Title

Reducing the overfitting in the gROC curve estimation

Abstract

The generalized receiver-operating characteristic, gROC, curve considers the classification ability of diagnostic tests when both larger and lower values of the marker are associated with higher probabilities of being positive. Its empirical estimation implies to select the best classification subsets among those satisfying particular condition. Both strong and weak consistency have already been proved. However, using the same data for both to select the classification subsets and to calculate its gROC curve leads to an over-optimistic estimate of the real performance of the diagnostic criteria on future samples. In this work, the bias of the empirical gROC curve estimator is explored through Monte Carlo simulations. Besides, two cross-validation based algorithms are proposed for reducing the overfitting. The practical application of the proposed algorithms is illustrated through the analysis of a real-world dataset. Simulation results suggest that the empirical gROC curve estimator returns optimistic approximations, especially, in situations in which the diagnostic capacity of the marker is poor and the sample size is small. The new proposed algorithms improve the estimation of the actual diagnostic test accuracy, and get almost unbiased gAUCs in most of the considered scenarios. However, the cross-validation based algorithms reported larger L1-errors than the standard empirical estimators, and increment the computational cost of the procedures. As online supplementary material, this manuscript includes an R function which wraps up the implemented routines.

Authors

Martínez-Camblor P.; Díaz-Coto S.

Author full names

Martínez-Camblor, Pablo (24462229000); Díaz-Coto, Susana (57194854806)

Author(s) ID

24462229000; 57194854806

Year

2024

Source title

Computational Statistics

Volume

39.0

Issue

2

Page start

1005

Page end

1022

Page count

17.0

DOI

10.1007/s00180-023-01344-6

Link

https://www.scopus.com/inward/record.uri?eid=2-s2.0-85149467932&doi=10.1007%2fs00180-023-01344-6&partnerID=40&md5=50d7850f162df68b4bba45cdf3b7d70a

Affiliations

Department of Anesthesiology, Geisel School of Medicine at Dartmouth, 7 Lebanon Street, Suite 309, Lebanon, 03751, NH, United States; Faculty of Health Sciences, Universidad Autonoma de Chile, Providencia, Chile; Department of Epidemiology, Geisel School of Medicine at Dartmouth, Lebanon, NH, United States

Authors with affiliations

Martínez-Camblor P., Department of Anesthesiology, Geisel School of Medicine at Dartmouth, 7 Lebanon Street, Suite 309, Lebanon, 03751, NH, United States, Faculty of Health Sciences, Universidad Autonoma de Chile, Providencia, Chile; Díaz-Coto S., Department of Epidemiology, Geisel School of Medicine at Dartmouth, Lebanon, NH, United States

Author Keywords

Binary classification problem; Cross-validation; Diagnostic problem; gROC curve; Overfitting

Funding Details

Ministerio de Ciencia e Innovación, MICINN

Funding Texts

This work was supported from the Grants GRUPIN AYUD/2021/50897 from the Asturies Government and PID2020-118101GB-I00 from Ministerio de Ciencia e Innovación (Spanish Government).

References

Airola A., Pahikkala T., Waegeman W., De Baets B., Salakoski T., An experimental comparison of cross-validation techniques for estimating the area under the ROC curve, Comput Stat Data Anal, 55, 4, pp. 1828-1844, (2011); Bantis L., Tsimikas J.,

Chambers G., Capello M., Hanash S., Feng Z., The length of the receiver operating characteristic curve and the two cutoff Youden index within a robust framework for discovery, evaluation, and cutoff estimation in biomarker studies involving improper receiver operating characteristic curves, Stat Med, 40, 7, pp. 1767-1789, (2021); Copas J.B., Corbett P., Overestimation of the receiver operating characteristic curve for logistic regression, Biometrika, 89, 2, pp. 315-331, (2002); Gong M., Studies on Classification Problems and Application in Consumer Lending Optimization System., (2022); Hanley J., McNeil B., The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology, 143, pp. 29-36, (1982); Krzanowski W., Hand D., ROC curves for continuous data, (2009); Lisans-Tezi Y., Belirteç değerlerinin monoton olmamasi durumunda generlleştirilmiş ROC eğrilerinin parametrik ve parametrik olmayan yöntemlerle kestirilmesi ve en iyi kesim noktalarinin saptanmas, In T.C. Hacettepe Üniversitesi Sağlik Bilimleri Enstitüsü, (2022); Lusted L., Signal detectability and medical decision-making, Science, 171, 3977, pp. 1217-1219, (1971); Ma H., Bandos A., Gur D., Informativeness of diagnostic marker values and the impact of data grouping, Comput Stat Data Anal, 117, pp. 76-89, (2018); Martinez-Camblor P., Corral N., Rey C., Pascual J., Cernuda-Morollon E., Receiver operating characteristic curve generalization for non-monotone relationships, Stat Methods Med Res, 26, 1, pp. 113-123, (2017); Martinez-Camblor P., Pardo-Fernandez J., Parametric estimates for the receiver operating characteristic curve generalization for non-monotone relationships, Stat Methods Med Res, 28, 7, pp. 2032-2048, (2019); Martinez-Camblor P., Perez-Fernandez S., Diaz-Coto S., Improving the biomarker diagnostic capacity via functional transformations, J Appl Stat, 46, 9, pp. 1550-1566, (2019); Martinez-Camblor P., Perez-Fernandez S., Diaz-Coto S., The area under the generalized receiver-operating characteristic curve, Int J Biostat, 18, (2021); McIntosh M.W., Pepe M.S., Combining several screening tests: optimality of the risk score, Biometrics, 58, 3, pp. 657-664, (2002); Montoya-Perez I.A., Airola P.J.B.,

Jambor I., Pahikkala T., Tournament leave-pair-out cross-validation for receiver operating characteristic analysis, Stat Methods Med Res, 28, 10-11, pp. 2975-2991, (2019); Parodi S., Pistoia V., Muselli M., Not proper ROC curves as new tool for the analysis of differentially expressed genes in microarray experiments, J Math Psychol, 9, 410, pp. 1-30, (2008); Pepe M., The statistical evaluation of medical tests for classification and prediction, (2003); Perez-Fernandez S., Martinez-Camblor P., Filzmoser P., Corral N., Visualizing the decision rules behind the ROC curves: understanding the classification process, AStA Adv Stat Anal, 105, (2021); Rutter C., Gatsonis C., A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations, Stat Med, 20, 19, pp. 2865-2884, (2001); Leave-One-Out Cross-Validation, pp. 600-601, (2010); Shah E., Pelletier E., Greeley C., Sieglinger E., Sanchez J., Northam K., Perrone J., Curley M., Navas C., Ostler T., Burnett Greeley A., Martinez-Camblor P., Baker J., Harris A., Siegel C., Chey W., An office-based, point-of-care test predicts treatment outcomes with community-based pelvic floor physical therapy in patients with chronic constipation, Clin Gastroenterol Hepatol, (2022); Spanos A., Harrell F., Durack D., Differential diagnosis of acute meningitis: an analysis of the predictive value of initial observations, J Am Med Assoc, 262, 19, pp. 2700-2707, (1989); Vanda I., Rodriguez-Alvarez M., Gayoso-Diz P., Statistical evaluation of medical tests, Ann Rev Stat Appl, 8, 1, pp. 41-67, (2021); Zhou X., Obuchowski N., Clish D., Statistical methods in diagnostic medicine, (2002)

Correspondence Address

P. Martínez-Camblor; Department of Anesthesiology, Geisel School of Medicine at Dartmouth, Lebanon, 7 Lebanon Street, Suite 309, 03751, United States; email: Pablo.Martinez-Camblor@hitchcock.org

Publisher

Springer Science and Business Media Deutschland GmbH

ISSN

09434062

Language of Original Document

English

Abbreviated Source Title

Comput. Stat.

Document Type

Article

Publication Stage

Final

Source

Scopus

EID

2-s2.0-85149467932