Mach Learn (2013) 92:479–502
DOI 10.1007/s10994-013-5349-4

Multi-stage classifier design

Kirill Trapeznikov · Venkatesh Saligrama · David Castañón

Received: 9 January 2013 / Accepted: 2 April 2013 / Published online: 14 May 2013

Abstract In many classification systems, sensing modalities have different acquisition costs. It is often *unnecessary* to use every modality to classify a majority of examples. We study a multi-stage system in a prediction time cost reduction setting, where the full data is available for training, but for a test example, measurements in a new modality can be acquired at each stage for an additional cost. We seek decision rules to reduce the average measurement acquisition cost. We formulate an empirical risk minimization problem (ERM) for a multi-stage reject classifier, wherein the stage k classifier either classifies a sample using only the measurements acquired so far or rejects it to the next stage where more attributes can be acquired for a cost. If we restrict ourselves to binary classification setting then, to solve the ERM problem, we show that the optimal reject classifier at each stage is a combination of two binary classifiers, one biased towards positive examples and the other biased towards negative examples. We use this parameterization to construct stage-by-stage global surrogate risk, develop an iterative algorithm in the boosting framework and present convergence and generalization results. We test our work on synthetic, medical and explosives detection datasets. Our results demonstrate that substantial cost reduction without a significant sacrifice in accuracy is achievable.

Keywords Multi-stage classification · Sequential decision · Boosting · Cost sensitive learning

Editors: Zhi-Hua Zhou, Wee Sun Lee, Steven Hoi, Wray Buntine, and Hiroshi Motoda.

K. Trapeznikov (✉) · V. Saligrama · D. Castañón
8 Saint Mary's Street, Boston, MA 02215, USA
e-mail: ktrap@bu.edu

V. Saligrama
e-mail: srv@bu.edu

D. Castañón
e-mail: dac@bu.edu