

Marta de Antonio-Cuscó^{a,*}, Agustí Albalat-Torres^a,
Lucía Picazo Moreno^b, Javier Mateu-de Antonio^a

^a Pharmacy Department, Hospital del Mar, Barcelona, Spain

^b Critical Care Department, Hospital del Mar, Barcelona, Spain

* Corresponding author.

E-mail address: 61323@psmar.cat (M. de Antonio-Cuscó).

<https://doi.org/10.1016/j.medin.2024.08.001>

0210-5691/ Published by Elsevier España, S.L.U.

Comparisons are odious



Las comparaciones son odiosas

Dear Editor,

We would like to thank the authors of the article “*Sepsis mortality prediction with Machine Learning Techniques*”¹ for their valuable contribution to mortality prediction in septic patients using machine learning, with the aim of shedding new light on the heterogeneity of sepsis.

The authors aimed to evaluate machine learning models based on a local database and a public database (MIMIC-III).² After their analysis, they found that lactate levels, urine output, and acid–base balance variables were the most relevant for predicting mortality.

The authors identified similar predictive variables across both databases, with strong results at the local level but mediocre outcomes when compared with the other dataset. They argue that this discrepancy is due to the differences in variables used in each model or the reduction in the number of variables, which is a mathematically sound explanation. However, we believe that a more in-depth analysis of the article’s findings could provide greater insight, as medicine cannot always be fully explained through mathematical models.

- 1) COMPARED POPULATIONS: The data presented by the authors suggest that the two septic patient populations are markedly different. The local database shows an approximate mortality rate of 44%, while in MIMIC-III, the mortality rate is 16.25%. This difference in mortality likely reflects variations in patient populations, underlying pathologies, and healthcare systems. Therefore, we may be addressing different healthcare challenges under the same label. Mathematically, population heterogeneity affects the model’s generalizability.
- 2) METRICS USED: The metric employed to evaluate the models is the area under the curve (AUC), which may be appropriate for measuring the event of interest in the local population, where the event rate is close to 50%. However, AUC becomes a “less informative” metric in MIMIC-III, where the event of interest is imbalanced.³
- 3) VARIABLE IMPORTANCE: Furthermore, the metrics used to quantify variable importance—namely, the “*Mean Decrease Accuracy (MDA)/Gini (MDG)*” complex—are a “joint metric” that provides a more robust view of variable importance, as each captures different aspects⁴: MDA measures how the variable affects the overall predictive accuracy, while MDG measures how it impacts the quality of splits in the model. A closer examination of the values reveals that, although the highlighted variables

may be similar, the degree to which they improve prediction or reduce impurity differs quantitatively, implying a different classification capacities.

In summary, we would like to add to the discussion that not only the variables used explain the models’ reduced accuracy but also the consideration of socio-health factors, event distributions, and the quantification of variable importance should be addressed.

Funding

None.

Conflicts of interest

The authors declare that none of them have any conflicts of interest.

References

1. Pérez-Tome JC, Parrón-Carreño T, Castaño-Fernández AB, Nievas-Soriano BJ, Castro-Luna G. Sepsis mortality prediction with Machine Learning Techniques. *Med Intensiva* (Engl Ed). 2024;48:584–93, <http://dx.doi.org/10.1016/j.medicine.2024.05.009>.
2. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3:160035, <http://dx.doi.org/10.1038/sdata.2016.35>.
3. Jeni LA, Cohn JF, De La Torre F. Facing imbalanced data – recommendations for the use of performance metrics. In: 2013 humane association conference on affective computing and intelligent interaction. IEEE; 2013. p. 245–51, <http://dx.doi.org/10.1109/ACII.2013.47>.
4. Han H, Guo X, Yu H. Variable selection using mean decrease accuracy and mean decrease gini based on random forest. In: 2016 7th IEEE international conference on software engineering and service science (ICSESS). IEEE; 2016. p. 219–24, <http://dx.doi.org/10.1109/ICSESS.2016.7883053>.

Marcos Valiente Fernández*, Alejandro Caballo Manuel,
José Ginestal Calvo, Lidia Orejon García

Servicio de Medicina Intensiva, Hospital Universitario 12 de Octubre, Avenida de Córdoba, s/n, 28041 Madrid, Spain

* Corresponding author.

E-mail address: mvalientefernandez@gmail.com (M. Valiente Fernández).

<https://doi.org/10.1016/j.medin.2024.10.004>

0210-5691/ © 2024 Elsevier España, S.L.U. y SEMICYUC. Se reservan todos los derechos, incluidos los de minería de texto y datos, entrenamiento de IA y tecnologías similares.