Balanced models from unbalanced data: an illustrative case in cardiovascular risk

B. de Otto[1], I. Pedrosa[1*] P. Quirós[1] and J. Pascual[1]

(1) CTIC Technology Centre. W3C LATAM Office host,
Ada Byron 39, Gijón, 33203, Asturias, Spain.

Abstract

While AI has the potential to mitigate the impact of biases, it also has the capacity to exacerbate the problem by integrating and implementing biases on a larger scale in critical domains, such as the healthcare domain. Therefore, fairness should be integrated into the training process itself [1]. Class imbalance in classification models is a common issue within clinical data analysis [2]. Rather than using the common approaches of undersampling or oversampling, identifying an appropriate threshold corrects the effect of class imbalance without disrupting the training sample. In this paper we consider the potential negative outcomes from poor data quality on healthcare and the effect when applying an innovative strategy to tackle class imbalance by replacing the standard 0.5 threshold. To reach this goal, we have chosen the threshold that minimises the difference between the sensitivity and specificity of the classifier because: 1) in the absence of expert medical knowledge that allows an informed assessment of the costs associated with each type of error, the chosen threshold assigns the same importance to false positives and false negatives, and 2) when sensitivity and specificity are very similar, the overall accuracy rate, located between the first two metrics, also takes a very close value to both, reducing the disparity between the most common performance metrics evaluation and facilitating the comparison between different specifications of the same model.This method has been applied to a model for predicting the risk of ischemic heart disease (i.e. heart attack and angina) in the US population using low-quality data from almost 400,000 people surveyed within the *Behavioral Risk Factor Surveillance System* survey. Positive class accounts approximately for 10% of the participants. The chosen strategy puts the models tested with different classification algorithms and specifications on an equal footing, reducing the disparity between the usual performance metrics for each single model. This simplifies the choice of the most suitable model to a comparison of the accuracy, sensitivity, and specificity rates on one hand, and the precision rate on the other. These results highlight both the relevance of defining appropriate techniques that deal with low-quality data, particularly in sensitive areas such as healthcare; and the potential negative consequences for clinical decision-making when unbalancing, biases or unfairness of the data are not addressed.

References

[1] Mehrabi, N., Morstatter, F., Saxena. N., Lerman, K., Galstyan. A. A Survey on Bias and Fairness in Machine Learning. *ACM Comput. Surv.* 54(6): Article 115, 2022. https://doi.org/10.1145/3457607

[2] Tasci E., Zhuge Y., Camphausen K., Krauze AV. Bias and Class Imbalance in Oncologic Data—Towards Inclusive and Transferrable AI in Large Scale Oncology Data Sets. *Cancers*. 14(12):2897, 2022. https://doi.org/10.3390/cancers14122897.

* Corresponding author ignacio.pedrosa@fundacionctic.org