated variables. As a result, the coefficients of these features can be inflated or deflated, leading to unreliable importance scores. This can mislead researchers and practitioners into overvaluing or undervaluing certain features.

Moreover, linear regression does not inherently account for feature interactions or complex interdependencies among multiple features. Without methods to identify and model these interactions, the model may simplify the relationships among variables, which could lead to biased importance rankings. Crucially, this simplification can mask the true drivers of the target variable, preventing a comprehensive understanding of the underlying mechanisms.

Lastly, the selection of features based on statistical significance without integrating domain knowledge can perpetuate biases. Researchers may inadvertently favor features that yield significant p-values in linear regression analysis, overlooking potentially important predictors that do not meet arbitrary significance thresholds. This selective inclusion again biases the feature importance assessments and can impede the formulation of valid conclusions.

In summary, biases in feature importance derived from linear regression stem from several factors, including its assumptions of linearity, the presence of multicollinearity, the inability to account for feature interactions, and a reliance on statistical significance without the incorporation of domain knowledge. Acknowledging these limitations is essential for researchers aiming to accurately interpret the relationships between features and outcomes. It is important to remember that the primary objective of machine learning is to make accurate predictions of the target variable; however, feature importances do not necessarily reflect true associations between the target and features, often serving

https://doi.org/10.1016/j.atherosclerosis.2024.119049
Received 29 October 2024; Received in revised form 5 November 2024; Accepted 7 November 2024
Available online 8 November 2024

merely as byproducts of the modeling process. This distinction underscores the need for more robust methodologies to derive genuine insights in predictive modeling.

## 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.

## References

[1] T. Garcia, A. Petrera, S.M. Hauck, et al., Relationship of proteins and subclinical cardiovascular traits in the population-based LIFE-Adult study, Atherosclerosis 398 (2024) 118613, https://doi.org/10.1016/j.atherosclerosis.2024.118613.

[2] J. Chen, L.Q.R. Ooi, T.W.K. Tan, et al., Relationship between prediction accuracy and feature importance reliability: an empirical and theoretical study, Neuroimage 274 (2023) 120115, https://doi.org/10.1016/j.neuroimage.2023.120115.

[3] J. Ochoteco Asensio, M. Verheijen, F. Caiment, Predicting missing proteomics values using machine learning: filling the gap using transcriptomics and other biological features, Comput. Struct. Biotechnol. J. 20 (2022) 2057–2069, https://doi.org/10.1016/j.csbj.2022.04.017.

[4] F. Mohtasham, M. Pourhoseingholi, S.S. Hashemi Nazari, et al., Comparative analysis of feature selection techniques for COVID-19 dataset, Sci. Rep. 14 (2024) 18627, https://doi.org/10.1038/s41598-024-69209-6.

[5] V.N. Dang, A. Cascarano, R.H. Mulder, et al., Fairness and bias correction in machine learning for depression prediction across four study populations, Sci. Rep. 14 (2024) 7848, https://doi.org/10.1038/s41598-024-58427-7.

[6] P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, Explainable AI: a review of machine learning interpretability methods, Entropy (Basel). 23 (1) (2020) 18, https://doi.org/10.3390/e23010018. Published 2020 Dec 25.

[7] J. Zeng, C. Zhou, Q. Yi, et al., Validation of the Rome severity classification of chronic obstructive pulmonary disease exacerbation: a multicenter cohort study, Int. J. Chronic Obstr. Pulm. Dis. 19 (2024) 193–204, https://doi.org/10.2147/COPD.S442382. Published 2024 Jan 17.

[8] E. Landfeldt, A. Alemán, S. Abner, et al., Predictors of loss of ambulation in duchenne muscular dystrophy: a systematic review and meta-analysis, J. Neuromuscul. Dis. 11 (3) (2024) 579–612, https://doi.org/10.3233/JND-230220.

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