World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:11, 2022

A Framework for Auditing Multilevel Models Using Explainability Methods

Authors : Debarati Bhaumik, Diptish Dey

Abstract : Multilevel models, increasingly deployed in industries such as insurance, food production, and entertainment within functions such as marketing and supply chain management, need to be transparent and ethical. Applications usually result in binary classification within groups or hierarchies based on a set of input features. Using open-source datasets, we demonstrate that popular explainability methods, such as SHAP and LIME, consistently underperform inaccuracy when interpreting these models. They fail to predict the order of feature importance, the magnitudes, and occasionally even the nature of the feature contribution (negative versus positive contribution to the outcome). Besides accuracy, the computational intractability of SHAP for binomial classification is a cause of concern. For transparent and ethical applications of these hierarchical statistical models, sound audit frameworks need to be developed. In this paper, we propose an audit framework for technical assessment of multilevel regression models focusing on three aspects: (i) model assumptions & statistical properties, (ii) model transparency using different explainability methods, and (iii) discrimination assessment. To this end, we undertake a quantitative approach and compare intrinsic model methods with SHAP and LIME. The framework comprises a shortlist of KPIs, such as PoCE (Percentage of Correct Explanations) and MDG (Mean Discriminatory Gap) per feature, for each of these three aspects. A traffic light risk assessment method is furthermore coupled to these KPIs. The audit framework will assist regulatory bodies in performing conformity assessments of AI systems using multilevel binomial classification models at businesses. It will also benefit businesses deploying multilevel models to be future-proof and aligned with the European Commission's proposed Regulation on Artificial Intelligence.

Keywords : audit, multilevel model, model transparency, model explainability, discrimination, ethics
Conference Title : ICAAIB 2022 : International Conference on Applications of Artificial Intelligence in Business
Conference Location : Amsterdam, Netherlands
Conference Dates : November 03-04, 2022