ROC curves in cost space

José Hernández-Orallo · Peter Flach · César Ferri

Abstract ROC curves and cost curves are two popular ways of visualising classifier performance, finding appropriate thresholds according to the operating condition, and deriving useful aggregated measures such as the area under the ROC curve (*AUC*) or the area under the optimal cost curve. In this paper we present new findings and connections between ROC space and cost space. In particular, we show that ROC curves can be transferred to cost space by means of a very natural threshold choice method, which sets the decision threshold such that the proportion of positive predictions equals the operating condition. We call these new curves *rate-driven curves*, and we demonstrate that the expected loss as measured by the area under these curves is linearly related to *AUC*. We show that the rate-driven curves are the genuine equivalent of ROC curves in cost space, establishing a point-point rather than a point-line correspondence. Furthermore, a decomposition of the rate-driven curves is introduced which separates the loss due to the threshold choice method from the ranking loss (Kendall τ distance). We also derive the corresponding curve to the ROC convex hull in cost space; this curve is different from the lower envelope of the cost lines, as the latter assumes only optimal thresholds are chosen.

Keywords Cost curves · ROC curves · Cost-sensitive evaluation · Ranking performance · Operating condition · Kendall tau distance · Area Under the ROC Curve (*AUC*)

J. Hernández-Orallo · C. Ferri (✉)
Departament de Sistemes Informàtics i Computació, Universitat Politècnica de València, València, Spain
e-mail: cferri@dsic.upv.es

J. Hernández-Orallo
e-mail: jorallo@dsic.upv.es

P. Flach
Department of Computer Science, University of Bristol, Bristol, UK
e-mail: peter.flach@bristol.ac.uk